UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Statistics of High-Throughput Characterization of Microbial Interactions

Permalink
https://escholarship.org/uc/item/8900x1s7

Author
Krinsman, William Edward

Publication Date
2022

Peer reviewed|Thesis/dissertation

Statistics of High-Throughput Characterization of Microbial Interactions

by

William Edward Krinsman

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biostatistics

and the Designated Emphasis

in

Computational and Data Science and Engineering

and the Designated Emphasis

in

Computational and Genomic Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mark van der Laan, Chair
Professor Adam Arkin
Professor Lexin Li

Spring 2022

Statistics of High-Throughput Characterization of Microbial Interactions

Abstract

Statistics of High-Throughput Characterization of Microbial Interactions

by

William Edward Krinsman

Doctor of Philosophy in Biostatistics

and the Designated Emphasis in

Computational and Data Science and Engineering

and the Designated Emphasis in

Computational and Genomic Biology

University of California, Berkeley

Professor Mark van der Laan, Chair

An active area of research interest is the inference of ecological models of complex microbial communities. Inferring such ecological models entails understanding the interactions between microbes and how they affect each other's growth. This dissertation employs a statistical perspective to contribute further to the knowledge currently addressing this problem.

Part I   Part I explains how high-throughput droplet-based microfluidics technology can be used to screen for microbial interactions. An explicit, statistical framework is motivated and developed that can guide the analysis of data from such experiments. Chapter 1 investigates the specific questions that need to be answered to study microbial interactions. Chapter 1 explains why high-throughput droplet-based microfluidics technology overcomes previous limitations to answering these questions. It is shown how answering these questions can be recast as the statistical problem of estimating a network with mixed-sign edge weights. Chapter 2 investigates how to approach these questions using statistical models. Chapter 2 explains that the data from the noisy dynamical systems corresponding to each droplet can be understood as censored observations of a multivariate Markov process. The statistical understanding of a droplet's initial state is identified as crucial to overcoming the main limitation of these experiments, the uncontrolled assignment of microbes to droplets.

Part II Part II explains how it might be possible to predict, based on the experimental setup, how much data will be produced to infer given microbial interactions. Running the experiment once without incubating the droplets turns out to be necessary to make such predictions. Chapter 3 investigates which statistical (working) models can be used to describe a droplet's initial state. Chapter 3 shows that specific assumptions justify a default working model. New working models are derived by relaxing each of these assumptions. Chapter 4 investigates whether the failure of any of these assumptions leads to substantially new behavior. Chapter 4 demonstrates that log likelihood ratios can be used to answer this question. Failure of the sampling without replacement assumption turns out to have negligible effects in practice, but failures of the other assumptions could be important. Chapter 5 investigates how failures of the relevant assumptions affect the targeted estimands that enable the prediction of how much data will be produced to infer given microbial interactions. Chapter 5 confirms that more severe failures of these assumptions lead to more severe discrepancies with the predictions derived from the default working model. The nature of the effect depends on the chosen grouping of droplets defining the targeted estimands. Chapter 6 investigates how to estimate failures of the relevant assumptions from the data produced by unincubated droplets. Chapter 6 presents both plugin and maximum likelihood estimators for doing so. Failures of these assumptions are shown to be understandable non-parametrically.

Part III Part III demonstrates the feasibility of inferring microbial interactions from the data produced by these experiments. Relevant ideas from the microbiological and ecological literature are recast into an explicit, statistical framework. Chapter 7 investigates how a particular measure of relative fitness can be recast into the statistical framework of average treatment effects. Chapter 7 explains how violations of positivity assumptions are inevitable for this problem, making controlling for confounding difficult. Explicit assumptions are given under which the estimands are identifiable from the observed data produced by incubated droplets, even though initial states of the droplets are not directly observed. Chapter 8 investigates how comparisons of ecological interactions can be recast into the statistical framework of loss functions for signed networks. Chapter 8 explains how avoiding unexpected behavior requires loss functions for signed networks to satisfy what is called herein "the double penalization principle". Starting from loss functions of unsigned networks, several examples of loss functions for signed networks are derived that satisfy this property.

Conclusion Future work will explore further choices that can be made when modelling this problem, how to connect these ideas to more sophisticated statistical methodologies, and the biological interpretation of the results of applying these ideas to real experimental datasets. This work demonstrates the plausibility of characterizing microbial interactions using high-throughput droplet-based microfluidics technologies, and hopefully will guide the analysis of data produced by such experiments in the future.

For Falcon and Ruby. We will always miss you.

Contents

Code

Relevant source code and Jupyter notebooks can be found in several public archives.

GitLab: https://gitlab.com/krinsman/dissertation

GitHub: https://github.com/krinsman/dissertation

Zenodo: https://doi.org/10.5281/zenodo.6539845

FigShare: https://doi.org/10.6084/m9.figshare.19747609

Excerpted source code and Jupyter notebooks relevant for Part II can be found in the relevant GitLab repository at https://gitlab.com/krinsman/droplets.

Excerpted source code and Jupyter notebooks relevant for chapter 8 can be found in the relevant GitLab repository at https://gitlab.com/krinsman/mixed-sign-networks.

List of Figures

List of Tables

Notational Conventions

In most cases bound variables will be denoted by lower-case Latin letters, and their corresponding "maximum possible value" (which is a free variable) will be denoted by the corresponding Latin letter. For example, $s \in [S]$, or $d \in [D]$. In cases where the free variable will later be used to define the bound variable for another expression, the free variable will usually be denoted with a lower-case Latin letter (e.g. s), and the bound variable will be denoted with the corresponding lower-case Greek letter (e.g. σ).

Deterministic scalars are generally denoted with lower-case, regular-font variable names. Random scalars are generally denoted with upper-case, regular-font variable names. Deterministic vectors are generally denoted with lower-case, bold-face, and arrow-topped variable names. Random vectors are generally denoted with upper-case, bold-face, and arrow-topped variable names. (Deterministic) matrices are generally denoted with upper-case, bold-face variable names (*without* arrows). Random matrices do not have a notational convention.

$\mathbb{N}$ is the natural numbers $\{1, 2, 3, \dots\}$, $\mathbb{N}_{\geq 0}$ is the whole numbers $\{0, 1, 2, 3, \dots\}$, $\mathbb{R}$ is the real numbers, and $\mathbb{R}_{\geq 0}$ is the non-negative real numbers $[0, \infty)$.

Given a positive integer $N \in \mathbb{N}$, define $[N] := \{1, \dots, n, \dots, N\}$. In general, the entries of a (non-negative) vector $\vec{\mathbf{v}} \in \mathbb{R}^S_{\geq 0}$ are denoted $v^{(s)}$, i.e.

$$\vec{\mathbf{v}} \stackrel{def}{=} \left(v^{(1)}, \dots, v^{(s)}, \dots, v^{(S)} \right).$$

The shorthand $v := \sum_{s=1}^S v^{(s)}$ is used for the sum of the entries of $\vec{\mathbf{v}}$.

Acknowledgments

Dr. Fangchao Song for suggesting the topic of the dissertation and generous consultations, feedback, and advice throughout.

Dr. Lauren M. Lui for invaluable recommendations of strategies and resources for scientific writing.

Professor Perry de Valpine for the useful suggestion of investigating the Dirichlet-Multinomial and related models for weakening the 'well-mixed' droplets assumption.

Part I

Introduction

The ideas explored in Part I belong to the general fields of both:

(1) using droplet-based microfluidics technology for high-throughput biological assays, and

(2) quantitatively comparing the growth of microbes under different conditions.

Droplet-based microfluidics technology has already been used for several different kinds of high-throughput biological assays. See e.g. [KvVDH10] or [GRHW12] for reviews of such studies. However, droplet-based microfluidics technology has been particularly successful for the analysis of single cells. See e.g. [LE13] or [MRH20] for reviews of such studies.

Experiments quantitatively comparing the growth of microbes under different conditions are often called "competition assays" [WL15]. At a high level, one can think of there being roughly two different kinds of competition assays. I describe both kinds below.

Competition assays where the treatment conditions correspond to mutant strains of a microbe, and the control condition is the "wild type" of the microbe, are very popular. Often the experiment is augmented by comparing the growth of mutant strains and wild types not just under one set of external conditions, but repeatedly under multiple sets of external conditions. The studies [WPW+15], [HRM+19], and [MKBDB00] are some (possibly non-representative) examples. This kind of competition assay can be thought of as a "bipartite interaction network" [CK21] [WBGLR+21] [TKA20], where there is a clear distinction between "predictor variable nodes" (the sets of external conditions, "environmental variable nodes") and "response variable nodes" (the numbers of individuals for the various mutant strains). This kind of competition assay corresponds to a "gene × environment interactions problem" in the framework described in item (4) of section 2.1.1.

In contrast, there are also competition assays where the treatment conditions correspond to co-cultures of different types of microbes, and the control conditions are monocultures of the types of microbes. (See [GFP14] for a review of methodologies for co-culturing microbes.) Studies such as [VCF+18] that are based on this kind of competition assay allow us to study how different microbes affect each other's growth. The studies [SS19] and [BEG+18] are other (possibly non-representative) examples. This kind of competition assay can be thought as an "unpartitioned interaction network" [AML+17], where there is no distinction between "predictor variable nodes" and "response variable nodes" (both being the numbers of individuals for the types of microbes). This kind of competition assay corresponds to a "gene × gene interactions problem" in the framework described in item (4) of section 2.1.1.

The dichotomy described above may not always be entirely clear-cut in practice. For example, the studies [LYS+19] and [NW18] describe competition assays where the external conditions are the presence of microbes different from those whose growth is being measured.

Inferring ecological models of microbial communities entails understanding the interactions between microbes and how they affect each other's growth. Therefore high-throughput competition assays using droplet-based microfluidics could potentially allow us to infer ecological models of complex microbial communities consisting of numerous types of microbes. Hence the specific focus of Part I is on the new subfield of using high-throughput droplet-based microfluidics assays to study how different microbes affect each other's growth. Pre-

vious studies in this subfield, e.g. [KKO+19] and [HCT+19], had important limitations. (Cf. the discussion in sections 1.3 and 1.4.3.) The new MOREI platform (cf. section 1.4) aims to overcome most of these limitations. In doing so, the form of the resulting data is substantially different from that produced by previous experimental platforms. Thus the specific problem of this dissertation is to identify and address the data analysis challenges for the MOREI platform. This includes those data analysis challenges that are also applicable to more general experimental platforms but that are insufficiently addressed by previous work.

Chapter 1 The broader field that chapter 1 belongs to is the study of ecological models of microbial interactions. In particular, chapter 1 explores the subfield of how droplet-based microfluidics technology can be used to infer microbial interactions. This chapter begins its exploration by investigating the specific questions that need to be answered to study microbial interactions when using any technology. Then the chapter follows this by explaining why high-throughput droplet-based microfluidics technology may overcome previous limitations to answering these questions. The chapter proceeds further by describing how the problem of inferring ecological models of microbial interactions can be recast as a special case of the statistical problem of estimating a network with mixed-sign edge weights. Finally, the chapter overviews of the details of the particular experiment considered by this dissertation.

Chapter 2 Hypothetically, the ideas of chapter 2 belong to a broader field of "*statistical inference of (time-evolving) interactions between variables of a (noisy) dynamical system for which the temporal evolution may be incompletely observed*". However, such a concrete field of research appears not to exist[1] in the literature. In practice, the relevant ideas are spread between numerous disparate fields that are currently unconnected to each other. See section 2.1 for a further elaboration. Hence any of those disparate broader fields may be chosen, according to one's preferences, as the general context within which the ideas of chapter 2 may be understood. Starting from any such general context, chapter 2 then focuses on the important issue of statistical inference of (time-dependent) microbial interactions based on pseudo-longitudinal data. In particular, chapter 2 investigates plausible statistical models that could be used to describe the data produced by this experiment. This chapter explains that the data from the noisy dynamical systems corresponding to each droplet can be understood as censored observations of a multivariate Markov process. The statistical understanding of a droplet's initial state is identified as crucial to overcoming the main limitation of these experiments, the uncontrolled assignment of microbes to droplets.

---

[1] If it does exist, hopefully it has a name that is pithier and catchier than the phrase written above.

Chapter 1

Biological Context of Experiment

I show how inferring an ecological model of a microbial community can be understood as a special case of the statistical problem of estimating a network with mixed-sign edge weights. See section 1.2. Herein I explain why high-throughput droplet-based microfluidics technology overcomes previous limitations to studying microbial interactions. See section 1.3. Finally, I explain why and how the data-generating process for the experiment motivating this dissertation can be divided into (at least) three phases. See section 1.4.

Section 1.1 clarifies some issues motivating the scientific questions this dissertation helps to address. Section 1.2 explains how these scientific questions are simplified to a statistical questions for the purposes of this work. Section 1.3 discusses the kinds of data necessary for answering such questions in general. Finally, section 1.4 describes the particular experiment being used to answer these questions and specific features of the data it generates.

Chapter 1 Contents

1.1   Why and What of Microbial Ecology

Microbial ecology is the study of microbial interactions with other microbes, with non-microbial organisms, and with the environment.

A central assumption of this thesis is that microbial interactions are interesting and therefore that characterizing microbial interactions is valuable. While there are many reasons why that is the case, many more than I can list here, I want to share with the unfamiliar reader some such reasons why one might be interested in microbial ecology. There are also reviews discussing this, cf. e.g. [FR12], [GFP14], or [AGH⁺17]. Readers already familiar with or interested in microbial ecology might consider skipping the rest of this section.

Why   One general reason to study microbial interactions is advancing fundamental biological knowledge. Microbial interactions have greatly shaped the formation and development of (macroscopic) ecosystems throughout deep time. Microbial interactions crucially influence biogeochemical nutrient cycles and thus Earth's climate[FFD08], for example via the marine carbon pump or the oxygenation of Earth's atmosphere and oceans billions of years ago.

Microbial interactions also greatly shaped selective pressures that favored many key evolutionary innovations. Examples include multicellularity, endosymbiosis (including eukaryogenesis), pathogenicity, and antibiotic resistance.

The above two categories are not mutually exclusive. Microbial interactions impact the fitness of macroscopic organisms [BML⁺15] and can lead to key evolutionary innovations that greatly shape macroscopic ecosystems. For example, microbial associations are also important for plants' being "ecosystem engineers" and for animals to exploit new food sources.

Another general reason to study microbial interactions is the potential to develop new technological applications that benefit humanity [GFP14] [AML19]. Via their influence on non-microscopic organisms and biogeochemical nutrient cycles, understanding microbial interactions helps to develop new techniques in agriculture [MTFH16], medicine and health [PM17], climate modelling and protection of the environment [WL02] [GWV⁺11] [YSP⁺19], and astrobiology and human space colonization [LPR19], among other areas.

From a purely anthropocentric point of view, understanding the role of microbial interactions in the evolution of life on Earth also means better understanding our origins.

Microbial interactions with other microbes have historically been the most difficult to study because all participants in the interactions are microscopic. The great diversity and number of microbes leads to a great diversity and complexity of microbial interactions with other microbes. It is arguably more important to understand microbial interactions with other microbes than to understand microbial interactions with non-microbial organisms and the environment. The former interactions mediate the latter, while also being more diverse, more complex, and more understudied. Herein, unless specified otherwise, "microbial interactions" refers to microbial interactions with other microbes.

What   The general field that chapter 1 belongs to is the study of ecological models of microbial interactions. The review [FR12] describes how such models are often formulated as pairwise interaction networks, an approach also espoused elsewhere, cf. e.g. [AML19] or [AML⁺17]. (Although see e.g. [MXS17] or [BKK16] regarding potential limitations of models that consider only pairwise interactions.) Ecological interactions have been classified

at least since [Lid79] in terms of signs ($\pm$) describing how organisms affect each other's growth. The review [GFP14] describes several technologies that can be used for studying how various microbes affect the growth of other microbes with their presence. See also [VCF$^+$18] for an example. The notion of a network describing ecological interactions, and ecological interactions being described by ($\pm$) signs, leads naturally to the notion of a signed network, something that has been discussed (in other contexts) at least since [Har53]. The signed values of the edges can also be interpreted as the coefficients of some (parametric) dynamical system, see e.g. [GCLF18] for a review or either of [GHG$^+$18] and [XAF$^+$17] for discussion. Cf. the introduction to Part I. Chapter 1 explores the subfield of how droplet-based microfluidics technology can be used to infer microbial interactions. Published studies in this subfield include [KKO$^+$19], [HCT$^+$19], and [KOK$^+$20]. In this chapter I investigate the specific questions that need to be answered to study microbial interactions.

1.2   Modelling Microbial Interactions

We need both (1) an estimand to estimate and (2) data from which to make estimates. Then a (parametric) statistical model posits how the estimand specifies the data generating distribution. In this section I loosely describe the estimand to estimate for characterizing microbial interactions. Later sections discuss the data used for making estimates.

We can use a network to characterize the pairwise interactions of a microbial community (cf. [FR12]). The approach to do this used herein is as follows. Each of the S nodes of the network corresponds uniquely to one of the S strains[1], where S denotes the number of strains in the microbial community. An edge is directed from (the node for) strain s_1 to (the node for) strain s_2 if strain s_1 affects the growth of strain s_2. If strain s_1 promotes the growth of strain s_2 then the weight of the edge is positive, whereas if strain s_1 suppresses the growth of strain s_2 then the weight of the edge is negative. The estimand for characterizing microbial interactions is a directed network with mixed-sign edge weights [Har53] [Lid79] [GHG$^+$18], whose number of nodes is the same as the number of strains. A "microbial interaction" is operationally defined as one strain affecting the growth of another strain. Cf. figure 1.1.

---

[1]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.
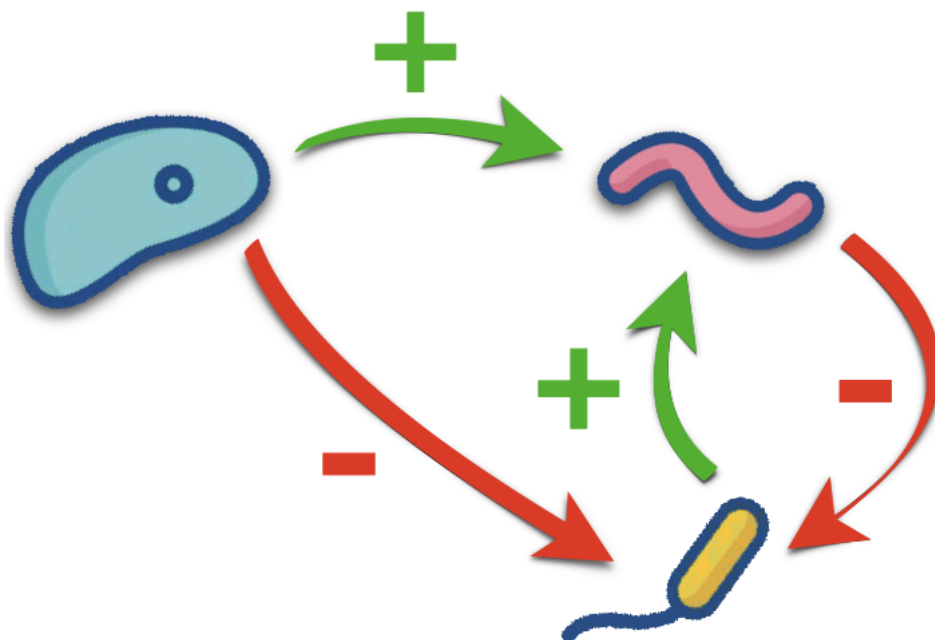
Figure 1.1: A directed network with mixed-sign edge weights that represents the interactions within a given microbial community containing three strains. Arrows on the edges indicate the directions of the interactions. Green indicates positive edge weights (interaction promotes growth). Red indicates negative edge weights (interaction suppresses growth).

Herein I only consider how to characterize pairwise microbial interactions. There are a panoply of formalisms generalizing networks to objects that describe "higher order interactions". (Hypergraphs, simplicial complexes, multilayered networks, etc.) Choosing among them requires us to be much more precise about the notion of "higher order" microbial interaction in which we are interested. Moreover, compared to pairwise interactions, there is usually less data available to infer higher order interactions[2]. Thus both aspects of formulating "higher order" microbial interactions as a statistical estimation problem are substantially more difficult. Hence these tasks are left to future work. Nevertheless, failing to model or account for the effects of "higher order interactions" might limit the usefulness or validity of any ecological interpretations of our results; cf. [BKK16] or [MXS17] for arguments along these lines. For a contrasting viewpoint, see [AML19] which (using simplifying assumptions) argues that it is sufficient to characterize only the pairwise interactions of a microbial community to be able to manipulate it (e.g. for biotechnological applications, cf. [GFP14]).

In chapter 8 I discuss methods for comparing two directed networks with mixed-sign edge weights. This allows us to quantify how similar an estimated microbial interaction network is to the true microbial interaction network.

---

[2] Cf. the discussions in sections 2.3.5 and 7.7.3, which both make arguments to this effect.

1.3   How of Microbial Ecology

In section 1.2 above I described the estimand to estimate for characterizing microbial interactions. Now I will discuss the data used for making estimates.

As described in section 1.2, herein "microbial interactions" are operationally defined as one strain affecting the growth of another strain, and we only characterize pairwise microbial interactions. Therefore to characterize microbial interactions, we want to quantify for all (ordered) pairs of microbial strains the effect that one strain has on the other strain's growth[3].

While there are many approaches that have been used for this problem, herein I focus exclusively on controlled laboratory studies. Biology is immensely complicated with many "unknown unknowns", and ecology even more so. Study designs for this problem based on observational data are exposed to many more potentially uncontrolled confounders. Therefore the conclusions from study designs based on observational data are much less reliable[4].

Another reason to not use observational data (e.g. metagenomic studies) for *attempting* to infer microbial interactions is that these studies are problematic [MMV$^+$08] [BBI20] inasmuch as they usually only produce "compositional data" consisting of relative abundances, see section 1.4.3. Cf. e.g. [CHL$^+$20], [SBT$^+$13], [GHG$^+$18], or [MBH$^+$14] for counterpoints.

To infer the effect of strain s_1 on the growth of strain s_2, we can compare the growth of strain s_2 when strains s_1 and s_2 co-occur with the growth of strain s_2 when strains s_1 and s_2 do *not* co-occur. If growth under both conditions occurs in a controlled laboratory setting, we can conclude that any difference in growth of strain s_2 between the two conditions is most likely the effect of strain s_1. The condition where strain s_2 co-occurs with strain s_1 can be thought of as the "treatment", and the condition where strain s_2 does *not* co-occur with strain s_1 can be thought of as the "control". Cf. figure 1.2 or the discussion of "competition assays" from the introduction to Part I. See also [GFP14] for a (perhaps now somewhat dated) review of such "co-culture" systems for studying cellular interactions.

---

[3] Note that inferring self-interactions is a problem. The "competition assay" framework described below provides no obvious way to do that. A priori it's unclear whether self-interactions are actually identifiable from such observations without more information (or assumptions). Previous work has either assumed values for self-interactions (e.g. [XAF$^+$17]) or used models that require longitudinal data (e.g. [VCF$^+$18] or [SBT$^+$13]). See [GCLF18] for an overview. Preliminary results (data not shown) suggested that attempting to modify methodology similar to that used in [VCF$^+$18] [SBT$^+$13] [MBH$^+$14] [GHG$^+$18] [FM14] [MMV$^+$08] to not require longitudinal data, and apply to pseudo-longitudinal data, leads to very poor estimates of all of the interactions, including the self-interactions. (The attempted modification is described in detail in appendix A.3.) Other approaches that are unable to estimate self-interactions led to much better results (data not shown) for the remaining interactions. Addressing this shortcoming is left to future work.

[4] One argument for observational data is that conclusions based on the synthetic microbial communities ("SynComs") used in controlled laboratory studies might not transfer to naturally occurring systems. Thus it is useful for SynComs to be as realistic as possible, e.g. vis a vis the number of strains in the community.
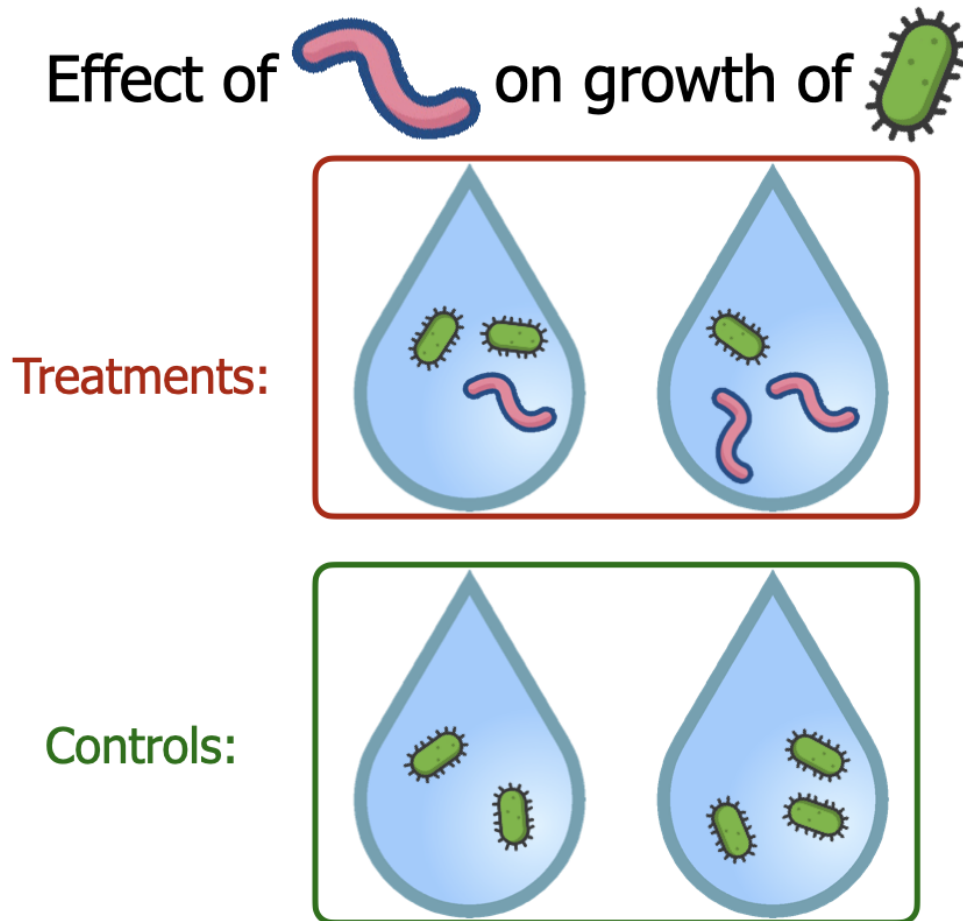
Figure 1.2: Inferring the effect of one strain on the growth of another strain. Cf. the discussion of "competition assays" from the introduction to Part I.

As before, let S denote the total number of strains in the microbial community. Then to implement this framework for all (ordered) pairs of strains, we need at least $\binom{S}{2} + S$ replicates. $\binom{S}{2}$ of the replicates correspond to "treatments" where two distinct strains co-occur, while S of the replicates correspond to "controls" where a single strain does not co-occur with other strains. Therefore, to characterize all pairwise microbial interactions, a lower bound on the number of required replicates grows quadratically with the total number of strains[5].

_____

[5] Conceivably the only way we could escape quadratic scaling for this problem is with a "sparsity assumption". Otherwise this appears to be an "inherently quadratic" problem. More specifically, to escape quadratic scaling it seems we would need some a priori biological knowledge that only a subquadratic number of microbial interactions can actually exist. (The other interactions would be "zero", i.e. correspond to missing edges in the microbial interaction network.) That would allow us to get away with having a subquadratic number of replicates. One conceivable way this could occur is if the microbial community was structured into "subcommunities" with minimal overlaps, such that pairwise interactions could only occur for two strains within the same subcommunity. Taking the number of pairwise interactions within

The most controlled way to generate these replicates is to manually plate the cells from the aforementioned $\binom{S}{2} + S$ combinations of strains. However such an approach is fairly labor intensive and does not scale well with the number of plates. Therefore the approach becomes practically infeasible for even relatively small numbers of strains. For example, the current state of the art as performed in e.g. [VCF+18] is not much larger than 10 strains.

Therefore to characterize all pairwise microbial interactions in a community with a reasonably large number of strains, we need to substantially increase the throughput of replicates. One way to do this is using droplet-based microfluidics technology. Using droplet-based microfluidics, the number of replicates (which corresponds to the number of microfluidic droplets containing cells) can be in the hundreds of thousands or even millions. See any of [TLHL08] [KvVDH10] [GRHW12] [ZW17] [SCZ17] [SkM20], sorted in roughly chronological order, for a review of droplet-based microfluidics technology[6].

Droplet-based microfluidics has already proven useful for single-cell, single-type experiments. See, for example, [LE13] or [MRH20], for a review of single-cell studies. New experimental platforms, such as kChip [KKO+19] [KOK+20] or MINI-Drop [HCT+19], use droplet microfluidics to probe the interactions of multiple types of cells. For kChip and MINI-Drop specifically the motivation is to enable more systematic explorations of microbial interactions. Microbial ecology would benefit from experimental, rather than merely observational, data characterizing the interactions of large numbers of microbial strains. The current state of the art for such experiments, by manually culturing and plating combinations of cell types, probes the interactions of fewer than 20 microbial strains simultaneously [VCF+18]. Because droplet microfluidics devices produce large numbers of microfluidic droplets, multi-cell, multi-type droplet microfluidics experimental platforms potentially could drastically increase the number of microbial strains whose interactions can be probed simultaneously. This has not yet happened with current droplet-based microfluidics studies of microbial interactions, which were severely limited in the number of strains considered [HCT+19] [KKO+19] [KOK+20] due (in part) to requiring fluorescently labelled strains. Cf. section 1.4.3.

A major tradeoff associated with this approach, compared to manually culturing and plating combinations of cell types, is that droplet microfluidics chips form microfluidic droplets randomly. Thus compared to manually plating cells we lose some control over each individual replicate. Cf. figure 3.1. Therefore, even starting with known microbial communities, it is a priori unclear approximately how many droplets will be produced containing any given combination of cell types. Filling this gap will increase the relative advantage of multi-cell,

---

the largest subcommunity as a constant, it follows that the scaling would be roughly linear in the number of subcommunities. The scaling following from this "divide and conquer" approach could in practice be much more favorable than the quadratic scaling in the total number of strains. That being said, we would need biological justification for why either (1) no pairwise microbial interactions occur between strains from different subcommunities, or (2) why we don't care about pairwise microbial interactions between strains from different subcommunities and are content to consider the microbial interaction *sub*network where all of the edges corresponding to these interactions are missing.

[6] Droplet-based microfluidics is a special case of a wider variety of microfluidics technologies developed in recent decades. See any of [Whi06] [TF08] [SFB14] [CdC+17] [CG19], sorted in roughly chronological order, for a review of general microfluidics technology.

multi-type droplet microfluidics experimental platforms over manually culturing and plating combinations of cell types. This will help ensure that the drastic increase in throughput in the number of replicates more than compensates for any loss of control. Cf. chapters 3 and 5. Hopefully this can abet future dramatic increases in the state of the art for the number of microbial strains whose interactions can be probed simultaneously. Being able to characterize all of the interactions in more realistically sized microbial communities under controlled laboratory settings could substantially advance the field of microbial ecology.

1.4   Description of Experiment

MOREI (pronounced "more-ray") stands for "More Interactions" or (treating MORE itself as a backronym) "Microfluidics Offers Replicate Experiments for Interactions". Starting from a microbial community sample, MOREI creates on the order of $10^5 - 10^6$ microfluidic droplets for characterizing microbial interactions.

Multiple batches of droplets are created using droplet-based microfluidics chips[7]. Droplets within each batch are incubated for the same amount of time, while different batches are incubated for different amounts of time. Cf. figure 1.3.

The "life cycle" of each droplet can be divided into roughly three phases.

1. The first phase is "**initial droplet formation**". In the first phase, fluid from the microbial community sample is pumped from a syringe into the microfluidics chip. The microfluidics chip uses oil to encapsulate the fluid, as well as any cells that may have been inside the fluid, into a droplet.

2. The second phase is "**growth of cells inside of the droplet**". In the second phase, the droplet is incubated for the amount of time prescribed for its batch, during which any cells inside are able to divide and grow.

3. Finally the third phase is "**preparation for sequencing**". In the third phase, the droplet is fed into another microfluidics chip, where it is merged with another droplet containing materials that enable the sequencing and unique identification of the contents of the original droplet.

After all three phases are completed, all of the (merged) droplets are pooled together to undergo PCR[8] amplification and finally sequencing. Cf. figure 1.4 below. I will also describe all three phases in more detail below. Readers uninterested in the technical details of the experiment or of the resulting data might consider skipping the rest of this section.

---

[7] If you are unfamiliar with this technology, cf. [PSG$^+$21] for a recent helpful review.

[8] polymerase chain reaction
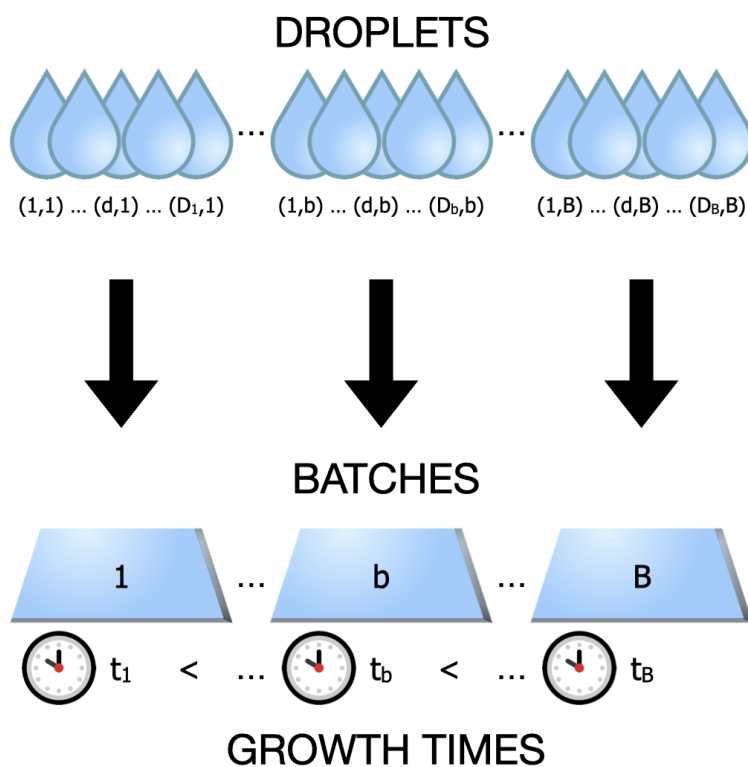
Figure 1.3: Notation; Organization of Droplets into Batches

1.4.1 Initial Droplet Formation

The microfluidics chip[9] encapsulates cells into droplets randomly. This makes each droplet's initial condition, both the number of cells in the droplet, as well as which the strains the cells belong to, random. The growth and division of cells during the second phase means that the droplets' random initial conditions are obscured in the final data. Therefore the best we can do is to make predictions of the droplets' unobserved random initial conditions based on average behavior.

We can control the concentration/dilution of the microbial community sample to guarantee a specified average number of cells per droplet. The average number of cells per droplet is usually chosen to be around two. This is intended to guarantee that, when there are two (or more) strains present in a droplet, the most likely scenario is that there is at most one cell of each strain.

The small average number of cells per droplet has other effects. The data is incredibly sparse. Even in a typical non-empty droplet, almost all strains will be absent. This is very much unlike, for example, typical single-cell RNAseq data. Another effect is that for any

---

[9] If you are unfamiliar with this technology, cf. [PSG$^+$21] for a recent helpful review.
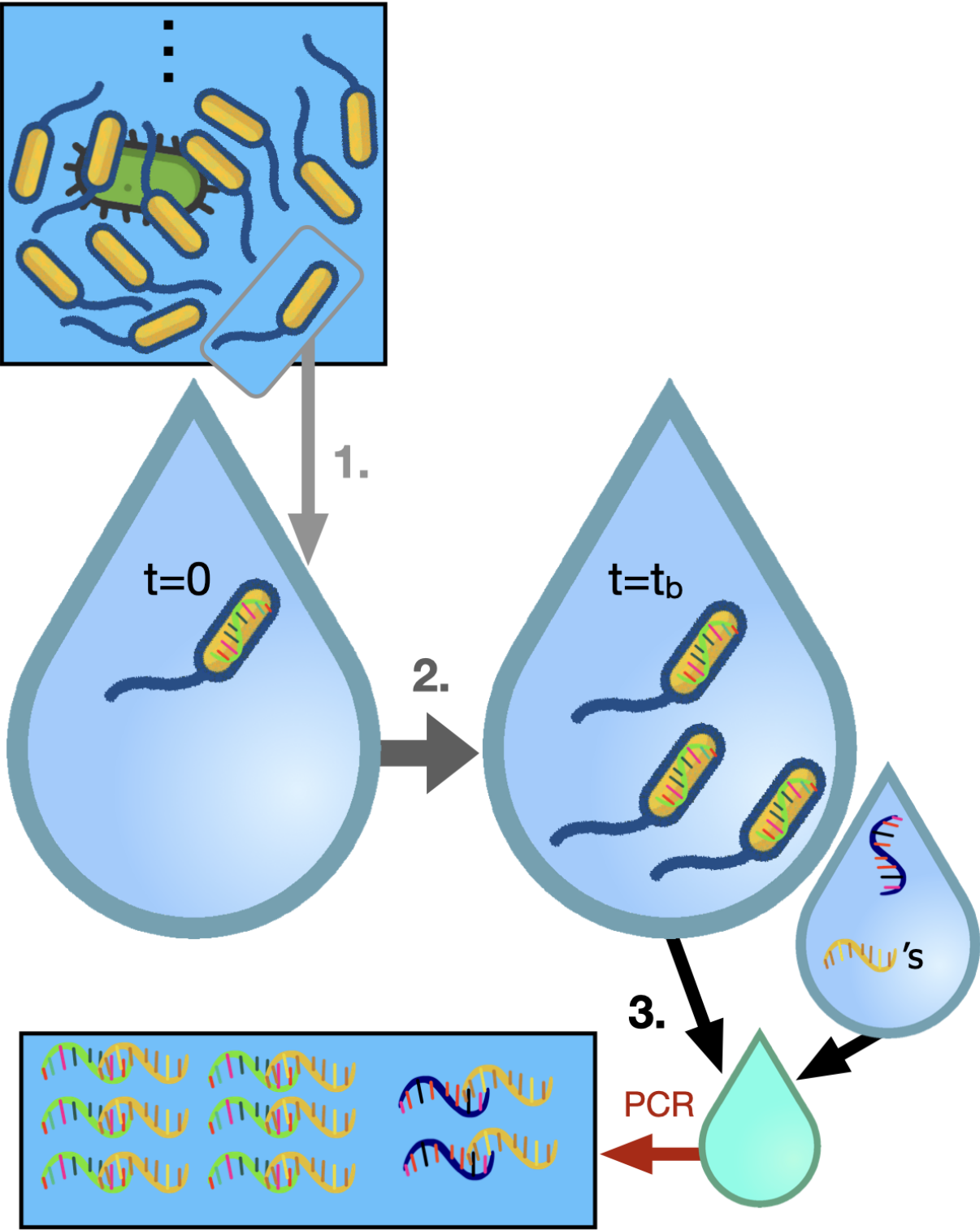
Figure 1.4: Schematic of the "life cycle" of a droplet. See text for details.

given interaction, the number of droplets that serve as controls will usually vastly exceed the number of droplets that serve as treatments.

1.4.2 Growth of Cells Inside of the Droplet

The contents of the droplets are only observed once, after the end of the second phase. In particular, the details of what occurs during the second phase inside of any given droplet are completely opaque to us. The resulting data is therefore not longitudinal. Thus many standard methods for inferring microbial interactions, which assume time series data[10], are not applicable. Currently there appears to be no scalable method that can measure the abundances of microbial strains within the same microfluidic droplet at multiple time points [DiSd17], so the limitation posed by the absence of longitudinal data appears unavoidable.

The droplets are all small, and thus nutrient-limited. Models of microbial interactions like the generalized Lotka-Volterra equations cannot make accurate predictions for the second phase, because they allow for indefinite growth. Cf. appendix A.1.2. The maximum possible number of cells that could exist inside of a droplet after incubation is *not* unlimited.

How long the droplets are incubated will determine whether cells are still actively dividing at the end of the second phase. It could be more difficult to infer the strengths of interactions from cells that exhausted the nutrients in their droplet and are undergoing logistic-like growth. This makes it important to incubate distinct batches for distinct amounts of time, to get "pseudo-longitudinal" data. Future work needs to determine how to best combine data from different batches, for example how to best balance averaging results within batches and averaging results between batches, while also accounting for how batches that were incubated longer are more likely to correspond to logistic-like growth.

1.4.3 Preparation for Sequencing

During the third phase, a second microfluidics chip[11] merges each droplet with a "PCR mix" droplet, whose contents prepare the contents of the original droplet for sequencing.

First there is a lysis buffer, which lyses the cells in the droplet. This kills any cells which are still alive, and makes the genetic material inside all cells (living or dead) accessible. The third phase is necessary for observing the contents of the droplet, but the lysis buffer also ensures that it is destructive and cannot be repeated. Therefore the third phase is what forces the data to not be longitudinal. Because dead cells are also lysed and have their genetic material available for sequencing, this also introduces potential biases from counting such "relic" genetic material the same as if it had come from living cells. Cf. figure 1.5.

---

[10] See e.g. [SBT+13] or [MBH+14]. An introduction to time series for microbial ecologists is [CHL+20].

[11] If you are unfamiliar with this technology, cf. [PSG+21] for a recent helpful review.
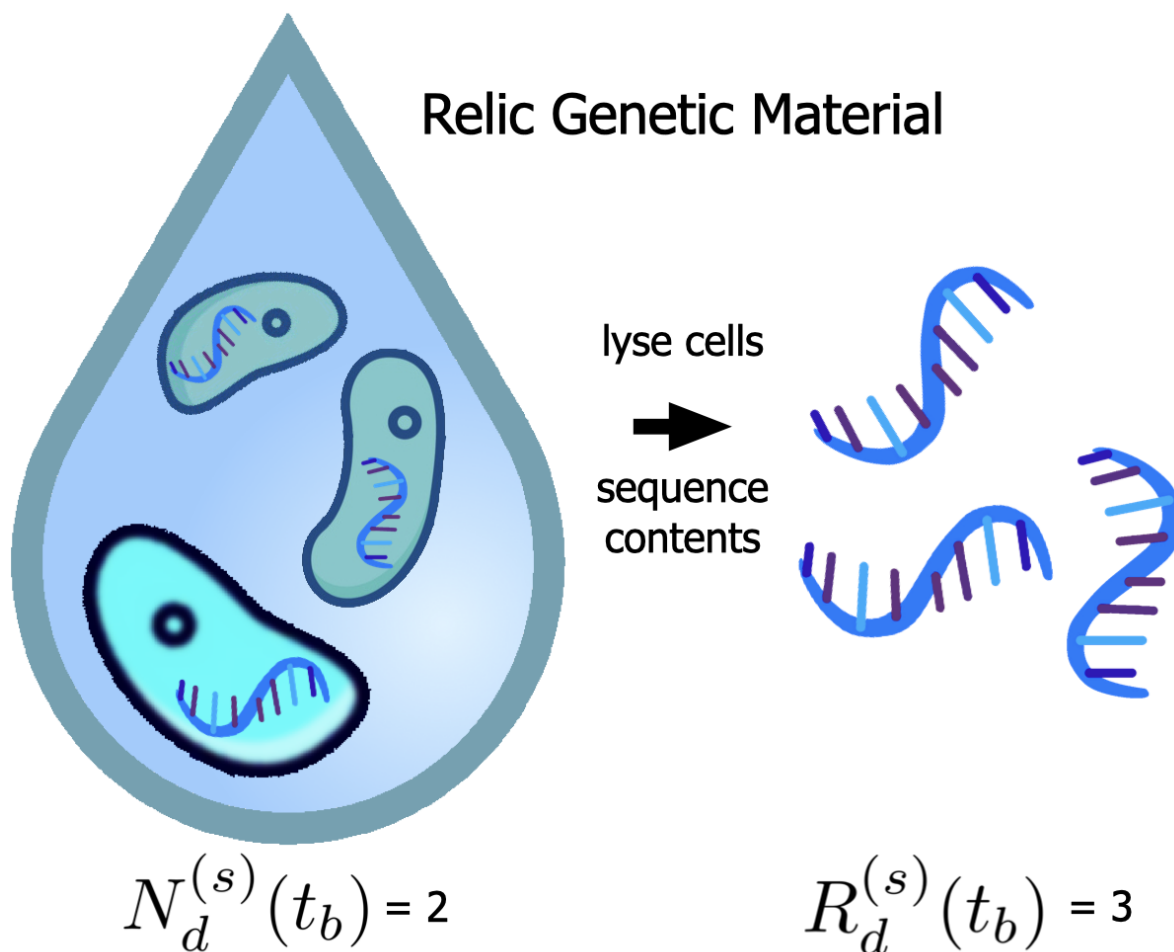
Figure 1.5: Relic genetic material can inflate inferred count of cells, potentially introducing biases. This droplet had two living cells and one dead cell after incubation. However, all three cells (living and dead) are lysed and thus also have their contents sequenced.

Second there are the barcodes, represented by the yellow nucleotides in figure 1.4. (For more details about barcode methodology, and examples of applications, see for example any of the papers [LHYA16], [BMK+18], [GR18], [RGG+21], or [DA21].) More precisely, the barcodes consist of (1) an Illumina Miseq adapter, (2) a barcode sequence, and (3) a 515F primer. The adapter ensures that the sequence and any other nucleotides that bind to it will be sequenced. The barcode sequences are (approximately[12]) unique for each "PCR mix" droplet. Therefore the barcode sequences identify everything they bind to as having come from the same droplet. Finally the 515F primers bind to 16S rRNA genes from the

---

[12]Choosing a barcode sequence of L nucleotides, there are 4^L possible barcode sequences of length L. Therefore, choosing a long enough length L for the barcode sequences, and generating them randomly, in principle one will have with high probability that the barcode sequence in each "PCR mix" droplet is unique, or that almost all are. This is much easier said than done, of course, but still feasible with today's technology.

lysed microbial cells. This ensures that primarily the 16S rRNA genes from the cells will be amplified in the PCR step that occurs following the third phase. Pooled replicate reactions may be unnecessary to avoid "jackpot effects and chimera formation" during PCR, at least for 16S data [MSH$^+$19]. Hence, we should be able to use the PCR-amplified 16S rRNA gene sequences to distinguish different strains and (approximately) count reads within droplets.

Third there are spikein genes, represented by the blue nucleotides in figure 1.4. A random, but approximately constant[13], number of these are incorporated into each "PCR mix" droplet. Like the 16S genes from inside of the lysed cells, these will also bind to the barcodes, and therefore will also be amplified in the PCR step that occurs following the third phase. Assuming that all sequences from the same droplet that bind to the barcodes will be amplified by roughly the same amount during the PCR step, for each droplet we can divide all of the PCR amplified counts by the ratio of the PCR amplified spikein count to the average number of spikein gene copies originally in each "PCR mix" droplet.

While details of the spikein genes may seem unimportant from a purely statistical perspective, they actually imply a very important feature of the data.

Specifically, the spikein gene protocol enables us to normalize by an (approximate) PCR amplification factor, which means that we get **absolute abundance** values from the sequencing results. Cf. [BBI20] for details of a related spikein protocol. Without estimates for the PCR amplification factor, normally one can only infer *relative abundance* values. However relative abundance data may be problematic for inferring microbial interactions [MMV$^+$08] or [BBI20, Figure 1]. This also implies that *none* of the vast, pre-existing literature on methods for studying the so-called "compositional data" coming from relative abundances is (directly) relevant to this problem. (Cf. e.g. [SWMD17], [CHL$^+$20], [MGSD19], [HCSB20], or the introduction of [PGB21] for examples of methods for dealing with compositional data.) This is true although previous work for inferring microbial (or genetic) interactions uses working models (superficially) similar[14] to those discussed herein.

A previous attempt, MINI-drop [HCT$^+$19], to study microbial interactions using microfluidic droplets while getting absolute abundance measurements was limited to only 3 distinct strains. Studies using kChip [KKO$^+$19] [KOK$^+$20] were able to investigate larger number of strains (e.g. 20 in the case of [KOK$^+$20]), but still nothing vastly surpassing the state of art using cell plating [VCF$^+$18]. The reason for this is the reliance of these methods on fluorescent labeling of cells to get absolute abundance measurements. That means that the total number of strains that can possibly be distinguished, and thus included in the experiment, is

---

[13] Assuming a roughly Poisson distribution[CNd$^+$15] (for which the mean equals the variance), the larger the mean, the smaller the ratio of the standard deviation to the mean. For example, $\frac{\pm 10}{100}$ is a better relative uncertainty than $\frac{\pm 5}{25}$, even though ± 10 is clearly a worse absolute uncertainty than ± 5. Relative uncertainty is all that matters in this context because we use the PCR amplified count of spikein genes as a normalization factor. Thus, as long as the average number of copies of spikein gene per "PCR mix" droplet is large enough, we may treat it as "approximately constant". However, we don't want to add too many copies of the spikein gene on average, because then after PCR amplification the spikein genes will likely "swamp out" any signal from strains that are (relatively) less abundant inside of the droplet.

[14] E.g. working models related to the Poisson or Dirichlet-Multinomial distributions, or working models using generalized Lotka-Volterra models, cf. [MBH$^+$14], [GHG$^+$18], and [SBT$^+$13].

limited by the number of distinct possible fluorescent proteins available whose fluorescence can be distinguished by automated imaging techniques. Moreover, it also requires either pre-existing transgenic strains of interest that express (compatible) fluorescent proteins, or genetically modifying de novo the organisms one wants to study with the hope of being able to get them to express fluorescent proteins. (This seems to be why [KOK$^+$20] only studied strains of *Enterobacterales* and *Pseudomonadales* γ-Proteobacteria. Such bacteria are closely related to intensively studied "lab rats"of microbiology, namely *E. Coli* and *Pseudomonas* respectively.) In contrast, MOREI is in principle capable of studying microbial communities with many more strains, while still producing absolute abundance measurements. This is because MOREI does not use automated fluorescence imaging to get absolute abundance measurements[15], and instead uses a spikein gene protocol (cf. [BBI20]). Therefore MOREI is not limited to studying transgenic strains, nor does the available number of distinguishable fluorescent proteins limit the number of strains that MOREI can study.

1.5 Conclusion

Findings and Contributions I explained why high-throughput droplet-based microfluidics technology overcomes previous limitations to answering these questions.

I showed how answering these questions can be recast as the statistical problem of estimating a network with mixed-sign edge weights.

Practical Implications MOREI needs to increase the number of strains whose interactions can be simultaneously characterized to advance the state of the art (cf. section 1.3). Being able to characterize the interactions of more strains requires characterizing the interactions of strains with low relative abundance (a.k.a. frequency) in the microbial community. For example, if all strains have the same frequency, then simultaneously characterizing the interactions of 20 strains requires characterizing the interactions of strains with 5% relative abundance, while simultaneously characterizing the interactions of 2,000 strains requires characterizing the interactions of strains with .05% relative abundance. Thus understanding how to use MOREI to advance the state of the art in microbial ecology requires understanding how to analyze the data produced by MOREI in a way that characterizes the interactions of the least abundant strains as accurately as possible.

Next Steps and Open Questions Given highly accurate simulations of the data-generating process for MOREI, we could compare the performance of different estimators to see which most accurately characterizes the interactions of the least abundant strains. This would determine the methods that are best for analyzing the data produced by MOREI. That in turn would determine how far we can realistically expect to advance the state of the art in microbial ecology using MOREI.

---

[15] Cf. [TVL$^+$22] for one fluorescence imaging protocol for counting bacteria in microfluidic droplets.

Chapter 2

Formulation as a Statistical Problem

Herein I explain that the data from the noisy dynamical systems corresponding to each droplet can be understood as censored observations of a multivariate Markov process. See section 2.2. I identify the statistical understanding of a droplet's initial state as crucial to overcoming the main limitation of these experiments, the uncontrolled assignment of microbes to droplets. See section 2.3.2. I also give the precise definitions of the targeted parameters that will be the focus of part II. See sections 2.3.4, 2.3.6.1, and 2.B.

Section 2.2 discusses the statistical model and questions for the entire experiment, when droplets are incubated. This corresponds to all three phases of the data generating process from section 1.4. Section 2.3 discusses the statistical model and questions for the important "$t = 0$" experiment special case, when droplets are *not* incubated. This corresponds only to the first and third phases of the data generating process from section 1.4.

In this chapter I make no attempt to explicitly model the third phase of the droplet life cycle described in section 1.4 (although see appendix A.2 for some very preliminary thoughts). The problem of choosing good statistical models of PCR amplification is applicable to an extremely wide range of biological experiments. It should be its own project and fully addressing it is outside of the scope of this work. Previous work that has more substantially considered similar problems includes [YRCS06] [KMP+07] [YWSJ08] [DH15][MWC19][MSH+19].

Chapter 2 Contents

2.1   Background

Section 2.1.1 overviews the general context of related problems across many fields. Section 2.1.2 narrows this down to only issues that are directly relevant to analysis of the data produced by MOREI. Finally, section 2.1.3 overviews the ideas behind the concrete proposals that will be made further in this chapter and throughout the dissertation.

2.1.1   Broader field

The ideas of this chapter belong to the hypothetical broader field of "*statistical inference of (time-evolving) interactions between variables of a (noisy) dynamical system for which the temporal evolution may be incompletely observed*". Cf. the discussion earlier from the

introduction to Part I. [GW11] appears to describe something very similar as *"inference for the parameters of complex nonlinear multivariate stochastic process models"*. The ideas in this chapter can be understood as belonging to any of the following general fields:

(1) The statistical inference of dynamical systems. See for example [ECR16], [RH17], and [Wil18] for introductory monographs. The papers [IBK06], [GW11], [SMY$^+$12], [EVL$^+$22], [AML$^+$17], and [LLLW11] are examples of important contributions to the field. This field apparently overlaps with both functional data analysis, see [RHCC07] and [DM21] for examples, and with ecological modelling, see [YEJ$^+$07], [Woo10], [RDC$^+$06], [PE00] for examples. Notice that this field appears to be unrelated[1] to the "dynamical system approach to multivariate data analysis" described in [Tre06].

(2) The study and mathematical theory of random dynamical systems. See [HS02] for a relatively accessible introduction, or [Arn10] for a comprehensive reference. This appears to include much of the theory of stochastic differential equations (SDEs) as a special case, cf. [Arn10, section 2.3] or [GW11]. Thus I do not discuss modelling and inference with SDEs as its own field below, although one could reasonably do so.

(3) The field of functional data analysis. See any of [WCM16], [UF13], [RHG09], [KR21], [RS05], [HK12], [HE15], or [Gre81] for introductions of varying levels of difficulty. As mentioned above, this apparently overlaps with the statistical inference of dynamical systems, with [RHCC07] and [DM21] as examples. The overlap of this field with manifold learning may be particularly important for multivariate data like this problem, perhaps because the phase space of the dynamical system is a (random) manifold. See [CM13], [CM12], and [ZYLS18] for examples.

(4) The study of variable interactions, or the inference of interaction networks. See [LR20], [AML$^+$17], or even [MSH$^+$08] for examples of the kinds of studies that I intend to describe. In regression terminology, in this context all of the response variables are also predictor variables for each other, so given n (response=predictor) variables, there are potentially $O(n^k)$ k-th order interactions. In the terminology of genetics studies, this corresponds only to "gene $\times$ gene" interactions (pairwise interactions grow *quadratically* with the number of genes), but *not* to e.g. "gene $\times$ environment" interactions (pairwise interactions grow only *linearly* with the number of genes). The

---

[1] [Tre06] explains how many problems in multivariate data analysis can be formulated as matrix fitting constrained optimization problems, which in turn can be reformulated into initial value problems for (ordinary) differential equations on the matrix manifolds defined by the constraints of the optimization problem. To clarify, "matrix manifold" refers to an abstract manifold whose points can be thought of as elements of a(n abstract) space of matrices. Then the idea is that dynamical systems theory can often be used to better understand the corresponding initial value problem. However, given the connection of the theory of statistical inference of dynamical systems with functional data analysis mentioned above, and the connection between functional data analysis and multivariate data analysis described in e.g. [RS05], this may still suggest an *indirect* (and non-trivial) connection between statistical inference of dynamical systems and the "dynamical system approach to multivariate data analysis". Of course this still remains to be confirmed.

former, "response=predictor" or "gene × gene", interactions problem is much more difficult than the latter problem, even if only for combinatorial reasons[2]. Unfortunately, this more difficult "gene × gene" interactions problem also seems to be less frequently studied than the easier "gene t environment" interactions problem. In the terminology from the Part I introduction, "gene × gene" interactions are the same as "unpartitioned interactions", whereas "gene × environment" interactions are the same as "bipartite interactions". The variables of the "environment" can also be described as "features", such that the "gene × environment" interactions problem also corresponds to the "feature enrichment" or "differential analysis" problem commonly studied with high-throughput biological assays, see [GCS+21] for a discussion.

Gene regulatory networks [WBGLR+21] [CK21] [MJS20] [TKA20] [LLLW11] are a common example of the "gene × environment" interactions problem, with transcription factors corresponding to the "environmental variable nodes". The preprint [LR20] suggests [HZ17] or [HFZ18] as recommendations for papers on regression with large numbers of (interaction) covariates, and any of [Cor09], [KL08], or [ZHSL12] as recommendations for papers studying the inference of "gene × gene" interactions. The paper [LLLW11], which takes a dynamical systems approach, also appears to be related. See also [MGSD19] vis a vis the "gene × gene" interactions problem.

(5) The study of pseudotemporal, or "pseudotemporally ordered", cross-sectional data[3]. While no textbook account thereof seems to exist, the references [TCG+14], [QMT+17], [CY18], and [SL21] are all important contributions. Current studies in this field appear to all be conceived within the context of analyzing data produced by high-throughput droplet-based microfluidics experiments.

In particular, valuable future work would further clarify the connections that these fields have not only with each other, but also with the fields of

(i) the statistical analysis of longitudinal (also known as "time series") data (for longitudinal studies of similar biological problems, cf. e.g. [PE00], [SBT+13], [GHG+18], [RGG+21], or cf. [CHL+20] for an introduction geared towards microbial ecologists),

(ii) the mathematical theory of Markov (or specifically Feller) stochastic processes (cf. e.g. [IBK06], [Gre81, chapter 6], [Arn10, appendix A.4], or [KP10, section 1.6] for discussion of how this mathematical theory is relevant to several of the aforementioned fields),

---

[2] Let n be the number of genes. Then a "gene × gene" interactions problem amounts to n instances of "(one) gene × (other genes)", or "gene × environment", interactions problems. For each of the n genes, construe the $(n-1)$ "other genes" as the "environment" to get a "gene × environment" interactions problem. Cf. [TKA20] for other ideas about relating "gene × environment" and "gene × gene" problems.

[3] Described in [CY18] as datasets where "individuals in the cross-sectional cohort behave asynchronously and each is at a different stage of progression". This should (in some sense) apply to the data produced by this experiment due to the use of batches incubated for different amounts of time. Cf. section 1.4.

(iii) (macro)ecological modelling studies (for relevant examples compare e.g. any of the following papers: [BKK16], [SMY$^+$12], [WSCR03], [SBT$^+$13], [GCLF18], [MXS17], [ELW$^+$20], [GLB$^+$18], [ESGVS21], [EVL$^+$22], [GWDM21], [MBH$^+$14], [LBG10], [YEJ$^+$07], [Woo10], [RDC$^+$06], [PE00], [GW11], [BSM$^+$21], [MJS20]), and

(iv) "repeated cross-sectional" or "pseudo-longitudinal" epidemiological studies.

Some amount of "synthesis" or "unification" of the relevant ideas from *all* of these fields would probably lead to a substantial increase in the breadth and sophistication of the techniques available for many problems. Nevertheless, it would be difficult to accomplish.

From this more general context, chapter 2 focuses on the important issue of statistical inference of (time-dependent) microbial interactions, using either longitudinal or pseudo-longitudinal data. The interacting variables are the populations of microbes. See [HFH21] for one example of this type of problem, and [GCLF18] for a helpful review of related work.

2.1.2 Specific problem

Herein I investigate how to approach the data produced by MOREI using statistical models. We want an explicitly specified statistical model to be the basis of our statistical analyses. This allows us to keep track of which assumptions we are making and how realistic they are. Previous work has implicitly (but not explicitly) specified longitudinal statistical models for inferring microbial interactions that amount to discrete-time Markov processes. See any of [FM14] [MMV$^+$08] [VCF$^+$18] [MBH$^+$14] for examples. However, explicitly specifying something like this would not be enough for MOREI because in this case the data is only pseudo-longitudinal. As mentioned in the review [GCLF18], most previous work using dynamical system models does not claim the ability to infer microbial interactions without access to (strictly) longitudinal data. The paper [XAF$^+$17] is a noteworthy exception. The paper begins by making strong differentiability assumptions about the population dynamics. Based on these assumptions, it then claims that it should be possible to infer ecological interactions from entirely cross-sectional data, not even necessarily pseudo-longitudinal, provided the dynamical system is always observed in a steady state or local equilibrium. Similarly, the pseudotime method from [CY18] also claims to be able to make similar inferences (cf. the discussion from (4) of 2.1.1 above) from entirely cross-sectional data.

The assumptions required by [XAF$^+$17] are questionable within the context of the data produced by MOREI. For example, the use of a pseudo-longitudinal data structure (from making multiple batches) already implies that we do not generally believe the equilibrium assumptions to be valid for most droplets. Likewise the method from [CY18] also fails to exploit the pseudo-longitudinal structure of the MOREI data. Moreover, the proposed algorithms from [XAF$^+$17] and [CY18] both do not exploit the sparsity of the MOREI data, and both fail to scale well with the number of observations (which for MOREI is massive). These issues lead to prohibitively severe performance problems when applied to MOREI data. Given these problems, I decided to not attempt identifying and describing explicit

models used implicitly in [XAF$^+$17] and [CY18]. The models would be unrealistic, and the inference methods they would suggest would not actually be useful for MOREI data.

Interestingly, [SSB$^+$17] claims to have developed a method based on the generalized Lotka-Volterra equations, analogous to e.g. [FM14], yet which also claims to work on cross-sectional data. *However*, the claim does not seem to be that they can make inferences about microbial interactions based on pseudo-longitudinal data. Instead, [SSB$^+$17] seems to essentially require cross-sectional data that is analogous to longitudinal data but with some other quantity as the "independent variable" besides time:

> series of cross-sectional data which originate from samples along gradients in environmental parameters but do not have a temporal dimension

This does not characterize the data produced by MOREI any more than longitudinality does. Hence the methodology of [SSB$^+$17] also seems to not be applicable to this problem. Nevertheless it does seem to warrant closer comparison in future work.

2.1.3 Particular approach

This chapter characterizes the droplet life cycle using a discrete-time Markov process whose initial state is in $\mathbb{N}_{\geq 0}^S$ but which thereafter can take values in $\mathbb{R}_{\geq 0}^S$ (cf. equation (A.2.1.1) from appendix A.2). This choice of model can be thought of as "implicitly" discretizing a continuous-time Markov process that is the "true model".

The idea of modelling a system similar to this as a Markov process observed at discrete time points has precedent in [IBK06], as well as in references [24, 25, 27, 30, 32] of [GW11]. Cf. also [KP10, Part IV], which explains and discusses procedures for discrete-time approximations of stochastic differential equation solutions. Similarly, precedent for applying models that are essentially stochastic differential equations (SDEs) to Lotka-Volterra-like systems and to microbial (prokaryotic) population dynamics can be found in [GW11] itself. Cf. also [KP10, section 7.1], which discusses applying SDEs to model population dynamics. The paper [BSM$^+$21] borrows tools from mathematical finance to investigate SDE models of bacterial growth that include "hedging" via persisters (slow-growing but resilient cells).

However, much previous work, such as [VCF$^+$18] [FM14] [MBH$^+$14] [SBT$^+$13] [MMV$^+$08], implicitly[4] used discrete-time Markov processes. So it is left to future work to fully consider e.g. [KP10, chapter 6], which discusses the applications of stochastic differential equations[5] to modelling problems, in particular [KP10, section 6.2] that discusses the continuous-time "diffusion limits" of discrete-time Markov processes. For example, it is doubtful that equation (A.2.1.2) from appendix A.2 is actually the appropriate diffusion limit of the discrete-time Markov processes used in previous work. A starting point might be [Arn10, section 2.1], which discusses how discrete-time random dynamical systems can be derived from processes like the discrete-time Markov process models from previous work, and vice versa.

---

[4]The references do not explicitly state the underlying statistical models, which would be needed to precisely define or justify the regression procedures that they propose.

[5]Which correspond to special kinds of continuous-time Markov processes.

The assumption of a discrete-time Markov process is made not only for the purpose of facilitating comparisons with prior work, but also for the purpose of model simplicity. In particular, there is the issue of the distinction between (using the terminology of [Arn10]) "random differential equations" [Arn10, section 2.2] and "stochastic differential equations" [Arn10, section 2.3]. Specifically, even if there might only be one standard way, via the diffusion limit, to go from a discrete-time Markov process to a continuous-time Markov process, there appears to be at least two standard ways to go from a (deterministic) differential equation to a "non-deterministic" differential equation, corresponding to "random differential equations" and "stochastic differential equations" respectively. The discrete-time Markov process models implicit in [VCF+18] [FM14] [MBH+14] [SBT+13][MMV+08] were all derived by discretizing and then "stochasticizing" a (deterministic) differential equation[6]. Hence we can ask whether, when extending these models to continuous-time, it might be better to "skip the intermediary" and seek to instead directly "stochasticize" the (deterministic) differential equation. If we believe it is better to do so, then it becomes incumbent upon us to clearly articulate a reason for a preference of a "random differential equation" model versus a "stochastic differential equation" model. To ensure that all of this is done correctly and well requires considerable mathematical sophistication and hence is left to future work.

At the same time, some choice of a continuous-time Markov process model (cf. equation (A.2.1.2) from appendix A.2) would still probably be more accurate or realistic than the current discrete-time model. Another benefit of using a continuous-time model would be that time periods of unequal length between the batches' incubation times would not complicate the interpretation of such a continuous-time model, cf. the corresponding argument made in [HFH21]. The references [IBK06] and [GW11] also seem relevant to this question.

There is precedent for using continuous-time stochastic processes to model the growth of bacterial populations in microfluidics droplets. See [TVL+22] which claims to find empirical support for such a model using automated fluorescence imaging, elaborating on prior work [BSJAB19] that generalized the "classical" Bellman-Harris branching process to apply to the growth of bacteria inside of microfluidic droplets. The Bellman-Harris process (as well as its generalizations in [BSJAB19] and [TVL+22]) is a continuous-time stochastic process with *discrete* state space, i.e. taking values in $\mathbb{N}$. Contrast this with an SDE model, which is continuous-time with continuous state space, and the model proposed below in section 2.2, where initial states are required to be discrete but subsequent states are allowed to be "pseudocounts" i.e. continuous. Note that both the models in [TVL+22] and [BSJAB19] are explicitly designed for the count of cells of a single, non-interacting strain, i.e. they are single-type branching processes. Hence future work trying to apply similar models to MOREI would need to begin by identifying the generalization of these models as multi-type branching processes. Cf. [AN72] for a reference about multi-type branching processes.

While this model was intended to describe MOREI specifically, in principle it should not be too difficult to modify in a way that makes it applicable to other "co-culture" systems as well. (See [GFP14] for a perhaps somewhat dated review of examples of "co-culture"

---

[6] Cf. the procedure used for deriving equation (A.2.1.1) in the appendix section A.2.1.

systems for studying cellular interactions.) These modified models might even describe their corresponding systems more accurately than this model describes MOREI. For example, the assumption that different experimental replicates (in the case of MOREI, droplets) are statistically independent is probably more realistic for other systems.

2.2 "Full" Experiment

Below I propose a general form for likelihoods characterizing the first and second phases of the droplet life cycle that were described in section 1.4. This corresponds to a discrete time Markov process whose initial state is in $\mathbb{N}_{\geq 0}^S$ but which thereafter can take values in $\mathbb{R}_{\geq 0}^S$.

2.2.1 Preliminaries

Let "S" for "**s**trains" denote the total number of microbial strains[7] in the microbial community sample. For each strain s, let $f^{(s)}$ ("f" for "**f**requency") denote the relative abundance of strain s in the sampling pool. In particular, we necessarily have that $\sum_{s=1}^S f^{(s)} = 1$.

Given a positive integer $N \in \mathbb{N}$, define $[N] := \{1, \ldots, n, \ldots, N\}$. In general, the entries of a (non-negative) vector $\vec{v} \in \mathbb{R}_{\geq 0}^S$ are denoted $v^{(s)}$, i.e.

$$\vec{v} \overset{def}{=} \left(v^{(1)}, \ldots, v^{(s)}, \ldots, v^{(S)} \right). \tag{2.2.1}$$

The shorthand $v := \sum_{s=1}^S v^{(s)}$ is used for the sum of the entries of $\vec{v}$.

For a given droplet d ("d" for "**d**roplet"), the random variable $N_d(0)$ is the number of cells from the population encapsulated in the droplet d during its formation ("N" for "**n**umber"). Equivalently, $N_d(0)$ is the number of cells in droplet d at time 0. This is the sum of the entries of the random vector:

$$\vec{\mathbf{N}}_d(0) := \left(N_d^{(1)}(0), \ldots, N_d^{(s)}(0), \ldots, N_d^{(S)}(0) \right), \tag{2.2.2}$$

where for each strain s ("s" for "**s**train") of the S total strains, the random variable $N_d^{(s)}(0)$ is the number of cells from strain s in droplet d at time 0.

For the sake of brevity, often the subscript d is omitted, i.e. $N(0)$, $\vec{\mathbf{N}}(0)$, and $N^{(s)}(0)$ are used in place of $N_d(0)$, $\vec{\mathbf{N}}_d(0)$, and $N_d^{(s)}(0)$, respectively. This highlights how the random variables corresponding to the droplets are assumed (for the sake of simplicity) to be identically distributed. Therefore the corresponding statements should not depend on the specific droplet d, making it unnecessary or even misleading to specify d. Cf. figure 3.2.

For simplicity, the $\vec{\mathbf{N}}_d(0)$ are assumed to not only be identically distributed, but also mutually statistically independent. This becomes relevant in sections 2.2.3 and 2.2.4 below,

---

[7]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.

and is implicit throughout everything that follows, e.g. the derivations in chapter 6. In practice this is probably unrealistic.

Let "B" for "**b**atches" denote the total number of batches. For any given batch $b \in [B]$, let $D_b \in \mathbb{N}$ for "**d**roplets" denote the total number of microfluidic droplets belonging to the batch. With current droplet microfluidics technology, D_b can be on the order of 10^6 or greater. Whether one assumes for simplicity that $D_1 = \cdots = D_B$ does not substantially affect the analysis. Droplets in different batches are measured at different times, and for the entire experiment the total number of droplets, denoted D, is

$$D := \sum_{b=1}^{B} D_b = B\bar{D}, \tag{2.2.3}$$

where $\bar{D}$ denotes the arithmetic mean of the D_b. Compare Figure 1.3.

During the second phase that was described in section 1.4, the cells of all D_b droplets in a given batch b, for any $b \in [B]$, are incubated until time t_b has elapsed ("t" for "**t**ime"), until the droplets are measured (which begins the third phase). The incubation time t_b for any given batch b can be subdivided into b time periods, corresponding to $0 =: t_0 < t_1 < \cdots < t_\beta < \cdots < t_b$, the incubation times for all batches that are not incubated longer than batch b. Again, compare Figure 1.3. Each time period does not necessarily need to have the same length, although that could simplify interpretation of the analysis.

2.2.1.1 Observables and Format of Data

After the end of the experiment, for each batch $b \in [B]$ what we observe are (roughly[8]) the D_b vectors $\vec{\mathbf{N}}_d(t_b) \in \mathbb{R}_{\geq 0}^S$, one for each $d \in [D_b]$. These data of course can be formatted into a $S \times D_b$ or $D_b \times S$ matrix.

2.2.2 Likelihood of a Single Droplet

For a given droplet, at time $t = 0$ the unobserved likelihood $\mathbb{N}_{\geq 0}^S \to [0,1]$ for the initial strain distribution vector, i.e. the random vector $\vec{\mathbf{N}}(0) \in \mathbb{N}_{\geq 0}^S$ giving the initial counts of the number of living cells in the droplet from each of the S strains in the population, is for any $\vec{n} \in \mathbb{N}_{\geq 0}^S$ denoted

$$\mathbb{P}\left(\vec{\mathbf{N}}(0) = \vec{n}\right), \tag{2.2.4}$$

where $\vec{n} =: (n^{(1)}, \ldots, n^{(S)})$ (such that for all $s \in [S]$ one has $n^{(s)} \in \mathbb{N}_{\geq 0}$). This corresponds to the first phase that was described in section 1.4.

Various possible likelihood models specifying $\mathbb{P}\left(\vec{\mathbf{N}}(0) = \vec{n}\right)$ are derived and described in chapter 3. Chapter 4 compares these various models. Chapter 5 discusses the different "downstream" impact that each of these models has on the amount of data available for

---

[8]Cf. appendix A.2 for some distortions that the third phase from section 1.4 causes.

inferring interactions. Finally, chapter 6 discusses how we might go about choosing which of these models is most realistic based on observations of $\vec{\mathbf{N}}(0)$ (which can be accomplished by not incubating the droplets and thereby skipping the second phase of the droplet life cycle that was discussed above in section 1.4.2).

The unobserved full likelihood $\mathbb{N}_{\geq 0}^S \times \left(\prod_{\beta=1}^b \mathbb{R}_{\geq 0}^S \right) \to [0, 1]$ for a given droplet in batch b over the entire duration of the growth of its cells is assumed to correspond to a discrete time Markov process. In other words the unobserved full likelihood for a single droplet of a given batch b is assumed to be of the form:

$$
\begin{aligned}
&\mathcal{L}\left(\vec{\mathbf{N}}(0) = \vec{n}(0), \vec{\mathbf{N}}(t_1) = \vec{n}(t_1), \dots, \vec{\mathbf{N}}(t_b) = \vec{n}(t_b) \right) \\
&= \left[\prod_{\beta=1}^b \mathcal{L}\left(\vec{\mathbf{N}}(t_\beta) = \vec{n}(t_\beta) \ \middle| \ \vec{\mathbf{N}}(t_{\beta-1}) = \vec{n}(t_{\beta-1}) \right) \right] \cdot \left[\mathbb{P}\left(\vec{\mathbf{N}}(0) = \vec{n}(0) \right) \right],
\end{aligned}
\tag{2.2.5}
$$

where $\vec{n}(t_\beta) \in \mathbb{R}_{\geq 0}^S$ for all $\beta \in [b]$, $\vec{n}(0) \in \mathbb{N}_{\geq 0}^S$, and $\vec{n}(0) =: (n^{(1)}(0), \dots, n^{(S)}(0))$. The transition likelihoods $\mathcal{L}$ from (2.2.5) correspond to the second phase that was described in section 1.4. Non-integer values of the (pseudo-)counts $n^{(s)}(\beta) \in \mathbb{R}_{\geq 0}$ can be interpreted to correspond to e.g. incomplete cell division.

Since the contents of the droplets in a given batch b are only observed at time t_b (when the cells inside are lysed and their genetic material sequenced), starting from (2.2.5) it follows that the *observed* likelihood for a single droplet of batch b is of the form:

$$
\begin{aligned}
&\mathcal{L}\left(\vec{\mathbf{N}}(t_b) = \vec{n}(t_b) \right) \\
&= \int_{\mathbb{R}_{\geq 0}^S} \cdots \int_{\mathbb{R}_{\geq 0}^S} \left[\sum_{\vec{m}(0) \in \mathbb{N}_{\geq 0}^S} \mathcal{L}\left(\vec{\mathbf{N}}(t_b) = \vec{n}(t_b) \ \middle| \ \vec{\mathbf{N}}(t_{b-1}) = \vec{m}(t_{b-1}) \right) \cdot \right. \\
&\quad \left. \left[\prod_{\beta=1}^{b-1} \mathcal{L}\left(\vec{\mathbf{N}}(t_\beta) = \vec{m}(t_\beta) \ \middle| \ \vec{\mathbf{N}}(t_{\beta-1}) = \vec{m}(t_{\beta-1}) \right) \right] \cdot \left[\mathbb{P}\left(\vec{\mathbf{N}}(0) = \vec{m}(0) \right) \right] \right] \mathrm{d}\vec{m}(t_1) \cdots \mathrm{d}\vec{m}(t_{b-1}),
\end{aligned}
\tag{2.2.6}
$$

where $\vec{m}(0) \in \mathbb{N}_{\geq 0}^S$ and $\vec{m}(t_\beta) =: (m^{(1)}(t_\beta), \dots, m^{(s)}(t_\beta), \dots, m^{(S)}(t_\beta)) \in \mathbb{R}_{\geq 0}^S$ for all $\beta \in [b-1]$ denote dummy variables, corresponding to the values from the full likelihood (2.2.5) that are unobserved and which therefore need to be marginalized out. Also, $\int_{\mathbb{R}_{\geq 0}^S} \cdots \mathrm{d}\vec{m}(t_\beta)$ is used as shorthand for the multiple integral $\int_{\mathbb{R}_{\geq 0}} \cdots \int_{\mathbb{R}_{\geq 0}} \cdots \mathrm{d}m^{(1)}(t_\beta) \cdots \mathrm{d}m^{(S)}(t_\beta)$ for all $\beta \in [b-1]$.

2.2.3 Likelihood of a Single Batch

Letting $\vec{\mathbf{N}}_{d,b}(t)$ denote the strain distribution vector of the d'th droplet of batch b at time t, and $N_{d,b}^{(s)}(t)$ denote the number of living cells of strain s in the d'th droplet of batch b at

time t, i.e.

$$\vec{\mathbf{N}}_{d,b}(t) := (N_{d,b}^{(1)}(t), \ldots, N_{d,b}^{(s)}(t), \ldots, N_{d,b}^{(S)}(t)) \in \mathbb{R}_{\geq 0}^{S}, \tag{2.2.7}$$

for all $d \in [D_b]$, $s \in [S]$, and $t \geq 0$, then the unobserved full likelihood for (all of the droplets of) a given batch b is of the form:

$$\mathcal{L}\left(\vec{\mathbf{N}}_{1,b}(0) = \vec{n}_{1,b}(0), \ldots, \vec{\mathbf{N}}_{1,b}(t_b) = \vec{n}_{1,b}(t_b), \vec{\mathbf{N}}_{2,b}(0) = \vec{n}_{2,b}(0), \ldots,\right.$$
$$\left.\vec{\mathbf{N}}_{D_b,b}(0) = \vec{n}_{D_b,b}(0), \ldots, \vec{\mathbf{N}}_{D_b,b}(t_b) = \vec{n}_{D_b,b}(t_b)\right)$$
$$= \prod_{d=1}^{D_b}\left(\left[\prod_{\beta=1}^{b} \mathcal{L}\left(\vec{\mathbf{N}}_{d,b}(t_\beta) = \vec{n}_{d,b}(t_\beta) \mid \vec{\mathbf{N}}_{d,b}(t_{\beta-1}) = \vec{n}_{d,b}(t_{\beta-1})\right)\right] \cdot \left[\mathbb{P}\left(\vec{\mathbf{N}}_{d,b}(0) = \vec{n}_{d,b}(0)\right)\right]\right), \tag{2.2.8}$$

where $\vec{n}_{d,b}(t_\beta) =: (n_{d,b}^{(1)}(t_\beta), \ldots, n_{d,b}^{(s)}(t_\beta), \ldots, n_{d,b}^{(S)}(t_\beta)) \in \mathbb{R}_{\geq 0}^{S}$ for all $\beta \in \{0\} \cup [b]$ and all $d \in [D_b]$. Extending equation (2.2.6) analogously to how equation (2.2.5) was extended to equation (2.2.8), the observed likelihood for (all of the droplets of) a given batch b is assumed to have the following form:

$$\mathcal{L}\left(\vec{\mathbf{N}}_{1,b}(t_b) = \vec{n}_{1,b}(t_b), \vec{\mathbf{N}}_{2,b}(t_b) = \vec{n}_{2,b}(t_b), \ldots, \vec{\mathbf{N}}_{D_b,b}(t_b) = \vec{n}_{D_b,b}(t_b)\right)$$
$$= \prod_{d=1}^{D_b}\left(\int_{\mathbb{R}_{\geq 0}^{S}} \cdots \int_{\mathbb{R}_{\geq 0}^{S}} \left[\sum_{\vec{m}_{d,b}(0) \in \mathbb{N}_{\geq 0}^{S}} \mathcal{L}\left(\vec{\mathbf{N}}_{d,b}(t_b) = \vec{n}_{d,b}(t_b) \mid \vec{\mathbf{N}}_{d,b}(t_{b-1}) = \vec{m}_{d,b}(t_{b-1})\right) \cdot \right.\right.$$
$$\left.\left[\prod_{\beta=1}^{b-1} \mathcal{L}\left(\vec{\mathbf{N}}_{d,b}(t_\beta) = \vec{m}_{d,b}(t_\beta) \mid \vec{\mathbf{N}}_{d,b}(t_{\beta-1}) = \vec{m}_{d,b}(t_{\beta-1})\right)\right] \cdot \left[\mathbb{P}\left(\vec{\mathbf{N}}_{d,b}(0) = \vec{m}_{d,b}(0)\right)\right]\right] \tag{2.2.9}$$
$$\left. \mathrm{d}\vec{m}_{d,b}(t_1) \cdots \mathrm{d}\vec{m}_{d,b}(t_{b-1})\right).$$

I claim it is valid, going from (2.2.8) to (2.2.9), to put the multiple integrals (and sum[9]) inside the product $\prod_{d=1}^{D}$ (thereby changing the order of integration) because the integrands (the likelihood factors) are non-negative, and because the values of the multiple integral $\int_{\mathbb{R}_{\geq 0}^{S}} \cdots \int_{\mathbb{R}_{\geq 0}^{S}} \sum_{\vec{m}_{d_1,b}(0) \in \mathbb{N}_{\geq 0}^{S}} \cdots \mathrm{d}\vec{m}_{d_1,b}(t_1) \cdots \mathrm{d}\vec{m}_{d_1,b}(t_{b-1})$ does not depend on the values of the dummy variables $\vec{m}_{d_2,b}(0), \vec{m}_{d_2,b}(t_1), \ldots, \vec{m}_{d_2,b}(t_{b-1})$ whenever $d_1 \neq d_2$. (The integrations "can be performed separately for each droplet and therefore distributed across the product".)

---

[9]Which from a measure-theoretic point of view is a special case of an integral.

2.2.4 Likelihood of Entire Experiment

For each given $b \in [B]$ and each $d \in [D_b]$, the d'th droplet of batch b and its strain distribution vector at time t:

$$\vec{\mathbf{N}}_{d,b}(t) := \vec{\mathbf{N}}_{d+\sum_{\beta=1}^{b-1} D_\beta}(t), \tag{2.2.10}$$

are the same as the $\left[d + \sum_{\beta=1}^{b-1} D_\beta \right]$'th droplet of the entire experiment and its strain distribution vector at time t. Consequently, for every $b \in [B], d \in [D_b]$, and $s \in [S]$:

$$N_{d,b}^{(s)}(t) := N_{d+\sum_{\beta=1}^{b-1} D_\beta}^{(s)}(t), \tag{2.2.11}$$

is the number of living strain s cells in the $\left[d + \sum_{\beta=1}^{b-1} D_\beta \right]$'th droplet at time t.

Even though it may be unrealistic to assume that the distributions of individual droplets within a batch are statistically independent, it is definitely realistic to assume that the distributions for distinct batches are statistically independent. Assuming the distributions for distinct batches are statistically independent motivates the derivation below.

Because we assume all droplets are sampled independently, and that they have the same distribution at any common given time, it follows that the unobserved full likelihood for the entire experiment must have the form:

$$\mathcal{L}\left(\vec{\mathbf{N}}_{1,1}(0) = \vec{n}_{1,1}(0), \ldots, \vec{\mathbf{N}}_{D_1,1}(t_1) = \vec{n}_{D_1,1}(t_1), \vec{\mathbf{N}}_{1,2}(0) = \vec{n}_{1,2}(0), \ldots, \right.$$

$$\vec{\mathbf{N}}_{1,b}(0) = \vec{n}_{1,b}(0), \ldots, \vec{\mathbf{N}}_{D_b,b}(t_b) = \vec{n}_{D_b,b}(t_b), \vec{\mathbf{N}}_{1,b+1}(0) = \vec{n}_{1,b+1}(0),$$

$$\left. \ldots, \vec{\mathbf{N}}_{1,B}(0) = \vec{n}_{1,B}(0), \ldots, \vec{\mathbf{N}}_{D_B,B}(t_B) = \vec{\mathbf{N}}_{D_B,B}(t_B) \right)$$

$$= \prod_{b=1}^{B} \left(\prod_{d=1}^{D_b} \left(\left[\prod_{\beta=1}^{b} \mathcal{L}\left(\vec{\mathbf{N}}_{d,b}(t_\beta) = \vec{n}_{d,b}(t_\beta) \mid \vec{\mathbf{N}}_{d,b}(t_{\beta-1}) = \vec{n}_{d,b}(t_{\beta-1}) \right) \right] \cdot \left[\mathbb{P}\left(\vec{\mathbf{N}}_{d,b}(0) = \vec{n}_{d,b}(0) \right) \right] \right) \right).$$

$$\tag{2.2.12}$$

Thus the observed likelihood for the entire experiment has the form:

$$
\mathcal{L}\left(\vec{\mathbf{N}}_{1,1}(t_1) = \vec{n}_{1,1}(t_1), \ldots, \vec{\mathbf{N}}_{D_1,1}(t_1) = \vec{n}_{D_1,1}(t_1), \vec{\mathbf{N}}_{1,2}(t_2) = \vec{n}_{1,2}(t_2), \ldots, \right.
$$
$$
\vec{\mathbf{N}}_{1,b}(t_b) = \vec{n}_{1,b}(t_b), \ldots, \vec{\mathbf{N}}_{D_b,b}(t_b) = \vec{n}_{D_b,b}(t_b), \vec{\mathbf{N}}_{1,b+1}(t_{b+1}) = \vec{n}_{1,b+1}(t_{b+1}),
$$
$$
\left. \ldots, \vec{\mathbf{N}}_{1,B}(t_B) = \vec{n}_{1,B}(t_B), \ldots, \vec{\mathbf{N}}_{D_B,B}(t_B) = \vec{n}_{D_B,B}(t_B) \right)
$$
$$
= \prod_{b=1}^{B} \left(\prod_{d=1}^{D_b} \left(\int_{\mathbb{R}_{\geq 0}^S} \cdots \int_{\mathbb{R}_{\geq 0}^S} \left[\sum_{\vec{m}_{d,b}(0) \in \mathbb{N}_{\geq 0}^S} \mathcal{L}\left(\vec{\mathbf{N}}_{d,b}(t_b) = \vec{n}_{d,b}(t_b) \,\middle|\, \vec{\mathbf{N}}_{d,b}(t_{b-1}) = \vec{m}_{d,b}(t_{b-1}) \right) \cdot \right. \right. \right.
$$
$$
\left[\prod_{\beta=1}^{b-1} \mathcal{L}\left(\vec{\mathbf{N}}_{d,b}(t_\beta) = \vec{m}_{d,b}(t_\beta) \,\middle|\, \vec{\mathbf{N}}_{d,b}(t_{\beta-1}) = \vec{m}_{d,b}(t_{\beta-1}) \right) \right] \cdot \left[\mathbb{P}\left(\vec{\mathbf{N}}_{d,b}(0) = \vec{m}_{d,b}(0) \right) \right] \Bigg]
$$
$$
\left. \left. \left. \mathrm{d}\vec{m}_{d,b}(t_1) \cdots \mathrm{d}\vec{m}_{d,b}(t_{b-1}) \right) \right) \right) .
$$

(2.2.13)

Note that, as implied by the above equations, each droplet has its contents measured exactly once, regardless of the batch to which it belongs. In particular the data is not longitudinal in the sense that no droplet is tracked across more than one time point. The batch a droplet belongs to only determines how long its contents are allowed to grow before they are destroyed (killed via cell lysis) in order to measure them. Cf. again section 1.4.2 above.

Going from (2.2.12) to (2.2.13), the multiple integrals and sum can be distributed across the product $\prod_{b=1}^{B} \prod_{d=1}^{D_b}$ for reasons analogous to those for going from (2.2.8) to (2.2.9).

2.2.5   Possible Form of the Transition Likelihoods $\mathcal{L}$

Cells within a microfluidic droplet have the opportunity to interact only while the droplet is being incubated, corresponding to the second phase of the experiment described in section 1.4. Because the transition likelihoods $\mathcal{L}$ correspond to the second phase, it is clear that the interactions between microbes must be encoded within the form of the transition likelihoods $\mathcal{L}$, regardless of whether that corresponds to a discrete time Markov process or an SDE. The challenge lies in (1) choosing a non-parametric estimand that sensibly quantifies the notion of "interaction" regardless of the underlying model[10], or (2) choosing a parametric model that is both biologically realistic and that has parameters directly quantifying the interactions. Neither strategy is easy to implement well. All comparable prior work that I know of corresponds to strategy (2), although my opinion is that in the long-term only strategy (1) is likely to have indefinite "marginal returns". Cf. [vdLR11] for a discussion of parametric vs. non-parametric models.

---

[10]Allowing us to escape the need to posit any explicit form for the transition likelihoods.

For the purposes of facilitating comparisons with prior work where equations similar to the following were either introduced explicitly [FM14] [SBT$^+$13] or implicitly [VCF$^+$18] [MBH$^+$14] [MMV$^+$08], the transition likelihoods in equations (2.2.8)(2.2.9)(2.2.12)(2.2.13) would most often be assumed below to be implicitly specified by the following relationships for a given droplet in a given batch b, applicable for all $\beta \in [b]$ and $s \in [S]$:

$$
\begin{cases}
\log(N^{(s)}(t_\beta)) - \log(N^{(s)}(t_{\beta-1})) = r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} N^{(\sigma)}(t_{\beta-1}) + \varepsilon & N^{(s)}(t_{\beta-1}) > 0 \\
N^{(s)}(t_\beta) = 0 & N^{(s)}(t_{\beta-1}) = 0
\end{cases}
\qquad (2.2.14)
$$

where the ε denote i.i.d. random variables with a Gaussian$(0, \eta^2)$ distribution ($\eta > 0$ being some "noise scale" parameter), and the r_s and $\alpha_{\sigma s} \in \mathbb{R}$ are parameters to be estimated from the data. Using induction, we can see that the above relationships (2.2.14) are equivalent to

$$
\begin{cases}
\log(N^{(s)}(t_\beta)) - \log(N^{(s)}(t_{\beta-1})) = r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} N^{(\sigma)}(t_{\beta-1}) + \varepsilon & N^{(s)}(0) > 0 \\
N^{(s)}(t_\beta) = 0 & N^{(s)}(0) = 0
\end{cases}
, \qquad (2.2.15)
$$

valid for all $\beta \in [b]$ and all $s \in [S]$. This implies that observations of $N^{(s)}(t_b)$ alone should in principle be sufficient to determine whether $N^{(s)}(0)$ was nonzero, except in cases where one of the strains in the droplet goes "extinct" during the course of incubation (but cf. also appendix A.2 as well as the extensive discussion of this in section 7.8.1).

It is worth noting explicitly that for this model the coefficients $\alpha_{s_1 s_2}$ encode the microbial interactions. Cf. appendix A.1 for more detail.

The equations (2.2.15) imply the biologically plausible assumption that cells from a given strain cannot spontaneously appear in a droplet if they were not previously present. They are motivated by the goal of creating a discretized and stochastic version of the generalized Lotka-Volterra equations (A.1.3.1) from appendix A.1 in such a way that the coefficients can be estimated from time series[11]/longitudinal data using some form of regression[VCF$^+$18] [FM14] [MBH$^+$14] [SBT$^+$13] [GHG$^+$18] [MMV$^+$08]. However, like the generalized Lotka-Volterra equations they also allow indefinite growth under some conditions, which as discussed in section 1.4.2 is unrealistic. Cf. also the discussion in appendices A.1 and A.2.

2.3 "$t = 0$" experiment

The "$t = 0$" ("sub-")experiment differs from the "full" experiment described in section 2.2 in that droplets are not incubated. Cells are not allowed to grow or divide. The droplets' contents are sequenced immediately after formation, skipping the second phase from 1.4. Section 2.3.1 reviews the relevant notation and observed data. Section 2.3.3 clarifies the

---

[11] See [CHL$^+$20] for an introduction to time series geared towards microbial ecologists.

statistical model. Section 2.3.4 discusses the original target parameter definitions. Section 2.3.5 explains why in practice one most likely has to use working models for this problem. Section 2.3.6 gives the default working model that an experimentalist would be expected to use, as well as corresponding projected target parameters. Section 2.3.7 outlines the reasons why we might expect this working model to be inadequate in practice. Section 2.3.8 defines a generalized working model which mitigates some of these problems, as well as corresponding projected target parameters.

2.3.1 Notation and Observed Data

Let "D" for "**d**roplets" denote the total number of microfluidic droplets formed from a sampling pool of cells, and S denote the total number of cell types in the sampling pool. In the motivating microbial ecology example, "S" for "**s**trains" is the total number of microbial strains in the microbial community sample.

Given a positive integer $N \in \mathbb{N}$, define $[N] := \{1, \ldots, n, \ldots, N\}$. Unless specified otherwise, given a vector $\vec{v} \in \mathbb{R}^S_{\geq 0}$, its entries are denoted

$$\vec{v} \stackrel{def}{=} \left(v^{(1)}, \ldots, v^{(s)}, \ldots, v^{(S)}\right). \tag{2.3.1}$$

The shorthand $v := \sum_{s=1}^S v^{(s)}$ is used for the sum of the entries of $\vec{v}$. Another useful notion is the **support** of a (non-negative) vector $\vec{v} \in \mathbb{R}^S_{\geq 0}$:

$$\text{supp}(\vec{v}) := \{s \in [S] : v^{(s)} > 0\} \subseteq [S]. \tag{2.3.2}$$

For a given droplet d ("d" for "**d**roplet") , the random variable $N_d(0)$ is the number of cells from the population encapsulated in the droplet d during its formation ("N" for "**n**umber"). Equivalently, $N_d(0)$ is the number of cells in droplet d at "time 0". This is the sum of the entries of the random vector:

$$\vec{\mathbf{N}}_d(0) := \left(N_d^{(1)}(0), \ldots, N_d^{(s)}(0), \ldots, N_d^{(S)}(0)\right), \tag{2.3.3}$$

where for each type s ("s" for "**s**train") of the S total types, $N_d^{(s)}(0)$ is the random variable equalling the number of cells from type s in droplet d at time 0. **These are the observed data.** For the sake of brevity, often the subscript d is omitted, i.e. $N(0)$, $\vec{\mathbf{N}}(0)$, and $N^{(s)}(0)$ are used instead. This highlights how the random variables corresponding to the droplets are assumed (for the sake of simplicity) to have **independent and identical probability distributions**, and thus how most statements should not depend on the specific droplet d.

For each cell type s, let $f^{(s)}$ ("f" for "**f**requency") denote the relative abundance of cell type s in the sampling pool. In particular, $\sum_{s=1}^S f^{(s)} = 1$.

2.3.2 Statistical Motivation for Targeted Estimands

The treatment group $\mathcal{T}^{(s_1, s_2)}$ for the effect of strain s_1 on strain s_2 consists of droplets containing both strains. Likewise, the control group $\mathcal{C}^{(s_1, s_2)}$ for the effect of strain s_1 on

strain s_2 consists of droplets containing strain s_2 but not strain s_1. With these groups, we can quantify the effect of strain s_1 on strain s_2 by targeting an average treatment effect (ATE) or other similar causal estimand.

However, even when we can implement estimators for such causal estimands, sometimes their results may still not be scientifically reliable. For example, consider a situation where our treatment group $\mathcal{T}^{(s_1,s_2)}$ consists of only $\mathbf{T}^{(s_1,s_2)} = 5$ droplets and our control group $\mathcal{C}^{(s_1,s_2)}$ consists of $\mathbf{C}^{(s_1,s_2)} \approx 50,000$ droplets. This example may seem unrealistic, but it is actually optimistic compared to some of the size imbalances between the treatment and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ that one may expect to see in practice. Cf. section 2.C. Regardless of the causal estimand we choose to quantify such interaction effects, in order for our estimates to be scientifically useful we need to estimate their reliability. We need to estimate their statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$.

Of course (even for a fixed causal estimand) the statistical power $\mathscr{P}$ function will depend on the particular estimator, so at this level of generality we cannot complete such a calculation. However, a priori we do also know that the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ will always depend on the sizes $\mathbf{T}^{(s_1,s_2)}$, $\mathbf{C}^{(s_1,s_2)}$ of the treatment group and control group, respectively. Thus, to estimate statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$, we always need to know the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$, regardless of the particular estimator.

Unfortunately, due to the nature of the experiment, the scientist can never directly control the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$. The treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ are random variables. Therefore, for any estimator of any causal estimand, the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ will itself be a random variable. This is not inherently a "deal-breaker", but we do still need to constrain the possibilities for the distribution of the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$. At the very least, we would like to be able to compute the expected statistical power $\mathbb{E}[\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})]$.

Specifically, for a given estimator of a given causal estimand, the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ will be a deterministic function $\mathscr{P}$ of the treatment and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ (and possibly other factors which for simplicity we will assume to be known and fixed). If we know specific values $\mathbf{T}^{(s_1,s_2)} = T^{(s_1,s_2)}$, $\mathbf{C}^{(s_1,s_2)} = C^{(s_1,s_2)}$ for the treatment and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$, then we can evaluate this statistical power function $\mathscr{P}$ and get a deterministic value $\mathscr{P}(T^{(s_1,s_2)}, C^{(s_1,s_2)})$. If instead the treatment and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ are random variables, then when evaluating this statistical power function we get a new random variable $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ whose distribution is determined by the pushforward measure.

Obviously, without specifying the particular causal estimand and its particular estimator, we still cannot complete the calculation of the expectation $\mathbb{E}\left[\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})\right]$ of a random statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$. Nevertheless, for any given estimator of any given causal estimand, the distribution of the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ will be determined by the distribution of the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$. Thus, for any estimator of any causal estimand, as a prerequisite for being able to estimate

expected statistical power $\mathbb{E}\big[\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})\big]$, we must first understand the distribution of the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$.

Note that the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ can be understood as sums of indicator random variables:

$$\mathbf{T}^{(s_1,s_2)} = \sum_{d \in [D]} \mathbb{I}\big\{d\text{'th droplet in } \mathcal{T}^{(s_1,s_2)}\big\}, \quad \mathbf{C}^{(s_1,s_2)} = \sum_{d \in [D]} \mathbb{I}\big\{d\text{'th droplet in } \mathcal{C}^{(s_1,s_2)}\big\}.$$
(2.3.4)

To completely characterize the distribution of indicator random variables it suffices to know their expectations, i.e. the probabilities:

$$\mathbb{P}\big(d\text{'th droplet in } \mathcal{T}^{(s_1,s_2)}\big), \qquad \mathbb{P}\big(d\text{'th droplet in } \mathcal{C}^{(s_1,s_2)}\big).$$
(2.3.5)

If we assume the droplets are identically distributed, then linearity of expectation allows us to compute the expected treatment and control group sizes:

$$\mathbb{E}\big[\mathbf{T}^{(s_1,s_2)}\big] = D \cdot \mathbb{P}\big(d\text{'th droplet in } \mathcal{T}^{(s_1,s_2)}\big), \quad \mathbb{E}\big[\mathbf{C}^{(s_1,s_2)}\big] = D \cdot \mathbb{P}\big(d\text{'th droplet in } \mathcal{C}^{(s_1,s_2)}\big).$$
(2.3.6)

This is already enough to compute $\mathscr{P}\big(\mathbb{E}\big[\mathbf{T}^{(s_1,s_2)}\big], \mathbb{E}\big[\mathbf{C}^{(s_1,s_2)}\big]\big)$, a quantity that we can use as a (potentially suboptimal[12]) stand-in for the expected statistical power $\mathbb{E}\big[\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})\big]$ in situations where we don't want to make additional assumptions about the distributions of the droplets.

However, if we additionally assume that the droplets are mutually independent, then we can compute the distributions of the treatment and control group sizes as marginals of a Multinomial distribution[13]. Given the statistical power function $\mathscr{P}$, i.e. choice of a particular causal estimand and estimator thereof, this would then in principle[14] be enough information to compute the distribution of the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ or any summary statistic thereof.

One might object to treating the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ as a random variable. Certainly, after the experiment has been run, the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ will be known, allowing us to condition on their observed values $T^{(s_1,s_2)}, C^{(s_1,s_2)}$ and "restore determinism" by evaluating $\mathscr{P}(T^{(s_1,s_2)}, C^{(s_1,s_2)})$. However, this is insufficient to help the scientist before running the experiment. The scientist has *no* direct control over what the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ will be.

However, the distribution of the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$ should depend on deterministic factors that the scientist can directly control, like the total number of droplets D, the expected number of cells per droplet λ, and the relative abundances $f^{(s)}$ of the various strains s in the sampling population. The scientist would like to

---

[12] Cf. the "arithmetic" estimands from section 7.E.

[13] Cf. sections 5.2.3 and 5.2.4.

[14] In practice, given realistic problem sizes, evaluating this exact formula would most likely be intractable, such that one would probably also need to use additional heuristics and approximations. The choice of those heuristics and approximations would depend on the particular causal estimand and estimator thereof, and thus is outside the scope of this discussion.

plan the experiment (*before* running it) in a way that allows them to (most likely) achieve certain thresholds of statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ for certain interactions of strains. Therefore the best way to help the scientist is to predict which values of statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ they will most likely be able to achieve as a function of the deterministic factors that the scientist can directly control. From that perspective it is inescapable that

(1) the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ must be treated as random,

(2) the most basic prerequisite for constraining the distribution of the statistical power $\mathscr{P}(\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)})$ is to characterize the distributions of the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}, \mathbf{C}^{(s_1,s_2)}$, and

(3) (if possible) the distributions of the treatment group and control group sizes $\mathbf{T}^{(s_1,s_2)}$, $\mathbf{C}^{(s_1,s_2)}$ should be described as functions of the deterministic factors that the scientist can directly control.

There is no need to incubate the droplets in order to estimate the probabilities (2.3.5). Indeed, it should be possible to estimate the probabilities using a much smaller number of total droplets D than what might be required in the final experiment to achieve certain (expected) statistical power thresholds. Therefore I propose that the scientist, to be able to design the experiment in a way that allows them achieve certain (expected) statistical power thresholds, first run a "$t = 0$" version of the experiment that skips incubating the droplets and uses a smaller number of total droplets D. This both (i) maximizes the usefulness of the data for estimating the distributions of the treatment group and control group sizes (because it should mitigate artifacts resulting from microbial growth and competition, e.g. censoring) and (ii) decreases the cost (in money and time) of this "pre-experiment" to the scientist.

2.3.3 Non-Parametric Statistical Model

The non-parametric statistical model $\mathcal{M}$ containing the distribution of $\vec{\mathbf{N}}(0)$ for any of the D independent and identically distributed droplets is

$$\mathcal{M} = \left\{ \mathcal{P} : \mathcal{P} \text{ is supported on } \mathbb{N}_{\geq 0}^S \right\} . \tag{2.3.7}$$

2.3.4 Target Parameter Definitions

The below sections define two different classes of target parameters. Cf. section 5.2.1 for more details.

2.3.4.1 "Gluttonous" Target Parameters

The treatment group for inferring the effect of type s_1 on the growth of type s_2 corresponds to droplets where s_1 and s_2 co-occur. The set of all droplets satisfying this condition is

$$\mathcal{T}_g^{(s_1,s_2)} := \left\{ d \in [D] : \mathrm{supp}(\vec{\mathbf{N}}_d(0)) \supseteq \{s_1, s_2\} \right\} . \tag{2.3.8}$$

This is called the "**gluttonous**" **treatment** group because *all* droplets where s_1 and s_2 co-occur are included. Cf. figure 5.1. Because of the i.i.d. assumption, this leads to the statistical queries[15] for all $s_1, s_2 \in [S]$, $s_1 \neq s_2$:

$$\begin{aligned}
\Psi_{\mathcal{T},g}^{(s_1,s_2)} &: \mathcal{M} \to [0,1], \\
\Psi_{\mathcal{T},g}^{(s_1,s_2)} &: \mathcal{P} \mapsto \mathbb{P}_{\mathcal{P}}\left(\mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right).
\end{aligned} \tag{2.3.9}$$

The control group for inferring the effect of type s_1 on the growth of type s_2 corresponds to droplets where s_2 occurs but s_1 does not. The set of all droplets satisfying this condition is

$$\mathcal{C}_g^{(s_1,s_2)} := \left\{d \in [D] : \mathrm{supp}(\vec{\mathbf{N}}_d(0)) \supseteq \{s_2\}, \; s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}_d(0))\right\}. \tag{2.3.10}$$

This is called the "**gluttonous**" **control** group because *all* droplets where s_2 occurs but s_1 does not are included. Cf. figure 5.1. Because of the i.i.d. assumption, this leads to the statistical queries for all $s_1, s_2 \in [S]$, $s_1 \neq s_2$:

$$\begin{aligned}
\Psi_{\mathcal{C},g}^{(s_1,s_2)} &: \mathcal{M} \to [0,1], \\
\Psi_{\mathcal{C},g}^{(s_1,s_2)} &: \mathcal{P} \mapsto \mathbb{P}_{\mathcal{P}}\left(\mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}, \; s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\right).
\end{aligned} \tag{2.3.11}$$

These definitions place no constraints on the presence or absence of types $s \in [S]$ besides s_1 and s_2. This can be favorable for example when studying very rare types, for which there may be no or very few droplets containing s_1 and s_2 only, but some containing s_1 and s_2 along with other types. These definitions give us the best possible chance of avoiding "data starvation".

2.3.4.2 "Picky" Target Parameters

The above "gluttonous" definitions correspond to more possible combinations of cell types than we probably would have included when studying the effect of type s_1 on the growth of type s_2 by manually plating combinations of cell types. This motivates the following definitions.

The set of droplets where s_1 and s_2 co-occur in a way corresponding to the combination of cell types we would use if we were manually plating combinations of cell types is

$$\mathcal{T}_p^{(s_1,s_2)} := \left\{d \in [D] : \mathrm{supp}(\vec{\mathbf{N}}_d(0)) = \{s_1, s_2\}\right\}. \tag{2.3.12}$$

This is called the "**picky**" **treatment** group because droplets where s_1 and s_2 co-occur are included *only when all other types are absent*. Cf. figure 5.1. Because of the i.i.d. assumption, this leads to the statistical queries[16] for all $s_1, s_2 \in [S]$, $s_1 \neq s_2$:

$$\begin{aligned}
\Psi_{\mathcal{T},p}^{(s_1,s_2)} &: \mathcal{M} \to [0,1], \\
\Psi_{\mathcal{T},p}^{(s_1,s_2)} &: \mathcal{P} \mapsto \mathbb{P}_{\mathcal{P}}\left(\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\right).
\end{aligned} \tag{2.3.13}$$

---

[15]Cf. section 2.A, in particular section 2.A.1 on the notation for functions used herein.

[16]Cf. section 2.A, in particular section 2.A.1 on the notation for functions used herein.

The set of droplets where *only* s_2 occurs, and all other types are absent, corresponding to the combination of cell types we would use if we were manually plating combinations of cell types, is

$$\mathcal{C}_p^{(s_1,s_2)} := \left\{ d \in [D] : \operatorname{supp}(\vec{\mathbf{N}}_d(0)) = \{s_2\} \right\} . \tag{2.3.14}$$

This is called the "**picky**" **control** group because droplets where s_2 occurs but s_1 does not occur are included *only when all other types are absent.* It is not enough for s_1 alone to be absent. Cf. figure 5.1. Because of the i.i.d. assumption, this leads to the statistical queries for all $s_1, s_2 \in [S]$, $s_1 \neq s_2$:

$$\begin{aligned} \Psi_{\mathcal{C},p}^{(s_1,s_2)} &: \mathcal{M} \to [0,1] , \\ \Psi_{\mathcal{C},p}^{(s_1,s_2)} &: \mathcal{P} \mapsto \mathbb{P}_{\mathcal{P}}\left(\operatorname{supp}(\vec{\mathbf{N}}(0)) = \{s_2\} \right) . \end{aligned} \tag{2.3.15}$$

Note that unlike any of the other group definitions, this one is shared between all questions investigating effects on the growth of strain s_2.

Unlike the gluttonous groups, the picky groups *do* place constraints on the presence or absence of types $s \in [S]$ besides s_1 and s_2. For both the treatments and the controls, for "picky" groups all types besides s_1 or s_2 must have zero counts. Every picky group is a subset of its gluttonous counterpart.

The picky definitions can be favorable when we have plenty of droplets from which to make estimates. They reduce the possibility of confounding effects on the growth of type s_2 that could be caused by types that are not type s_1. On the other hand, in instances where there are only very few or no droplets without "extra" types, for example for very rare types for which there are few droplets containing the type at all, the picky definitions could lead to "data starvation".

2.3.5 Why Empirical Estimates are not Useful

The empirical estimates for these probabilities are inefficient for several reasons. What connects these reasons is the fact that the empirical estimates do not provide any way to connect these estimands with each other. Thus information about one estimand, which should ideally be applicable for learning about related estimands, is wasted. More specifically, note that we have a combinatorial explosion in the number of estimands as a function of the number of strains S. This is at least quadratic if we only consider the subsets of S corresponding to the groups defined above, and exponential if we want to consider the probabilities associated with arbitrary subsets of $[S]$ (e.g. for the purpose of investigating higher-order interactions). Cf. section 7.7.3, which discusses a related issue. This is in contrast to the number of parameters in the working models which (if including the relative abundances $f^{(s)}$) only grow linearly with the number of strains S. Thus using the empirical estimates leads to not only exponential increases in "description length" or "information" required to characterize the initial distribution of droplets, but also to exponential increases in computational runtime or

algorithmic complexity for computing estimates of all estimands. The exponential number of parameters has practical implications. Cf. the problem motivating [TRKS$^+$17].

Related to this, especially for the treatment groups as well as types s with small relative abundances $f^{(s)}$, even when D is large, the expected counts can still be very small. In practice this leads to positivity violations, where the empirical estimates force us to state that certain unobserved events happen with zero probability when in principle information from related events should allow us make some small but non-zero estimate for their probabilities. This, combined with the combinatorial explosion, means that we would have to use very large values of D to make useful (i.e. non-zero and low-variance) estimates for all probabilities. Again, cf. the related discussion in section 7.7.3. This is very inefficient because we would prefer to use a smaller or medium sized D for the initial trial run (the "$t = 0$ pre-experiment", section 2.3). Fitting a model for predicting data throughput is simply a prerequisite for estimating the power we will have to answer certain questions as a function of D, rather than the main question itself. Cf. again section 2.3.2.

Similarly, the empirical estimates do not allow us to describe the probabilities in terms of parameters such as the relative abundances $f^{(s)}$ or the mean number of cells per droplet. That means that, e.g. if we slightly modify our experiment by swapping out two strains with five different strains (which will obviously have different relative abundances), then when using empirical estimates we will have to re-estimate everything from scratch. In contrast, to the extent that parameters like the density concentration ζ_D or compositional concentration ζ_C are believed to correspond to "intrinsic features" of the experimental apparatus itself, we could reasonably believe that we could re-use estimates for them even if we change the number or relative abundances of strains or the mean number of cells per droplet.

2.3.6 Default Working Model: hPoMu

The "hierarchical Poisson Multinomial" (hPoMu) distributions $\mathcal{M}_{\text{hPoMu}} \subset \mathcal{M}$ are hierarchical distributions. The total number of cells in the droplet has a Poisson distribution, while (*conditional upon the total number of cells*) the numbers of cells belonging to the multiple types have a joint Multinomial distribution. Cf. figure 3.2. Both choices of distribution correspond to assuming uniform and perfectly homogeneous sampling for each droplet. Cf. section 3.3.2 for two detailed arguments. The likelihood for the hPoMu distributions equals

$$\mathbb{P}_{\text{hPoMu}(\lambda, \vec{f})}\left(\vec{\mathbf{N}}(0) = \vec{\mathbf{n}}\right) = \mathbb{P}_{\text{hPoMu}(\lambda, \vec{f})}(N(0) = n) \cdot \mathbb{P}_{\text{hPoMu}(\lambda, \vec{f})}\left(\vec{\mathbf{N}}(0) = \vec{\mathbf{n}} \,\Big|\, N(0) = n\right)$$

$$= \frac{e^{-\lambda}\lambda^n}{n!} \cdot \binom{n}{n^{(1)} \cdots n^{(S)}} \prod_{s=1}^{S} (f^{(s)})^{n^{(s)}} .$$

$$(2.3.16)$$

An alternative form of the likelihood of hPoMu reveals that its marginal distributions are mutually independent Poisson distributions:

$$\mathbb{P}_{\text{hPoMu}(\lambda,\vec{f})}\left(\vec{\mathbf{N}}(0) = \vec{\mathbf{n}}\right) = \prod_{s=1}^{S} \frac{e^{-f^{(s)}\lambda}(f^{(s)}\lambda)^{n^{(s)}}}{(n^{(s)})!}\,. \tag{2.3.17}$$

Heuristically speaking, equation (2.3.17) says that during the formation of each droplet the expected number of cells, or "sampling rate", λ is evenly spread out among the multiple types according to their frequencies in the population.

2.3.6.1 Projected Target Parameters Under hPoMu Working Model

Let $\iota_{\text{hPoMu}} : \mathcal{M}_{hPoMu} \to \mathcal{M}$, $\iota_{\text{hPoMu}} : \mathcal{P} \mapsto \mathcal{P}$ denote the inclusion map[17] $\mathcal{M}_{\text{hPoMu}} \xrightarrow{\iota_{\text{hPoMu}}} \mathcal{M}$. Suppose $\Pi_{\text{hPoMu}} : \mathcal{M} \to \mathcal{M}_{\text{hPoMu}}$ is a projection[18]:

$$\Pi_{\text{hPoMu}} : \mathcal{P} \mapsto \operatorname*{argmin}_{\tilde{\mathcal{P}} \in \mathcal{M}_{\text{hPoMu}}} \int_{\mathbb{N}_{\geq 0}^S} L(\tilde{\mathcal{P}}(\vec{n})) d\mathcal{P}(\vec{n})\,, \tag{2.3.18}$$

for a given loss function L, e.g. $L : p \mapsto -\log(p)$, onto the working parametric submodel $\mathcal{M}_{\text{hPoMu}}$ from the nonparametric model $\mathcal{M}$. In particular one should have the relationships that $\Pi_{\text{hPoMu}} \circ \iota_{\text{hPoMu}} = \text{Id}_{\text{hPoMu}} : \mathcal{M}_{\text{hPoMu}} \to \mathcal{M}_{\text{hPoMu}}$ and that $\iota_{\text{hPoMu}} \circ \Pi_{\text{hPoMu}} : \mathcal{M} \to \mathcal{M}$ is idempotent.

Given the projection of a distribution $\mathcal{P}$ from the non-parametric model $\mathcal{M}$ onto the working parametric submodel $\mathcal{M}_{\text{hPoMu}}$, we can now evaluate any of the non-parametrically defined parameters (2.3.9), (2.3.11), (2.3.13), (2.3.15) from section 2.3.4 at the projection, defining a new set of target parameters. This new set of target parameters corresponding to $\mathcal{M}_{\text{hPoMu}}$ is defined below.

For a given $\mathcal{P} \in \mathcal{M}$, let $\iota_{\text{hPoMu}}(\Pi_{\text{hPoMu}}(\mathcal{P})) =: \text{hPoMu}(\lambda_{\mathcal{P}}, \vec{f}_{\mathcal{P}}) \in \mathcal{M}_{\text{hPoMu}}$. The projected target parameters for the gluttonous treatment groups $\Psi_{\mathcal{T},g,\text{hPoMu}}^{(s_1,s_2)} := \Psi_{\mathcal{T},g}^{(s_1,s_2)} \circ \iota_{\text{hPoMu}} \circ \Pi_{\text{hPoMu}}$ are defined as (cf. section 5.2.4)

$$\Psi_{\mathcal{T},g,\text{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$\begin{aligned} \Psi_{\mathcal{T},g,\text{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\text{hPoMu}}(\Pi_{\text{hPoMu}}(\mathcal{P}))}\left(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right) \\ &= \quad \mathbb{P}_{\text{hPoMu}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}})}\left(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right) \\ &= \quad \left(1 - e^{-f_{\mathcal{P}}^{(s_1)}\lambda_{\mathcal{P}}}\right) \cdot \left(1 - e^{-f_{\mathcal{P}}^{(s_2)}\lambda_{\mathcal{P}}}\right), \end{aligned} \tag{2.3.19}$$

---

[17]Cf. section 2.A, in particular the last part of section 2.A.2.

[18]Cf. section 2.A, in particular section 2.A.1 on the notation for functions used herein.

the projected target parameters for the gluttonous control groups $\Psi_{\mathcal{C},g,\mathrm{hPoMu}}^{(s_1,s_2)} := \Psi_{\mathcal{C},g}^{(s_1,s_2)} \circ$
$\iota_{\mathrm{hPoMu}} \circ \Pi_{\mathrm{hPoMu}}$ are defined as (cf. section 5.2.4)

$$\Psi_{\mathcal{C},g,\mathrm{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi_{\mathcal{C},g,\mathrm{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\mathrm{hPoMu}}(\Pi_{\mathrm{hPoMu}}(\mathcal{P}))}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}\,,\, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\Big) \\
&= \quad \mathbb{P}_{\mathrm{hPoMu}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}})}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}\,,\, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\Big) \\
&= \quad e^{-f_{\mathcal{P}}^{(s_1)}\lambda_{\mathcal{P}}} \cdot \big(1 - e^{-f_{\mathcal{P}}^{(s_2)}\lambda_{\mathcal{P}}}\big)\,,
\end{aligned}
\tag{2.3.20}
$$

the projected target parameters for the picky treatment groups $\Psi_{\mathcal{T},p,\mathrm{hPoMu}}^{(s_1,s_2)} := \Psi_{\mathcal{T},p}^{(s_1,s_2)} \circ \iota_{\mathrm{hPoMu}} \circ$
Π_{hPoMu} are defined as (cf. section 5.2.3)

$$\Psi_{\mathcal{T},p,\mathrm{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi_{\mathcal{T},p,\mathrm{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\mathrm{hPoMu}}(\Pi_{\mathrm{hPoMu}}(\mathcal{P}))}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\Big) \\
&= \quad \mathbb{P}_{\mathrm{hPoMu}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}})}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\Big) \\
&= \quad (1 - e^{-f_{\mathcal{P}}^{(s_1)}\lambda_{\mathcal{P}}})(1 - e^{-f_{\mathcal{P}}^{(s_2)}\lambda_{\mathcal{P}}}) \cdot e^{-(1-(f_{\mathcal{P}}^{(s_1)}+f_{\mathcal{P}}^{(s_2)}))\lambda_{\mathcal{P}}}\,,
\end{aligned}
\tag{2.3.21}
$$

and the projected target parameters for the picky control groups $\Psi_{\mathcal{C},p,\mathrm{hPoMu}}^{(s_1,s_2)} := \Psi_{\mathcal{C},p}^{(s_1,s_2)} \circ$
$\iota_{\mathrm{hPoMu}} \circ \Pi_{\mathrm{hPoMu}}$ are defined as (cf. section 5.2.3)

$$\Psi_{\mathcal{C},p,\mathrm{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi_{\mathcal{C},p,\mathrm{hPoMu}}^{(s_1,s_2)} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\mathrm{hPoMu}}(\Pi_{\mathrm{hPoMu}}(\mathcal{P}))}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\Big) \\
&= \quad \mathbb{P}_{\mathrm{hPoMu}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}})}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\Big) \\
&= \quad (1 - e^{-f_{\mathcal{P}}^{(s_2)}\lambda_{\mathcal{P}}}) \cdot e^{-(1-f_{\mathcal{P}}^{(s_2)})\lambda_{\mathcal{P}}}\,.
\end{aligned}
\tag{2.3.22}
$$

To estimate these, we just need to estimate $\lambda_{\mathcal{P}}$ and $\vec{f}_{\mathcal{P}}$ and then use the corresponding plugin estimators. Estimating $\lambda_{\mathcal{P}}$ and $\vec{f}_{\mathcal{P}}$ also gives us plugin estimators for the probabilities associated with *all* 2^S subsets of $[S]$. Thus "information is shared" between estimates for the probabilities of related events. Moreover, as long as all estimates of the $f_{\mathcal{P}}^{(s)}$ are nonzero, all estimated probabilities will also all be nonzero. Therefore these estimates should both be "smoother" and more efficient than the empirical estimates, especially when S is large (corresponding to some of the $f_{\mathcal{P}}^{(s)}$ being small). The large S regime would substantially advance the state of the art in microbial ecology, and thus is of particular interest.

2.3.6.2 Assumptions Justifying Default Working Model

Both of the two derivations justifying hPoMu as the default working model, cf. section 3.3.2, make the following implicit assumptions (among others):

- For *every* type s, the number of cells in the population which belong to type s is very large.

- Each cell (regardless of its type) has the same small probability of ending up in the droplet as any other cell.

- Whether a given cell (regardless of its type) ends up in the droplet is completely independent of what happens to any other cell.

The first can be thought of as stating that the population size for each type is effectively infinite. The second and third can together be thought of as stating that the sampling pool is perfectly homogeneous.

2.3.7 Failures of Default Working Model Assumptions

While hPoMu is a sensible working model to start with, its implicit assumptions failing to be realistic could undermine its usefulness in practice.

2.3.7.1 Populations are Finite

We assumed that *for each individual type s* the number of cells is "effectively infinite". For "rare" types it is a priori unclear whether this assumption is reasonable. However in practice, even for types with relative abundance as low as 0.01%, taking the finiteness of the populations into account does not lead to substantially different predictions than the default hPoMu working model. (Cf. sections 4.4.1 and 5.4.2.1 where this claim is substantiated.)

2.3.7.2 Density Heterogeneity

Different sections of the sampling pool could have average numbers of cells per unit volume that differ from the average number of cells per unit volume for the entire sampling pool (i.e. the total number of cells divided by the volume of the sampling pool). Herein I call this phenomenon "density heterogeneity". Higher density heterogeneity implies more variance in the numbers of cells per unit volume throughout the sampling pool. Cf. figure 3.4.

The default Poisson working model for the total number of cells per droplet can not account for additional variance that density heterogeneity might cause. Its variance is the smallest possible value that could hypothetically exist in practice, which seems unrealistic.

Because this describes extra variance in the total number of cells, regardless of their types, density heterogeneity is relevant even when there is only one type. Thus density heterogeneity is relevant even for the single-cell[19], single type experiments which have already been intensively developed.

---

[19]The goal for these experiments is to have one cell per droplet, whence the name "single-cell". However, because the droplet formation process is random, in practice some droplets in "single-cell" experiments may still end up with more than one cell.

2.3.7.3 Compositional Heterogeneity

The relative abundance of any given cell type s across the entire sampling pool is a fixed value $f^{(s)}$. Different sections of the sampling pool could have relative abundances of the cell types that differ from those for the entire sampling pool. Herein I call this phenomenon "compositional heterogeneity". Higher compositional heterogeneity implies more variance in the relative abundances for the cell types throughout the sampling pool. Cf. figure 3.5.

The default Multinomial working model for the relative abundances of the cell types per droplet (*conditional upon the total number of cells*) does not account for additional variance that would be introduced by compositional heterogeneity. Its variance is the smallest possible value that could hypothetically exist in practice, which is seems unrealistic.

Compositional heterogeneity is relevant only for multi-cell, multi-type experiments. Such experiments are newer and still less common than single-cell, single type experiments. It seems that compositional heterogeneity may not have been identified before as a potential issue affecting the data throughput of droplet microfluidics experiments.

2.3.8 Working Model to Address Assumption Failures: ghNBDM

Any of the three assumptions from section 2.3.6.2 can be unrealistic in practice. As mentioned before, even for extremely rare strains, taking the finiteness of the populations into account (section 2.3.7.1) does not lead to predictions substantially different than those made by the default hPoMu model. On the other hand, models for which the perfect homogeneity assumption is false *can* lead to substantially different predictions. This section introduces a generalization of the hPoMu working model that makes fewer homogeneity assumptions than hPoMu does. In particular, both nonzero density heterogeneity (section 2.3.7.2) and nonzero compositional heterogeneity (section 2.3.7.3) are allowed.

The "generalized hierarchical Negative Binomial Dirichlet-Multinomial" (ghNBDM) working model $\mathcal{M}_{\text{ghNBDM}} \subset \mathcal{M}$ is a family of hierarchical distributions. The number of cells in the droplet has a Negative Binomial distribution with "density concentration" parameter $\zeta_D S$, while (*conditional upon the total number of cells*) the numbers of cells belonging to the multiple types have a joint Dirichlet-Multinomial distribution with "compositional concentration" parameter $\zeta_C S$. The likelihood for this parametric submodel $\mathcal{M}_{\text{ghNBDM}}$ is

$$
\mathbb{P}_{\text{ghNBDM}(\lambda,\vec{f},\zeta_D,\zeta_C)}\left(\vec{\mathbf{N}}(0) = \vec{\mathbf{n}}\right)
$$

$$
= \quad \mathbb{P}_{\text{ghNBDM}(\lambda,\vec{f},\zeta_D,\zeta_C)}(N(0) = n) \quad \cdot \mathbb{P}_{\text{ghNBDM}(\lambda,\vec{f},\zeta_D,\zeta_C)}\left(\vec{\mathbf{N}}(0) = \vec{\mathbf{n}} \;\middle|\; N(0) = n\right)
$$

$$
= \quad \frac{\Gamma(\zeta_D S + n)}{\Gamma(\zeta_D S)} \cdot \frac{(\zeta_D S)^{\zeta_D S}}{(\zeta_D S + \lambda)^{\zeta_D S + n}} \cdot \frac{\lambda^n}{n!} \cdot \frac{\Gamma(\zeta_C S)}{\Gamma(\zeta_C S + n)} \cdot \left[\prod_{s=1}^{S} \frac{\Gamma(\zeta_C S f^{(s)} + n^{(s)})}{\Gamma\left(\zeta_C S f^{(s)}\right)}\right] \cdot \binom{n}{n^{(1)} \cdots n^{(S)}} \cdot
$$

$$
\tag{2.3.23}
$$

In general the marginal distributions of members of the ghNBDM family $\mathcal{M}_{\text{ghNBDM}} \subset \mathcal{M}$ have non-zero cross-covariance. Thus the entries $N^{(s)}(0)$ of the random vector $\vec{\mathbf{N}}(0)$ will in

general *not* be mutually independent. This is unlike the default working model (hPoMu) $\mathcal{M}_{\text{hPoMu}} \subset \mathcal{M}_{\text{ghNBDM}} \subset \mathcal{M}$, for which the $N^{(s)}(0)$ are mutually independent, cf. again section 2.3.6, particularly equation (2.3.17).

Notice how this working model has two additional parameters compared to the hPoMu working model, namely the density concentration ζ_D parameter and the compositional concentration ζ_C parameter. Higher values of the density concentration ζ_D parameter correspond to lower density heterogeneity (section 2.3.7.2). As $\zeta_D \to \infty$ the total number of cells in the droplet approaches a Poisson distribution. Cf. section 3.7.2.2. Similarly, higher values of the compositional concentration ζ_C parameter correspond to lower compositional heterogeneity (section 2.3.7.3). As $\zeta_C \to \infty$ the joint distribution of the numbers of cells belonging to the multiple types approaches a Multinomial distribution. Cf. section 3.6.2.2. The hPoMu model is approached in the limit as both $\zeta_D \to \infty$ and $\zeta_C \to \infty$ jointly. Cf. figure 3.6 and section 3.8.1. For details about the calculations demonstrating how the somewhat complicated expressions involving the Γ function do indeed generalize the likelihoods of the Poisson and Multinomial distributions, cf. sections 3.7.2, 3.6.2, and 3.A. At a high level, the idea is to use Stirling's approximation, in particular the formulations given by [Rob55] and [Gor94].

Projected Target Parameters Under ghNBDM Working Model

This is similar to what was done in section 2.3.6.1. See section 2.B.1 for details.

2.3.9   Intermediate Choice: hNBDM Working Model

Given estimates of $\lambda_{\mathcal{P}}$, $\vec{f}_{\mathcal{P}}$, $\zeta_{D,\mathcal{P}}$, and $\zeta_{C,\mathcal{P}}$, plugin estimators for the above target parameters (2.B.3), (2.B.4), (2.B.5), (2.B.6) can be computed to arbitrary accuracy, e.g. by truncating terms of the series expansions. Nevertheless, it is at best tedious to derive explicit bounds on the accuracy as a function of the number of terms in the truncated series expansions. An alternative is to use the "hierarchical Negative Binomial Dirichlet-Multinomial" (hNBDM) working model $\mathcal{M}_{\text{hNBDM}} \subset \mathcal{M}_{\text{ghNBDM}} \subset \mathcal{M}$.

The hNBDM working model is intermediate in complexity and flexibility between the ghNBDM and hPoMu working models, indeed $\mathcal{M}_{\text{hPoMu}} \subset \mathcal{M}_{\text{hNBDM}} \subset \mathcal{M}_{\text{ghNBDM}}$. Distributions in the hNBDM working model are distributions in the ghNBDM working model such that the density concentration ζ_D and compositional concentration ζ_C parameters are equal, $\zeta_D = \zeta_C =: \zeta$. Hence the hNBDM working model has one more parameter than the hPoMu working model, but also one less parameter than the ghNBDM working model.

The real "magic" of the hNBDM working model happens in its likelihood: when $\zeta_D = \zeta_C$, the resulting cancellations and simplifications in the ghNBDM likelihood cause it to factorize

over its marginal distributions:

$$
\begin{aligned}
&\mathbb{P}_{\mathrm{hNBDM}(\lambda,\vec{f},\zeta)}\Big(\vec{\mathbf{N}}(0) = \vec{\mathbf{n}}\Big) \\
&= \prod_{s=1}^{S} \frac{\Gamma(\zeta S f^{(s)} + n^{(s)})}{\Gamma(\zeta S f^{(s)})} \frac{(\zeta S f^{(s)})^{\zeta S f^{(s)}}}{(\zeta S f^{(s)} + f^{(s)}\lambda)^{\zeta S f^{(s)} + n^{(s)}}} \frac{(f^{(s)}\lambda)^{n^{(s)}}}{(n^{(s)})!} \, .
\end{aligned}
\tag{2.3.24}
$$

In other words, the marginal distributions are mutually independent Negative Binomial distributions. The "sampling rate" λ is proportionately "spread out" across the marginals according to the frequencies $f^{(s)}$, exactly analogous to what occurs for the hPoMu working model. Indeed, the hPoMu distributions can be considered the "boundary" of the hNBDM model reached by the limit $\zeta \to \infty$.

Projected Target Parameters Under hNBDM Working Model

This is similar to what was done in section 2.3.6.1. See section 2.B.2 for details.

2.4 Conclusion

Findings and Contributions   The data from the noisy dynamical systems corresponding to each droplet can be understood as censored observations of a multivariate Markov process. The statistical understanding of a droplet's initial state is crucial to overcoming the main limitation of these experiments, the uncontrolled assignment of microbes to droplets.

Practical Implications   Giving explicit descriptions of the intended underlying statistical model has several benefits. By formalizing intuitive insights, like the division of the experiment into three phases, it allows us use those insights as the basis for data analysis. It helps to clarify the similarities and differences between this problem and other problem, as well the similarities and differences between methods used for this problem and methods used for other problems. It makes explicit which assumptions we are making when analyzing the data, so that we might later question those assumptions if necessary. It guides us to understanding clearly what the most challenging aspects of the data analysis are. In short, it gives us a starting point that anchors everything else we might do.

Next Steps and Open Questions   There are several different choices we could have made when positing the explicit statistical models given in this chapter. It remains to be fully investigated what the consequences of such different choices would be. Reasonable possible choices, and their consequences, are explained in the next part. Later, methods for inferring interactions that would apply to observations produced by these models are considered.

Chapter 2.   Appendices

Appendix 2.A   Conventions and Notation for Functions

Section 2.A.1 discusses the notation used for functions. Section 2.A.2 discusses the convention used for function composition.

2.A.1   Function Notation

Herein I use a definition of function that requires not only (i) a rule of assignment, but also (ii) a specified domain (source) and codomain (target). A notion of function requiring only (i) corresponds more to "classical" mathematics, and in some sources is called an "intensional" definition of function[Sel13, section 1.1]. In contrast, a notion of function requiring (i) and (ii) corresponds more to "modern" mathematics, and in some sources is called an "extensional" definition of function[Sel13, section 1.1]. Herein I only use the "extensional" notion[20].

Given a function f whose domain is a set X and whose codomain is a set Y, thus sending every element $x \in X$ to a unique element $f(x) \in Y$, I write

$$\begin{aligned} f &: X \to Y\,, \\ f &: x \mapsto f(x)\,. \end{aligned} \tag{2.A.1}$$

The first line, $f : X \to Y$, indicates the domain and codomain of the function. The second line (using "anonymous function" notation from lambda calculus), $f : x \mapsto f(x)$, indicates the rule of assignment of the function.

In particular, I use the convention that for, two functions $f_1 : X_1 \to Y_1$ and $f_2 : X_2 \to Y_2$ to be equal, not only must they have the same rule of assignment, but they also must have the same domain and codomain. For example, the function $f_1 : x \mapsto x$ and $f_2 : x \mapsto x$ are considered different when $f_1 : \mathbb{N} \to \mathbb{R}$ and $f_2 : \mathbb{N} \to \mathbb{N}$, even though they both have the same rule of assignment and same domain, because their codomains are different. With this convention, a rule of assignment is insufficient to specify a function.

---

[20] In programming language terminology, the "intensional" notion is "untyped", whereas the "extensional" notion is "typed". Thus one motivation for using the latter notion more often in "modern" mathematics is that the resulting statements are more "type safe".

2.A.2 Conventions for Function Composition

There are at least two different possible conventions for when, given two functions

$$f_1 : X_1 \to Y_1 \qquad \text{and} \qquad f_2 : X_2 \to Y_2 \, ,$$

we say that the composition of the two functions, $f_2 \circ f_1 : X_1 \to Y_2$, is defined:

(i) whenever $Y_1 \subseteq X_2$,

(ii) whenever $Y_1 = X_2$.

Using either convention, the rule of assignment of $f_2 \circ f_1$ is $x_1 \mapsto f_2(f_1(x_1))$. However, with the first convention, there are potentially infinitely many functions f_1 such that $f_2 \circ f_1 : X_1 \to Y_2$ is defined and has the same rule of assignment, each corresponding to a distinct $Y_1 \subseteq X_2$. In particular, the first convention corresponds to defining a function only in terms of a rule of assignment, and not requiring that a specific domain and codomain be specified.

Hence the first convention for composition of functions is ill-suited for the definition of function explained in section 2.A.1 above. It is incompatible with the demand that not only the rule of assignment of a function must be specified but also its domain and codomain.

The second convention for composition of functions is compatible with the definition of function explained in section 2.A.1 above. The second convention requires a specific codomain for f_1 (namely exactly the domain of f_2) is specified in order for the composition $f_2 \circ f_1$ to even be defined. This convention is less ambiguous and is the convention I will use.

The choice between the two conventions can have practical consequences. For example, when using the first convention, the composition of two surjective (a.k.a. onto) functions need not again be surjective, and the composition of two bijections need not again be a bijection. This can be subtle and confusing and lead to hard to detect mistakes. On the other hand, when using the second convention, the composition of two surjective functions is always surjective, and the composition of two bijections is always a bijection.

For example, say that $f_1 : X_1 \to Y_1$ and $f_2 : X_2 \to Y_2$ are both surjective, with $Y_1 \subsetneq X_2$. If the restriction of f_2 to Y_1 is not surjective (onto) Y_2, then, when using the first convention, the composition $f_2 \circ f_1$ will not be surjective, even though f_1 and f_2 are both surjective. However, when using the second convention, the composition $f_2 \circ f_1$ is not even defined. Instead, we first need to define the inclusion function $\iota : Y_1 \hookrightarrow X_2$, $\iota : y_1 \mapsto y_1$, which is clearly injective but not surjective because $Y_1 \subsetneq X_2$ by assumption. Then the function $X_1 \to Y_2$ we are interested in is $f_2 \circ \iota \circ f_1$. It becomes clear then why the composite function is not surjective. Although f_1 and f_2 are both surjective, ι is not.

Appendix 2.B   Additional Projected Target Parameters

Section 2.B.1 discusses the analogue of section 2.3.6.1 but for the ghNBDM working model. Section 2.B.2 does likewise but for the hNBDM working model.

2.B.1 Projected Target Parameters Under ghNBDM Working Model

Let $\iota_{\text{ghNBDM}} : \mathcal{M}_{ghNBDM} \to \mathcal{M}$, $\iota_{\text{ghNBDM}} : \mathcal{P} \mapsto \mathcal{P}$ denote the inclusion map[21]

$$\mathcal{M}_{\text{ghNBDM}} \xhookrightarrow{\iota_{\text{ghNBDM}}} \mathcal{M} \,.$$

Suppose $\Pi_{\text{ghNBDM}} : \mathcal{M} \to \mathcal{M}_{\text{ghNBDM}}$ is a projection[22]:

$$\Pi_{\text{ghNBDM}} : \mathcal{P} \mapsto \operatorname*{argmin}_{\tilde{\mathcal{P}} \in \mathcal{M}_{\text{ghNBDM}}} \int_{\mathbb{N}_{\geq 0}^S} L(\tilde{\mathcal{P}}(\vec{n})) d\mathcal{P}(\vec{n}) \,, \tag{2.B.1}$$

for a given loss function L, e.g. $L : p \mapsto -\log(p)$, onto the working parametric submodel $\mathcal{M}_{\text{ghNBDM}}$ from the nonparametric model $\mathcal{M}$. One should have the relationships:

- $\Pi_{\text{ghNBDM}} \circ \iota_{\text{ghNBDM}} = \text{Id}_{\text{ghNBDM}} : \mathcal{M}_{\text{ghNBDM}} \to \mathcal{M}_{\text{ghNBDM}}$, and

- $\iota_{\text{ghNBDM}} \circ \Pi_{\text{ghNBDM}} : \mathcal{M} \to \mathcal{M}$ is idempotent.

Given the projection of a distribution $\mathcal{P}$ from the non-parametric model $\mathcal{M}$ onto the working parametric submodel $\mathcal{M}_{\text{ghNBDM}}$, we can now evaluate any of the non-parametrically defined parameters (2.3.9), (2.3.11), (2.3.13), (2.3.15) from section 2.3.4 at the projection, defining a new set of target parameters. This new set of target parameters corresponding to $\mathcal{M}_{\text{ghNBDM}}$ is defined below.

I use the notation $(x)^{\overline{n}}$ ostensibly popularized by Knuth for the *rising factorial*,

$$(x)^{\overline{n}} := \prod_{\nu=0}^{n-1} (x+i) = \frac{\Gamma(x+n)}{\Gamma(x)} \,, \tag{2.B.2}$$

assuming $x > 0$, real, and $n \in \mathbb{N}_{\geq 0}$. Cf. footnote 6 of section 6.2.2.

For a given $\mathcal{P} \in \mathcal{M}$, let $\iota_{\text{ghNBDM}}(\Pi_{\text{ghNBDM}}(\mathcal{P})) =: \text{ghNBDM}(\lambda_{\mathcal{P}}, \vec{f}_{\mathcal{P}}, \zeta_{D,\mathcal{P}}, \zeta_{C,\mathcal{P}}) \in \mathcal{M}_{\text{ghNBDM}}$. The projected target parameters for the gluttonous treatment groups

$$\Psi_{\mathcal{T},g,\text{ghNBDM}}^{(s_1,s_2)} := \Psi_{\mathcal{T},g}^{(s_1,s_2)} \circ \iota_{\text{ghNBDM}} \circ \Pi_{\text{ghNBDM}}$$

are defined as

$$\Psi_{\mathcal{T},g,\text{ghNBDM}}^{(s_1,s_2)} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi_{\mathcal{T},g,\text{ghNBDM}}^{(s_1,s_2)} : \quad \mathcal{P} \quad \mapsto \quad & \mathbb{P}_{\iota_{\text{ghNBDM}}(\Pi_{\text{ghNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1,s_2\}\Big) \\
= \quad & \mathbb{P}_{\text{ghNBDM}(\lambda_{\mathcal{P}}, \vec{f}_{\mathcal{P}}, \zeta_{D,\mathcal{P}}, \zeta_{C,\mathcal{P}})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1,s_2\}\Big) \\
= \quad & \sum_{n=2}^{\infty} \frac{(\zeta_{D,\mathcal{P}}S)^{\zeta_{D,\mathcal{P}}S}}{(\zeta_{D,\mathcal{P}}S+\lambda_{\mathcal{P}})^{\zeta_{D,\mathcal{P}}S+n}} \frac{(\zeta_{D,\mathcal{P}}S)^{\overline{n}}}{(\zeta_{C,\mathcal{P}}S)^{\overline{n}}} \frac{\lambda_{\mathcal{P}}^n}{n!} \Big[(\zeta_{C,\mathcal{P}}S)^{\overline{n}} - (\zeta_{C,\mathcal{P}}S(1-f_{\mathcal{P}}^{(s_1)}))^{\overline{n}} \\
& \qquad\qquad - (\zeta_{C,\mathcal{P}}S(1-f_{\mathcal{P}}^{(s_2)}))^{\overline{n}} + (\zeta_{C,\mathcal{P}}S(1-f_{\mathcal{P}}^{(s_1)}-f_{\mathcal{P}}^{(s_2)}))^{\overline{n}} \Big] \,,
\end{aligned}
\tag{2.B.3}
$$

---

[21]Cf. section 2.A, in particular the last part of section 2.A.2.

[22]Cf. section 2.A, in particular section 2.A.1 on the notation for functions used herein.

the projected target parameters for the gluttonous control groups

$$\Psi^{(s_1,s_2)}_{\mathcal{C},g,\text{ghNBDM}} := \Psi^{(s_1,s_2)}_{\mathcal{C},g} \circ \iota_{\text{ghNBDM}} \circ \Pi_{\text{ghNBDM}}$$

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{C},g,\text{ghNBDM}} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{C},g,\text{ghNBDM}} : \quad \mathcal{P} \quad \mapsto \quad & \mathbb{P}_{\iota_{\text{ghNBDM}}(\Pi_{\text{ghNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}, \, s_1 \notin \text{supp}(\vec{\mathbf{N}}(0))\Big) \\
= \quad & \mathbb{P}_{\text{ghNBDM}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}},\zeta_{D,\mathcal{P}},\zeta_{C,\mathcal{P}})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}, \, s_1 \notin \text{supp}(\vec{\mathbf{N}}(0))\Big) \\
= \quad & \sum_{n=1}^{\infty} \frac{(\zeta_{D,\mathcal{P}}S)^{\zeta_{D,\mathcal{P}}S}}{(\zeta_{D,\mathcal{P}}S+\lambda_{\mathcal{P}})^{\zeta_{D,\mathcal{P}}S+n}} \frac{(\zeta_{D,\mathcal{P}}S)^{\overline{n}}}{(\zeta_{C,\mathcal{P}}S)^{\overline{n}}} \frac{\lambda_{\mathcal{P}}^n}{n!} \cdot \\
& \left[(\zeta_{C,\mathcal{P}}S(1-f^{(s_1)}_{\mathcal{P}}))^{\overline{n}} - (\zeta_{C,\mathcal{P}}S(1-f^{(s_1)}_{\mathcal{P}}-f^{(s_2)}_{\mathcal{P}}))^{\overline{n}}\right],
\end{aligned}
$$

$$(2.\text{B}.4)$$

the projected target parameters for the picky treatment groups

$$\Psi^{(s_1,s_2)}_{\mathcal{T},p,\text{ghNBDM}} := \Psi^{(s_1,s_2)}_{\mathcal{T},p} \circ \iota_{\text{ghNBDM}} \circ \Pi_{\text{ghNBDM}}$$

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{T},p,\text{ghNBDM}} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{T},p,\text{ghNBDM}} : \quad \mathcal{P} \quad \mapsto \quad & \mathbb{P}_{\iota_{\text{ghNBDM}}(\Pi_{\text{ghNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_1,s_2\}\Big) \\
= \quad & \mathbb{P}_{\text{ghNBDM}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}},\zeta_{D,\mathcal{P}},\zeta_{C,\mathcal{P}})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_1,s_2\}\Big) \\
= \quad & \sum_{n=2}^{\infty} \frac{(\zeta_{D,\mathcal{P}}S)^{\zeta_{D,\mathcal{P}}S}}{(\zeta_{D,\mathcal{P}}S+\lambda_{\mathcal{P}})^{\zeta_{D,\mathcal{P}}S+n}} \frac{(\zeta_{D,\mathcal{P}}S)^{\overline{n}}}{(\zeta_{C,\mathcal{P}}S)^{\overline{n}}} \frac{\lambda_{\mathcal{P}}^n}{n!} \left[(\zeta_{C,\mathcal{P}}S(f^{(s_1)}_{\mathcal{P}}+f^{(s_2)}_{\mathcal{P}}))^{\overline{n}} \right. \\
& \left. -(\zeta_{C,\mathcal{P}}Sf^{(s_1)}_{\mathcal{P}})^{\overline{n}} - (\zeta_{C,\mathcal{P}}Sf^{(s_2)}_{\mathcal{P}})^{\overline{n}}\right],
\end{aligned}
$$

$$(2.\text{B}.5)$$

and the projected target parameters for the picky control groups

$$\Psi^{(s_1,s_2)}_{\mathcal{C},p,\text{ghNBDM}} := \Psi^{(s_1,s_2)}_{\mathcal{C},p} \circ \iota_{\text{ghNBDM}} \circ \Pi_{\text{ghNBDM}}$$

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{C},p,\text{ghNBDM}} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{C},p,\text{ghNBDM}} : \quad \mathcal{P} \quad \mapsto \quad & \mathbb{P}_{\iota_{\text{ghNBDM}}(\Pi_{\text{ghNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\Big) \\
= \quad & \mathbb{P}_{\text{ghNBDM}(\lambda_{\mathcal{P}},\vec{f}_{\mathcal{P}},\zeta_{D,\mathcal{P}},\zeta_{C,\mathcal{P}})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\Big) \\
= \quad & \sum_{n=1}^{\infty} \frac{(\zeta_{D,\mathcal{P}}S)^{\zeta_{D,\mathcal{P}}S}}{(\zeta_{D,\mathcal{P}}S+\lambda_{\mathcal{P}})^{\zeta_{D,\mathcal{P}}S+n}} \frac{(\zeta_{D,\mathcal{P}}S)^{\overline{n}}}{(\zeta_{C,\mathcal{P}}S)^{\overline{n}}} \frac{\lambda_{\mathcal{P}}^n}{n!} (\zeta_{C,\mathcal{P}}Sf^{(s_2)}_{\mathcal{P}})^{\overline{n}}.
\end{aligned}
$$

$$(2.\text{B}.6)$$

To estimate these, we just need to estimate $\lambda_{\mathcal{P}}$, $\vec{f}_{\mathcal{P}}$, $\zeta_{D,\mathcal{P}}$, and $\zeta_{C,\mathcal{P}}$ and then use the corresponding plugin estimators. Estimating $\lambda_{\mathcal{P}}$, $\vec{f}_{\mathcal{P}}$, $\zeta_{D,\mathcal{P}}$, and $\zeta_{C,\mathcal{P}}$ also gives us plugin estimators for the probabilities associated with *all* 2^S subsets of $[S]$. So again, like for the hPoMu working model, "information is shared" between estimates for the probabilities of related events. Again, like for the hPoMu working model, as long as all estimates of the $f_{\mathcal{P}}^{(s)}$ are nonzero, all estimated probabilities will also all be nonzero. Thus these estimates should again be useful particularly for the important large S regime.

2.B.2 Projected Target Parameters Under hNBDM Working Model

Let $\iota_{\text{hNBDM}} : \mathcal{M}_{hNBDM} \to \mathcal{M}$, $\iota_{\text{hNBDM}} : \mathcal{P} \mapsto \mathcal{P}$ denote the inclusion map[23]

$$\mathcal{M}_{\text{hNBDM}} \xrightarrow{\iota_{\text{hNBDM}}} \mathcal{M} \,.$$

Suppose $\Pi_{\text{hNBDM}} : \mathcal{M} \to \mathcal{M}_{\text{hNBDM}}$ is a projection[24]:

$$\Pi_{\text{hNBDM}} : \mathcal{P} \mapsto \operatorname*{argmin}_{\tilde{\mathcal{P}} \in \mathcal{M}_{\text{hNBDM}}} \int_{\mathbb{N}_{\geq 0}^S} L(\tilde{\mathcal{P}}(\vec{n})) d\mathcal{P}(\vec{n}) \,, \tag{2.B.7}$$

for a given loss function L, e.g. $L : p \mapsto -\log(p)$, onto the working parametric submodel $\mathcal{M}_{\text{hNBDM}}$ from the nonparametric model $\mathcal{M}$. One should have the relationships:

- $\Pi_{\text{hNBDM}} \circ \iota_{\text{hNBDM}} = \text{Id}_{\text{hNBDM}} : \mathcal{M}_{\text{hNBDM}} \to \mathcal{M}_{\text{hNBDM}}$, and

- $\iota_{\text{hNBDM}} \circ \Pi_{\text{hNBDM}} : \mathcal{M} \to \mathcal{M}$ is idempotent.

Given the projection of a distribution $\mathcal{P}$ from the non-parametric model $\mathcal{M}$ onto the working parametric submodel $\mathcal{M}_{\text{hNBDM}}$, we can now evaluate any of the non-parametrically defined parameters (2.3.9), (2.3.11), (2.3.13), (2.3.15) from section 2.3.4 at the projection, defining a new set of target parameters. This new set of target parameters corresponding to $\mathcal{M}_{\text{hNBDM}}$ is defined below.

For a given $\mathcal{P} \in \mathcal{M}$, let $\iota_{\text{hNBDM}}(\Pi_{\text{hNBDM}}(\mathcal{P})) =: \text{hNBDM}(\lambda_{\mathcal{P}}, \vec{f}_{\mathcal{P}}, \zeta_{\mathcal{P}}) \in \mathcal{M}_{\text{hNBDM}}$. The projected target parameters for the gluttonous treatment groups

$$\Psi_{\mathcal{T},g,\text{hNBDM}}^{(s_1,s_2)} := \Psi_{\mathcal{T},g}^{(s_1,s_2)} \circ \iota_{\text{hNBDM}} \circ \Pi_{\text{hNBDM}}$$

---

[23]Cf. section 2.A, in particular the last part of section 2.A.2.

[24]Cf. section 2.A, in particular section 2.A.1 on the notation for functions used herein.

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{T},g,\text{hNBDM}} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{T},g,\text{hNBDM}} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\text{hNBDM}}(\Pi_{\text{hNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1,s_2\}\Big) \\
&= \quad \mathbb{P}_{\text{hNBDM}(\lambda_\mathcal{P},\vec{f}_\mathcal{P},\zeta_\mathcal{P})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1,s_2\}\Big) \\
&= \quad \left(1 - \left(\frac{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P}}{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P} + f^{(s_1)}_\mathcal{P} \lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P}}\right) \cdot \left(1 - \left(\frac{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P}}{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P} + f^{(s_2)}_\mathcal{P} \lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P}}\right),
\end{aligned}
$$

$$(2.\text{B}.8)$$

the projected target parameters for the gluttonous control groups

$$\Psi^{(s_1,s_2)}_{\mathcal{C},g,\text{hNBDM}} := \Psi^{(s_1,s_2)}_{\mathcal{C},g} \circ \iota_{\text{hNBDM}} \circ \Pi_{\text{hNBDM}}$$

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{C},g,\text{hNBDM}} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{C},g,\text{hNBDM}} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\text{hNBDM}}(\Pi_{\text{hNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}\,,\, s_1 \notin \text{supp}(\vec{\mathbf{N}}(0))\Big) \\
&= \quad \mathbb{P}_{\text{hNBDM}(\lambda_\mathcal{P},\vec{f}_\mathcal{P},\zeta_\mathcal{P})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}\,,\, s_1 \notin \text{supp}(\vec{\mathbf{N}}(0))\Big) \\
&= \quad \left(\frac{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P}}{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P} + f^{(s_1)}_\mathcal{P} \lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P}} \cdot \left(1 - \left(\frac{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P}}{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P} + f^{(s_2)}_\mathcal{P} \lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P}}\right),
\end{aligned}
$$

$$(2.\text{B}.9)$$

the projected target parameters for the picky treatment groups

$$\Psi^{(s_1,s_2)}_{\mathcal{T},p,\text{hNBDM}} := \Psi^{(s_1,s_2)}_{\mathcal{T},p} \circ \iota_{\text{hNBDM}} \circ \Pi_{\text{hNBDM}}$$

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{T},p,\text{hNBDM}} : \quad \mathcal{M} \quad \to \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{T},p,\text{hNBDM}} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\text{hNBDM}}(\Pi_{\text{hNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_1,s_2\}\Big) \\
&= \quad \mathbb{P}_{\text{hNBDM}(\lambda_\mathcal{P},\vec{f}_\mathcal{P},\zeta_\mathcal{P})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_1,s_2\}\Big) \\
&= \quad \left(1 - \left(\frac{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P}}{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P} + f^{(s_1)}_\mathcal{P} \lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S f^{(s_1)}_\mathcal{P}}\right) \cdot \left(1 - \left(\frac{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P}}{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P} + f^{(s_2)}_\mathcal{P} \lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S f^{(s_2)}_\mathcal{P}}\right) \cdot \\
&\qquad \left(\frac{\zeta_\mathcal{P} S(1 - f^{(s_1)}_\mathcal{P} - f^{(s_2)}_\mathcal{P})}{\zeta_\mathcal{P} S(1 - f^{(s_1)}_\mathcal{P} - f^{(s_2)}_\mathcal{P}) + (1 - f^{(s_1)}_\mathcal{P} - f^{(s_2)}_\mathcal{P})\lambda_\mathcal{P}}\right)^{\zeta_\mathcal{P} S(1 - f^{(s_1)}_\mathcal{P} - f^{(s_2)}_\mathcal{P})},
\end{aligned}
$$

$$(2.\text{B}.10)$$

and the projected target parameters for the picky control groups

$$\Psi^{(s_1,s_2)}_{\mathcal{C},p,\text{hNBDM}} := \Psi^{(s_1,s_2)}_{\mathcal{C},p} \circ \iota_{\text{hNBDM}} \circ \Pi_{\text{hNBDM}}$$

are defined as

$$\Psi^{(s_1,s_2)}_{\mathcal{C},p,\text{hNBDM}} : \quad \mathcal{M} \quad \rightarrow \quad [0,1]$$

$$
\begin{aligned}
\Psi^{(s_1,s_2)}_{\mathcal{C},p,\text{hNBDM}} : \quad \mathcal{P} \quad &\mapsto \quad \mathbb{P}_{\iota_{\text{hNBDM}}(\Pi_{\text{hNBDM}}(\mathcal{P}))}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\Big) \\
&= \quad \mathbb{P}_{\text{hNBDM}(\lambda_{\mathcal{P}}, \vec{f}_{\mathcal{P}}, \zeta_{\mathcal{P}})}\Big(\text{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\Big) \\
&= \quad \left(1 - \left(\frac{\zeta_{\mathcal{P}} S f_{\mathcal{P}}^{(s_2)}}{\zeta_{\mathcal{P}} S f_{\mathcal{P}}^{(s_2)} + f_{\mathcal{P}}^{(s_2)} \lambda_{\mathcal{P}}}\right)^{\zeta_{\mathcal{P}} S f_{\mathcal{P}}^{(s_2)}}\right) \cdot \\
&\qquad \left(\frac{\zeta_{\mathcal{P}} S (1 - f_{\mathcal{P}}^{(s_2)})}{\zeta_{\mathcal{P}} S (1 - f_{\mathcal{P}}^{(s_2)}) + (1 - f_{\mathcal{P}}^{(s_2)}) \lambda_{\mathcal{P}}}\right)^{\zeta_{\mathcal{P}} S (1 - f_{\mathcal{P}}^{(s_2)})} .
\end{aligned}
\tag{2.B.11}
$$

Note that (cf. Lemma 3.3 from section 3.A) as $\zeta \to \infty$:

$$\lim_{\zeta \to \infty} \left(\frac{\zeta S f^{(s)}}{\zeta S f^{(s)} + f^{(s)} \lambda}\right)^{\zeta S f^{(s)}} = e^{-f^{(s)} \lambda}, \tag{2.B.12}$$

and therefore equations (2.B.8), (2.B.9), (2.B.10), (2.B.11) *do* approach equations (2.3.19), (2.3.20), (2.3.21), (2.3.22) respectively in the limit as $\zeta_{\mathcal{P}} \to \infty$, as we would have expected a priori. To estimate (2.B.8), (2.B.9), (2.B.10), (2.B.11), we just need to estimate $\lambda_{\mathcal{P}}$, $\vec{f}_{\mathcal{P}}$, and $\zeta_{\mathcal{P}}$ and then use the corresponding plugin estimators. Again, estimating $\lambda_{\mathcal{P}}$, $\vec{f}_{\mathcal{P}}$, and $\zeta_{\mathcal{P}}$ also gives us plugin estimators for the probabilities associated with *all* 2^S subsets of $[S]$. Thus like for the hPoMu and ghNBDM working models, again "information is shared" between estimates for the probabilities of related events and all estimated probabilities will be nonzero as long as all estimates of the $f_{\mathcal{P}}^{(s)}$ are nonzero. While the $\zeta_D = \zeta_C$ constraint provides less flexibility than the full ghNBDM working model, it still has both more flexibility than the hPoMu working model and plugin estimators with the same low level of complexity as those of the hPoMu working model.

Appendix 2.C Toy Model

Consider the following toy model. Droplets have one cell with 50% probability, two cells with two different strains with 50% probability. In both cases the distribution is assumed Multinomial with parameters corresponding to population relative abundances. Now consider the effect of strain s_1, with relative abundance 0.01%, on strain s_2, with relative abundance 20%. If the total number of droplets is $D = 500,000$, then we expect 50,000 droplets containing

strain s_2 only, whereas we would expect only $(50\% \times 20\% \times 0.01\%) \times 500,000 = 5$ droplets containing both strains. Cf. section 7.7.3 which expresses a similar idea.

Keep in mind that the reality is actually almost always *worse* than this toy model. For example, there will be a large and non-negligible number of empty droplets with zero cells. Moreover, when droplets are non-empty, all of the simplifying assumptions implicitly made above can and often will fail to be true. When any given strain is present, it's possible for there to be more than one cell belonging to that strain in the droplet (the "multiple representatives problem", cf. section 5.C). When the droplet is non-empty, there can be more than two strains present in the droplet. Finally, even when conditioning only on droplets with either one or two strains present, there usually will not be a $1 : 1$ ratio of droplets with one strain versus droplets with two strains.

Even the $1 : 1$ ratio from the toy model isn't ideal, because there are $O(S^2)$ possible two-strain treatment groups, and $O(S)$ possible one-strain control groups, so actually in an ideal world we would like a $\approx 1 : S$ ratio of droplets with one strain versus droplets with two strains. So even for the toy model, the sizes of the treatment groups compared to the sizes of the control groups would become increasingly imbalanced as the number of strains becomes larger. This is problematic to the extent that we want S as large as possible.

Part II

Modelling Initial Formation of Droplets

Inferring ecological models of complex microbial communities entails understanding the interactions between microbes and how they affect each other's growth. The initial counts of cells from each strain within a droplet are the baseline for comparing the growth of microbes under different conditions. Being better able to better characterize this baseline for growth means being better able to characterize effects on growth. Thus modelling and understanding these initial counts will help us to understand microbial interactions.

The broader field for part II is the study of statistical models for multivariate count data. One common approach for defining and categorizing these involves generalized linear models. Cf. [KTS$^+$15] or [ZZZS17]. Hierarchical models, often based in some way on the Poisson distribution (which is important for univariate count data), are also commonly used. Cf. [IYAR17] for a review. Models for multivariate count data are important to many problems, such as topic modelling. See [Zho18] or [KTS$^+$15] for examples.

Part II concerns more particularly statistical models for multivariate count data arising from two sources: high-throughput biological assays and ecological studies.

For multivariate count data arising from high-throughput biological assays, such models have been considered for (single-cell) droplet-based microfluidics experiments, and for RNA-seq experiments. Generally the Poisson distribution is used as the default statistical model for single-cell droplet-based microfluidics experiments [CNd$^+$15]. (Although this is univariate count data, not multivariate. See any of [LE13], [KvVDH10], [MRH20], or [GRHW12] for reviews of such experiments.) Examples of work considering statistical models for characterizing count data from RNA-seq experiments include [YHV13] and [ZZZS17].

The literature for multivariate count data arising from ecological studies is likely more developed. Relevant studies are often described as investigating "joint species abundances" or "joint species distribution". See [TRKS$^+$17], [OTN$^+$17], and [CMR21]. Work such as [BGF01] has considered the possible effects of preferential (dis)association of certain organisms (cf. section 3.6.1.2). Such work is not limited only to macroecology. Studies such as [BML$^+$15], [YS14], [HCSB20], or [MWC19] investigate count data from microbial ecology.

Specific Problem The next four chapters of this thesis discuss the "$t = 0$ experiment". This means focusing on candidate statistical working models for the first phase of the overall experiment described in section 1.4. Chapter 3 defines and discusses the candidate statistical working models. Chapter 4 qualitatively compares the candidate statistical working models. We want to determine whether any behave substantially differently from the default working model, which is the simplest. Motivated by the results of chapter 4, chapter 5 investigates and compares the target estimands for each of the candidate statistical working models. We want to determine whether substantially different behavior between the candidate working models translates into substantially different values of the target estimands in practice. Chapter 6 constructs estimators and obtains inference for all of the candidate statistical working models. Even under model misspecification, these estimators return sensible results.

Chapter 3   Chapter 3 belongs to the general field of describing and categorizing statistical models for multivariate count data. For any particular problem with multivariate count data, some thought is usually needed to find reasonable models. Chapter 3 investigates which statistical (working) models are reasonable choices for describing the initial state of a droplet. Chapter 3 shows that specific assumptions justify a default working model for this problem. New working models are derived by relaxing each of these assumptions.

Chapter 4   The problem considered in chapter 4 falls under the general field of "model selection". Herein we only need to consider the subfield of so-called "goodness of fit" tests, in particular those using the classic likelihood ratio statistic. Concretely, chapter 4 investigates whether the failure of any of the assumptions identified in chapter 3 leads to substantially new behavior, and demonstrates that log likelihood ratio statistics can be used to address this question. Failure of the sampling without replacement assumption turns out to have negligible effects in practice, but failures of the other assumptions could be important.

Chapter 5   Like chapter 4, the problem of chapter 5 also belongs both to the general field of model selection and to the subfield of goodness of fit tests in particular. However, the focus herein is on goodness of fit tests for contingency tables using the χ^2-divergence test statistic. (Cf. section 5.2.2 for terminology.) Chapter 5 investigates how failures of the assumptions identified in chapter 3 affect the targeted estimands from section 2.3.4. The results of this chapter confirm that more severe failures of these assumptions lead to more severe discrepancies with the predictions derived from the default working model. The nature of the effect depends on the chosen grouping of droplets defining the targeted estimands.

Chapter 6   The problem of chapter 6 belongs to the general field of point estimation. Chapter 6 is concerned in particular with the plug-in (also known as "method of moments") and maximum likelihood approaches to point estimation. The chapter investigates how to measure, from the data produced by unincubated droplets, failures of the assumptions identified in chapter 3. This is framed as a point estimation problem. The chapter presents both plugin and maximum likelihood estimators for solving this problem. In doing so, failures of the violations are also shown to be understandable non-parametrically.

Significance   We have to investigate specific methods (cf. part III) to predict how much data is needed to infer microbial interactions. To configure the experiment in accordance with such predictions, we first need to predict how much usable data will be produced. (Cf. again the discussion from section 2.3.2.) Part II helps with the latter kind of prediction. Running the experiment once without incubating the droplets allows us to predict how much usable data will be produced by understanding the initial distribution of cells within droplets. This includes assumptions for the distribution of the total numbers of cells within droplets, and for the distribution (given the total number of cells) of proportions of each strain.

Chapter 3

Motivation for and Definitions of Candidate Statistical Working Models for the Initial Formation of Droplets

Herein I show that specific assumptions justify a default working model for the initial state of droplets. See section 3.3.3. I derive new working models by relaxing each of these assumptions. See sections 3.5, 3.6, 3.7, and 3.8. I discuss concrete ways these assumptions could fail to be true in practice, so the relaxations are not purely hypothetical. See section 3.4.

Section 3.1 provides context for what follows. Section 3.2 proposes a general framework for defining any statistical model of the initial formation of droplets. Section 3.3 proposes a "default" working model of the initial formation of droplets. Section 3.4 explains three ways the assumptions underlying this default working model could fail to be true in practice.

The rest of chapter explains how to model these assumption failures. Section 3.5 proposes a working model accounting for the finiteness of the microbial population. Section 3.6 proposes a working model accounting for heterogeneity in the relative abundances of microbial strains throughout the microbial population. Section 3.7 proposes a working model accounting for heterogeneity in the average number of cells throughout the microbial population. Section 3.8 describes a family of distributions subsuming both of the previous two working models as well as the default working model. Section 3.9 summarizes the chapter.

Chapter 3 Contents

3.1 Background and Significance

Broader field This chapter belongs to the general field of describing and categorizing statistical models for multivariate count data. The references [Zho18] and [KTS+15] discuss some examples of these in the field of topic modelling. The article [IYAR17] is a helpful review of some other relevant statistical models. Approaches related to generalized linear models are often used to define statistical models for multivariate count data, cf. [ZZZS17] and [KTS+15] for examples. See also the discussion earlier from the introduction to Part II.

In general, some thought is required to find models that are reasonable for a given, specific problem involving multivariate count data. The references [BGF01], [BML+15], [OTN+17], and [CMR21] are examples of statistical models tailored to problems from ecological studies. Similarly, the references [YHV13], [CNd+15], and [ZZZS17] are examples of statistical models tailored to the analysis of data from high-throughput biological assays. (See any of [LE13], [KvVDH10], [MRH20], or [GRHW12] for reviews of some examples of such experiments.)

Specific problem Herein I investigate which statistical (working) models can be used to describe a droplet's initial state. A statistical model of the initial formation of droplets

helps us to characterize microbial interactions. One of the major challenges of the MOREI data is the randomness of the initial droplet formation process. Unlike manually plating everything, with MOREI the scientist has no *direct* control over which microbial interactions are observed. This means a loss of control over which questions that can be answered from the data. The premise of MOREI is that the vastly increased throughput makes the tradeoff of decreased control worthwhile. The more control we can retain for the scientist when using MOREI, compared to direct plating, the stronger the case for its superiority over existing methods, and the greater its ability to advance the state of the art. Cf. figure 3.1.
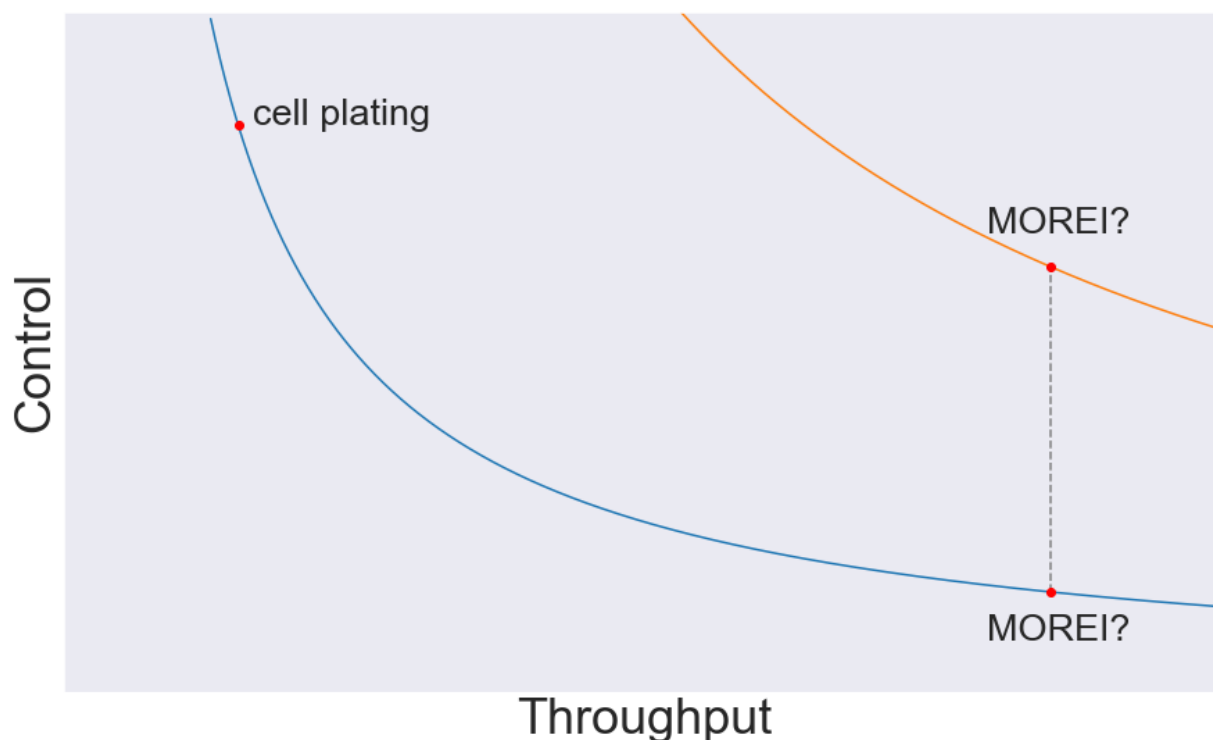
Figure 3.1: Control-Throughput Tradeoff Curve. Having a predictive statistical model of initial droplet formation would allow us to increase control without sacrificing throughput, moving us to the higher tradeoff curve.

To restore the scientist's control over which questions they can answer using MOREI data, we want to account for the effects which the randomness of the initial droplet formation process has on the data. To account for these effects, we need to make accurate predictions about what these effects will be. Cf. again section 2.3.2. To make such predictions, we need to understand the randomness of the initial droplet formation process. To understand its randomness, we need to have adequate statistical models for describing the initial droplet formation process. Similar previous work involving the selection of statistical models for count data has helped to improve the design of RNA-seq experiments, see e.g. [YHV13] or

[ZZZS17]. I hope that this work is an initial step towards analogous progress in the design of MOREI and related future experiments.

The ultimate goal of making accurate predictions about the droplet formation process may sound overly ambitious. Yet the amount of throughput that we can potentially achieve with a MOREI experiment is enormous. With such enormous sample sizes, the effects of asymptotic concentration of measure phenomena (such as the law of large numbers or the central limit theorem) almost certainly come into play. To phrase it very crudely, "randomness + enormous sample sizes = quasi-determinism". To exploit the "quasi-determinism" resulting from enormous sample sizes, all we need is to understand and characterize the underlying probability distribution well enough.

Particular approach   I review how we might model the initial droplet formation process. I begin with a general framework for how we might describe special multivariate count distributions as hierarchical models. Then, within this general framework, I derive specific distributions that follow from given assumptions about the initial droplet formation process.

An approach similar to that used in [BGF01] might be useful to explore for future work. However, such an approach requires us to infer a correlation pattern based on highly sparse observations[1]. This would most likely require substantial work to identify a suitable regularization technique to avoid over-fitting spurious correlations. Moreover, even assuming that issue was addressed, there would be the issue of the scientific interpretation. It would be tempting to infer interactions based on possible preferential grouping or distancing of strains in the sampling pool. Yet it is unclear how much such inferred parameters would actually reveal about the spatial preferences of strains, much less whether such spatial preferences actually correspond to interactions. Cf. section 3.6.1.2.

Other work, such as [OTN+17], [CMR21], [BML+15], or [IYAR17], has also presented possibly relevant working models for multivariate count data. However, like the situation with [BGF01], all of these fairly sophisticated approaches could possibly cause subtle complications when trying to apply them to the data produced by MOREI.

In any case, the working models considered herein are relatively simple and thus suitable as starting points for generalization, e.g. in the directions of [BGF01] or [IYAR17], and so should be considered complementary with those other approaches.

Providing this general framework, and detailed descriptions of simple working models within it, is our first step towards understanding the form of randomness that controls the droplet formation process, and thus ultimately towards allowing the scientist to control which questions they can answer using MOREI data. Cf. again section 2.3.2.

3.2   Hierarchical Count-Categorical Distributions

By **count distribution**, I mean a probability distribution $\mathcal{P}$ with values in the non-negative integers $\mathbb{N}_{\geq 0}$. $N(0)$ corresponds to a count distribution, as does $N^{(s)}(0)$ for all $s \in [S]$.

---

[1]In the sense that observations are vectors such that most of the entries are zero.

Similarly, I define a **multivariate count distribution** to be a probability distribution taking values in $\mathbb{N}_{\geq 0}^S := \bigtimes_{s=1}^S \mathbb{N}_{\geq 0}$. $\vec{\mathbf{N}}(0)$ corresponds to a multivariate count distribution. All marginal distributions of a multivariate count distribution are themselves count distributions.

A typical way to specify a multivariate count distribution is to specify the S count distributions that are its marginal distributions along with the statistical dependence structure amongst the marginals. There are infinitely many ways to define such a statistical dependence structure. Arguably the simplest of such ways is to declare the marginal distributions to be mutually independent, but making such a restriction drastically reduces the variety of multivariate count distributions that can be described. In general, directly specifying the statistical dependence structure amongst the marginal distributions can be quite complicated and difficult. Specifying S count distributions can also be cumbersome.

In this thesis I will use another framework for specifying multivariate count distributions. This framework decomposes the multivariate count distribution into a hierarchical[2] distribution. One of the distributions in the hierarchical distribution is always easy to describe. If the other distribution (more properly/technically speaking a family of distributions) in the hierarchical distribution is also easy to describe, then this hierarchical framework allows us to describe a multivariate count distribution easily using two distributions.

Consider a random vector $\vec{\mathbf{X}} \in \mathbb{N}_{\geq 0}^S$ corresponding to a multivariate count distribution. As mentioned before, this always corresponds to S count distributions, one for each of its S marginal distributions $X^{(s)}$. However, there is also always a way to associate with $\vec{\mathbf{X}}$ a single count distribution, namely the sum of the marginal distributions $X := \sum_{s \in [S]} X^{(s)}$. To define a hierarchical distribution from this, we need to specify a conditional distribution for every possible value of X. In other words, instead of considering $X^{(s)}$ for every $s \in [S]$ to specify $\vec{\mathbf{X}}$, I propose instead considering the pair X and $\vec{\mathbf{X}}|X$.

For any $n \in \mathbb{N}_{\geq 0}$, $\vec{\mathbf{X}}|X = n$ takes values in $\{0, \ldots, n\}^S$ such that the sum of the marginals equals n. I define a **categorical distribution** to be any [3] probability distribution $\mathcal{P}_n$ that takes values in $\{0, \ldots, n\}^S$ with the restriction that the sum of its marginals must equal n, or which equivalently takes as values partitions of n into S (possibly empty) distinct categories.

Two observations are worth making here. First, technically speaking this framework can be used to characterize any multivariate count distribution. No restrictions[4] on $\vec{\mathbf{X}}$ are

---

[2] Some sources call such distributions "compound". I use "hierarchical" here instead to avoid confusion. A "compound Poisson distribution" often refers to the distribution of a sum of N i.i.d. random variables where N itself is Poisson distributed. Using the latter notion, hPoMu is "compound Poisson Multinoulli", but using the former notion ("compound = hierarchical"), hPoMu is "compound Poisson Multinomial".

[3] This definition is much more general than what is more commonly referred to as "*the* categorical distribution" (sensu stricto), for which $n = 1$, and which is also called a Multinoulli distribution. Actually all categorical distributions (sensu lato) with $n = 1$ have to reduce to a Multinoulli distribution. This is analogous to how they must reduce to the Dirac Delta at 0 in the case where $n = 0$. Note also that while for $n > 1$ the sum of n independent Multinoulli random variables does correspond to a categorical distribution (sensu lato), specifically the Multinomial distribution, not every categorical distribution (sensu lato) for $n > 1$ is Multinomial. So one could reasonably say that categorical distributions (sensu lato) strictly generalize categorical distributions (sensu stricto) only for $n > 1$.

[4] Even the restriction that $\vec{\mathbf{X}}$ is a multivariate count distribution may be immaterial, although it does

required to define X and $\vec{\mathbf{X}}|X$. Second is that $\vec{\mathbf{X}}|X$ is not a proper random variable. At best $\vec{\mathbf{X}}|X$ is a family of random variables or perhaps more accurately a "random random variable". For every $n \in \mathbb{N}_{\geq 0}$, $\vec{\mathbf{X}}|X = n$ is a distinct (proper) random variable with its own distinct probability distribution (a categorical distribution). Thus this framework can only provide a simplification in the case that the infinite family of probability distributions for the $\vec{\mathbf{X}}|X = n$ for all $n \in \mathbb{N}_{\geq 0}$ can itself be described simply (ideally as a straightforward function of n). Otherwise this framework clearly complicates[5] the description of $\vec{\mathbf{X}}$.

Given a single count distribution $\mathcal{P}$ and a family of categorical distributions $\mathcal{P}_n$ for all $n \in \mathbb{N}_{\geq 0}$, we can construct a random vector $\vec{\mathbf{X}}$ following what I call a **hierarchical count-categorical distribution** as follows: given a random variable $X \sim \mathcal{P}$, draw from the categorical distribution $\mathcal{P}_n$ whenever the value of X is n. Observe how, by construction, the sum of the marginals of $\vec{\mathbf{X}}$ is X, and that $(\vec{\mathbf{X}}|X = n) \sim \mathcal{P}_n$ for all $n \in \mathbb{N}_{\geq 0}$. Again, any multivariate count distribution can be constructed this way, but whether this is a simple description of the multivariate count distribution depends entirely on whether the infinite family of categorical distributions $\mathcal{P}_n$ for all $n \in \mathbb{N}_{\geq 0}$ can be simply described.

In what follows, I indulge in an abuse of terminology by referring to the infinite family of categorical distributions corresponding to the $\vec{\mathbf{X}}|X = n$ for all $n \in \mathbb{N}_{\geq 0}$ as "the categorical distribution". This is to emphasize the henceforth implicit assumption that this infinite family of categorical distributions admits a simple parameterization in terms of n, simple enough to *not* require much (or any) more effort to describe than a single categorical distribution. When[6] this assumption holds, describing X and $\vec{\mathbf{X}}|X$ is not much more difficult than specifying a single count distribution and a single categorical distribution.

For MOREI, the count distribution part of a hierarchical count-categorical distribution specifies how many cells $N(0)$ are in the droplet. The categorical distribution part specifies which proportions of cells $\vec{\mathbf{N}}(0)|N(0) = n$ belong to each of the S possible strains[7]. Hence using the hierarchical count-categorical distribution framework allows us to split the problem of selecting a statistical model for initial droplet formation into two (smaller and easier) parts: (1) selecting which statistical model describes how many cells are in the droplet via the count distribution and (2) selecting which statistical model describes which strains any cells in the droplet belong to via the categorical distribution. Cf. figure 3.2.

---

imply that X is a discrete distribution. That makes rigorous mathematical proofs easier to achieve.

[5] In the worst case this framework requires specifying one count distribution and infinitely many completely unrelated categorical distributions. Compare this to specifying S count distributions and one description of their statistical dependence structure for the other aforementioned framework.

[6] This assumption may not be possible to satisfy for every multivariate count distribution. It is impossible to falsify either way without making the assumption more precise, of course.

[7] Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.
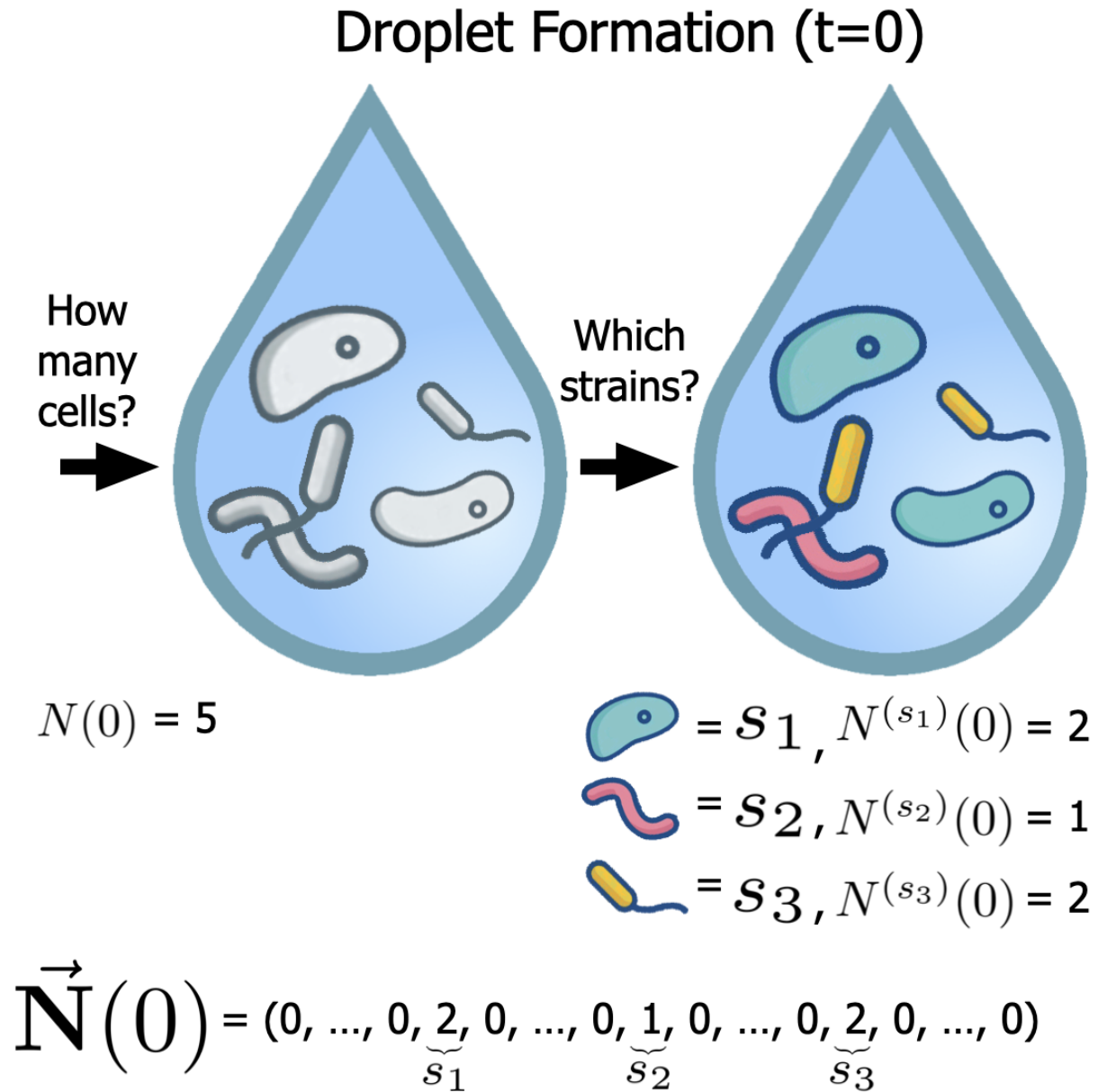
Figure 3.2: Schematic of count-categorical distributions and notation.

Splitting the problem up this way not only simplifies it, but also makes it easier to connect the model with details of the experimental setup. For example, because we expect the dilution of cells in the microbial community sample to only affect the number of cells that end up in each droplet, we can connect it with a parameter for the count distribution (without affecting the categorical distribution). Similarly, because we expect the relative abundances of the strains in the microbial community sample to only affect the proportions of the cells in each droplet belonging to each strain, we can connect them to parameters for

the categorical distribution (without affecting the count distribution).

3.2.1   Naming Convention

Hierarchical count-categorical distributions will be named (or abbreviated) as follows.

- First, a prefix "h" for "hierarchical".

- Two letters corresponding to the count distribution.

- Two letters corresponding to the categorical distribution.

For both the count and categorical distributions, the pairs of letters are assigned as follows.

- If the name of the distribution is one word, then the two letters are the first two letters of the word, the first upper case and the second lower case.

- If the name of the distribution is more than one word, then the two letters are the first two initials of the name, both upper case.

For example the "hierarchical geometric multivariate Wallenius' noncentral hypergeometric distribution" would be "hGeMW".

3.2.2   Connection with Topic Modelling

Imagine a document as a droplet, and the cells within the droplet as words. We can count the number of the words in the document the same way we can count the number of cells in a droplet. Similarly, the way cells can be assigned to strains, words within a document can be assigned to topics. Imagining the strains of the cells as topics of words, we can also count the number of words corresponding to each given topic. Cf. figure 3.3. Statistical questions related to modelling documents and the topics of words within them is already a well-studied field, given the name "topic modelling".
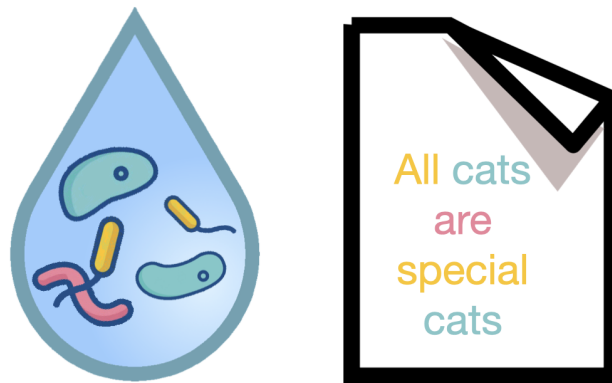
Figure 3.3: Analogy between documents with droplets, words with cells, and topics with strains. Yellow cells ↔ adjectives; blue cells ↔ nouns; pink cells ↔ verbs.

Given the above analogy of documents with droplets, words with cells, and topics with strains, a valuable opportunity for future work is to apply what is already known about topic modelling to modelling the initial formation of droplets in MOREI. See e.g. Theorem 1 of [Zho18] for one relevant example, which discusses both the hPoMu and hNBDM distributions (using different names). Cf. also [KTS⁺15]. The analogy plausibly suggests that mathematical abstractions relevant for one problem can be relevant for the other problem.

Of course this comes with caveats. One obvious difference is that parameter values for these models which are realistic for topic modelling are unlikely to be realistic for modelling the initial formation of droplets in MOREI, and vice versa. For example, the mean number of words in a document should generally be much larger than the mean number of cells in each droplet. A second caveat is that models like those in [Zho18] allow different documents to follow different distributions, whereas we might not want different droplets to follow different distributions (although doing so might be useful for e.g. modelling batch effects).

3.3   hPoMu: A Default Working Model

Section 3.3.1 gives the explicit definition of the hPoMu working model, mostly for future reference. Section 3.3.2 gives two heuristic arguments which establish hPoMu as the "baseline" or "default" (working) model for the initial formation of droplets against which other models are compared. Finally section 3.3.3 outlines some of the assumptions implicit in the derivation of the hPoMu working model, hinting at some of its potential weaknesses for accurately modelling the initial formation of droplets in practice.

3.3.1   hPoMu Definition

The "hierarchical Poisson Multinomial" (abbreviated as hPoMu) distribution is a hierarchical count-categorical distribution whose count distribution is Poisson and whose categorical distribution is Multinomial. Hence the probability mass function (PMF) for droplet d equals

$$
\begin{aligned}
\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) &= \mathbb{P}(N_d(0) = n) \cdot \mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right) \\
&= \frac{e^{-\lambda}\lambda^n}{n!} \cdot \binom{n}{n^{(1)}\cdots n^{(S)}} \prod_{s=1}^{S} (f^{(s)})^{n^{(s)}} .
\end{aligned}
\tag{3.3.1}
$$

An alternative form of expressing the PMF of hPoMu, which reveals that its marginals are mutually independent, is

$$
\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) = \prod_{s=1}^{S} \frac{e^{-f^{(s)}\lambda}(f^{(s)}\lambda)^{n^{(s)}}}{(n^{(s)})!} .
\tag{3.3.2}
$$

Via tedious algebra, one can show that equation (3.3.2) really does equal equation (3.3.1). The $n!$ from the multinomial coefficient cancels with that from the Poisson PMF. The remaining factorials from the multinomial coefficient can be distributed across the product,

and $e^{-\lambda}\lambda^n = \prod_{s=1}^{S} e^{-f^{(s)}\lambda}\lambda^{n^{(s)}}$, which can also be distributed across the product. So despite looking different, the first and second definitions are actually consistent with one another.

Heuristically, equation (3.3.2) says that, during the formation of a droplet, the sampling rate λ is spread out among the strains according to their frequencies in the population.

3.3.2 Derivations

The first derivation determines the count distribution and then the marginals. The second derivation determines the marginals and then the count distribution.

Assume we are sampling from a population of $K \in \mathbb{N}$ cells. Therefore we are sampling from a population of $f^{(1)}K \in \mathbb{N}$ cells of strain 1, ..., $f^{(s)}K \in \mathbb{N}$ cells of strain s, ..., and $f^{(S)}K \in \mathbb{N}$ cells of strain S.

When forming a droplet d, we assume each cell has the same small probability p of ending up in the droplet. Thus the indicator random variable of ending up in the droplet is a Bernoulli(p) random variable for every cell.

We also assume that whether a given cell ends up in the droplet d is completely independent of what happens to any of the other cells.

3.3.2.1 First Derivation

The number of cells that end up in the droplet d is the sum of all K of the indicator random variables. Because the sum of i.i.d. Bernoullis is binomial, the distribution of the number of cells that end up in the droplet is Binomial(p, K).

Due to the way the experiment was calibrated, we know a priori that the expected number of cells in the droplet d is (some fixed constant) λ. Because the number of cells follows a binomial distribution, the expected number of cells in the droplet d is also pK. (Thus $p = \frac{\lambda}{K}$.)

Because K is very large, we may as well use the approximating distribution as $K \to \infty$. By the "Law of Small Numbers", a.k.a. the Poisson Limit Theorem, this distribution is Poisson. In other words:

$$\mathbb{P}(N_d(0) = n) = \frac{e^{-\lambda}\lambda^n}{n!}. \tag{3.3.3}$$

Now that we know the number of cells in droplet d, we need to determine which strain each cell belongs to. Because any cell was equally likely to have been sampled, the strains follow a multivariate hypergeometric distribution with categories of sizes $f^{(1)}K, \ldots, f^{(S)}K$. Because the K is very large, it is very unlikely[8] that any given cell would be sampled twice. Thus we might as well use the approximating distribution as $K \to \infty$, multinomial with category probabilities $f^{(1)}, \ldots, f^{(S)}$, which corresponds to sampling with replacement. So

$$\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right) = \binom{n}{n^{(1)}\cdots n^{(S)}} \prod_{s=1}^{S} (f^{(s)})^{n^{(s)}}. \tag{3.3.4}$$

---

[8]Even more unlikely than being sampled once, which is already very unlikely.

Combining the count distribution and categorical distribution derived above, it follows that the joint distribution must approximately equal

$$
\begin{aligned}
\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) &= \mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right) \cdot \mathbb{P}(N_d(0) = n) \\
&= \binom{n}{n^{(1)} \cdots n^{(S)}} \prod_{s=1}^{S} (f^{(s)})^{n^{(s)}} \cdot \frac{e^{-\lambda} \lambda^n}{n!} .
\end{aligned} \tag{3.3.5}
$$

3.3.2.2 Second Derivation

For any given strain s, the number of cells *of that strain* that end up in the droplet d is the sum of all $f^{(s)}K$ of the indicator random variables belonging to cells *of that strain*. Because the sum of i.i.d. Bernoullis is binomial, the marginal distribution of the number of cells of strain s that end up in the droplet d is again binomially distributed, $\text{Binomial}(p, f^{(s)}K)$.

Because[9] $p = \frac{\lambda}{K}$, the expected number of cells of strain s that end up in the droplet d is $p(f^{(s)}K) = \frac{\lambda}{K}(f^{(s)}K) = f^{(s)}\lambda$. If we assume that K is sufficiently large such that $f^{(s)}K$ is also very large, then as before we may as well use the approximating distribution as[10] $(f^{(s)}K) \to \infty$. As in the first derivation, this distribution is Poisson as a result of the Poisson limit theorem, so for the marginal distributions:

$$
\mathbb{P}\left(N_d^{(s)}(0) = n^{(s)}\right) = \frac{e^{-f^{(s)}\lambda}(f^{(s)}\lambda)^{n^{(s)}}}{(n^{(s)})!} . \tag{3.3.6}
$$

Because what happens to any given cell is completely independent of what happens to any other cells, the marginal distributions for each strain are sums of mutually disjoint sets of i.i.d. Bernoulli random variables. Therefore the marginal distributions must be mutually independent Binomial random variables, and thus in the limit approach mutually independent Poisson random variables. Thus the joint distribution must approximately equal

$$
\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) = \prod_{s=1}^{S} \frac{e^{-f^{(s)}\lambda}(f^{(s)}\lambda)^{n^{(s)}}}{(n^{(s)})!} . \tag{3.3.7}
$$

3.3.3 Implicit Model Assumptions

Both of the above two derivations share the following implicit assumptions:

- The experiment can be calibrated to a priori guarantee a certain *average* number of cells per droplet.

- The total number of cells in the sampling pool, K, is very large.

---

[9]Due to the argument found in the first derivation in section 3.3.2.1.

[10]Equivalently, as $K \to \infty$ along a subsequence such that $f^{(s)}K \in \mathbb{N}$ (i.e. remains an integer).

- For *every* strain s, the number of cells $f^{(s)}K$ in the sampling pool which belong to strain s is very large.

- Each cell has the same small probability (regardless of its strain) of ending up in the droplet as any other cell.

- (Regardless of their strains), whether a given cell ends up in the droplet is completely independent of what happens to any other cell.

The second and third can be thought of as stating that the population size is effectively infinite. The fourth and fifth can be thought of as stating that the sampling pool is completely homogeneous, or perfectly "well-mixed".

My collaborator tells me that, using commercially available microfluidic devices, one can calibrate the experiment to a priori guarantee a certain average number of cells per droplet, and that the total number of cells in the sampling pool will be $\sim 10^{10}$. However, the remaining three assumptions are more questionable. In the next section I explain three ways these assumptions could fail to be true in practice.

3.3.4   Similar Work in Ecology

Similar (frameworks for) hierarchical working models for multivariate count data have been described for joint species distributions in (macro)ecology and observational metagenomic studies in microbial ecology [BGF01][BML$^+$15] [IYAR17] [OTN$^+$17] [CMR21]. However hPoMu appears to be far simpler than the other working models proposed in those studies. Moreover, hPoMu is also a much more direct generalization of the univariate Poisson distribution that is used as the default working model for single-cell (single-type) droplet microfluidics [CNd$^+$15]. Hence in my opinion it makes much more sense to propose hPoMu as the "default" working model for multivariate count data than any of these alternatives.

Nevertheless, the independent convergence of this and similar work onto the approach of hierarchical models (cf. also section 3.2.2) seems to support the claim from section 3.2 that hierarchical models are a good strategy for characterizing multivariate count data.

3.4   Failures of hPoMu Model Assumptions

While hPoMu is definitely a sensible starting point, in practice its implicit assumptions failing to be realistic could undermine its usefulness. Sections 3.4.1, 3.4.2, and 3.4.3 define three plausible ways that these assumptions could fail to be true in practice. Section 3.4.4 outlines the framework used to describe the latter failures of the hPoMu assumptions.

3.4.1   Sampling is from a Finite Population without Replacement

In both of the above derivations from section 3.3.2, we assumed that the number of cells being sampled from is "effectively infinite". This implied that no meaningful loss of accuracy

was occurred when invoking the Poisson limit theorem for the number of cells in each droplet. As mentioned before, with $\sim 10^{10}$ cells, this is a reasonable assumption.

However, we assumed more than this. We also assumed that *for each individual strain s* the number of cells is "effectively infinite". This assumption is reasonable and uncontroversial in the case of relatively abundant strains. However, in the case of "rare" strains, whose relative abundance is e.g. 0.1%, 0.01%, or even lower, whether this assumption remains reasonable is a priori unclear. Admittedly 0.01% of $\sim 10^{10}$ cells is still many, but perhaps not enough for the asymptotically approximate distributions to remain accurate.

3.4.2   Density Heterogeneity

In both of the above derivations we assumed that each cell everywhere in the sampling pool had the same probability of being sampled for every droplet. We also assumed that the probability of ending up in the droplet for any given cell was statistically independent of what happened to any other cell. This would make sense in a theoretical world where each cell corresponded to a scalar field uniformly spread across the entire sampling pool. In reality cells have finite volume and are limited at any given moment in time to a single location in the sampling pool. Only cells reasonably close to where the droplet is being formed have any nonzero probability of ending up in the droplet. Cells that are closer to one another are more likely to end up in the same droplet than those that are further apart.

These details hypothetically might not matter in the case where the average number of cells per unit volume[11] was constant throughout the entire sampling pool. However in practice we have no guarantee that this would be the case, and conceivably different sections of the sampling pool could have average numbers of cells per unit volume which differ from the average number of cells per unit volume for the entire sampling pool (i.e. the total number of cells divided by the volume of the sampling pool).

Herein I call this phenomenon "density heterogeneity". Compare figures 3.4a and 3.4b. In those hypothetical examples, the average number of cells per unit volume for the entire sampling pool is 2, but the average number of cells per unit volume differs between the specific sections. In figure 3.4a the differences are small, introducing not much more variance compared to what would be expected under the uniform/homogeneous density assumed by the default hPoMu model. In figure 3.4b the differences are larger, introducing much more variance compared to what would be expected under uniform/homogeneous density assumed by the default hPoMu model.

In the hierarchical count-categorical distribution framework, density heterogeneity is an issue which affects the count distribution. The default Poisson model for the count distribution will not account for additional variance which might be introduced in practice by density heterogeneity. See [YHV13] for modelling of density heterogeneity for RNA-seq.

---

[11]For volumes "large enough". This is obviously false for "infinitesimal volumes" as discussed above.
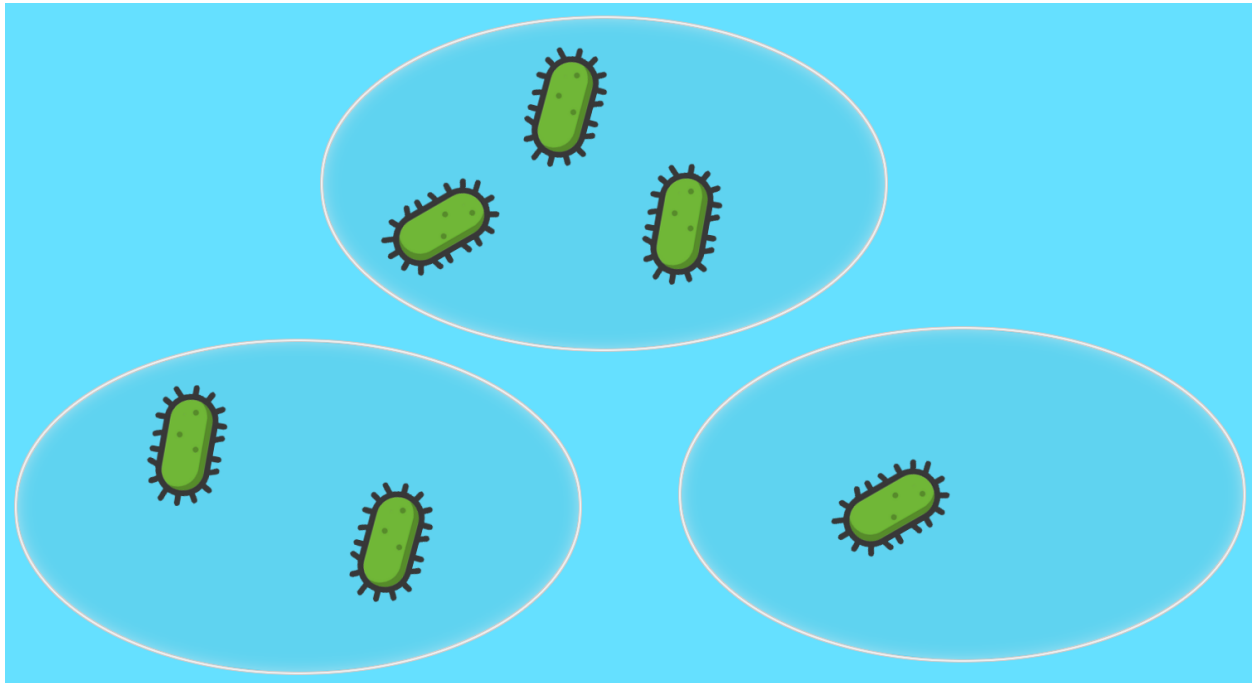
(a) Low Density Heterogeneity
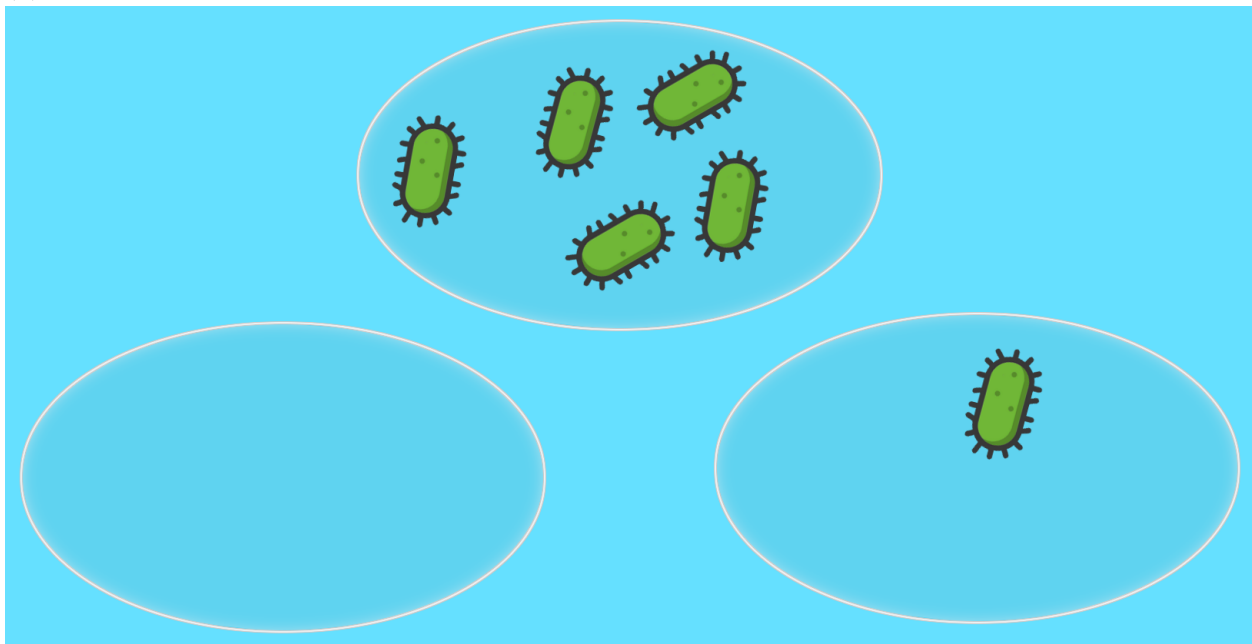
(b) High Density Heterogeneity

Figure 3.4: Schematic representation of varying levels of density heterogeneity.

3.4.3 Compositional Heterogeneity

In making the homogeneous assumptions that all cells have equal probability of ending up in the droplet, and that what happens to a given cell is completely independent of what happens to any other cells, we implicitly ignore the fact that cells belonging to different strains may tend to behave very differently.

In particular, although the relative abundance of strain s across the entire sampling pool may be $f^{(s)}$, conceivably cells of this strain could be slightly more concentrated or slightly less concentrated than this in different sections of the sampling pool. This means different sections of the sampling pool could have average relative abundances which differ from the average relative abundances for the entire sampling pool. Thus the fact that each droplet is formed only from cells nearby a certain location in the sampling pool could affect not only the number of cells likely to end up in the droplet, but also which strains are most likely to end up in the droplet. This is probably guaranteed to happen to some extent due to random chance given that cells occupy finite, not infinitesimal volume. The probabilities of belonging to given strains cannot be perfectly uniform scalar fields. For an example of one mechanism that could aggravate this, cells of a certain strain could preferentially aggregate with, or avoid, cells of another strain due to their interactions with one another.

Herein I call this phenomenon "compositional heterogeneity". Compare figures 3.5a and 3.5b. In those hypothetical examples, the average relative abundances of each strain for the entire sampling pool is $\frac{1}{2}$ blue, $\frac{1}{4}$ pink, $\frac{1}{4}$ yellow, but the average relative abundances of each strain differs between the specific sections. In figure 3.5a the differences are small, introducing not much more variance compared to what would be expected under the uniform/homogeneous density assumed by the default hPoMu model. In figure 3.5b the differences are larger, introducing much more variance compared to what would be expected under uniform/homogeneous density assumed by the default hPoMu model.

In the hierarchical count-categorical distribution framework, compositional heterogeneity is an issue which affects the categorical distribution. The default multinomial model for the categorical distribution (found in hPoMu) is a "null model" that does not account for additional variance that might be introduced in practice by compositional heterogeneity.

See 3.6.1.2 for a clarification of the intended notion of "compositional heterogeneity".

3.4.4 Modelling Heterogeneities

For both density heterogeneity and compositional heterogeneity, extra variance is introduced that is not accounted for by the default hPoMu model. In trying to predict and model the probability distribution of droplet formation, instead of trying to develop complicated mechanistic models describing these spatial heterogeneities whose impacts on the initial droplet formation distribution we might then infer, we can instead seek the humbler goal of modelling this extra variance. In other words, instead of directly modelling the causes of density and compositional heterogeneity, we can restrict ourselves to modelling their
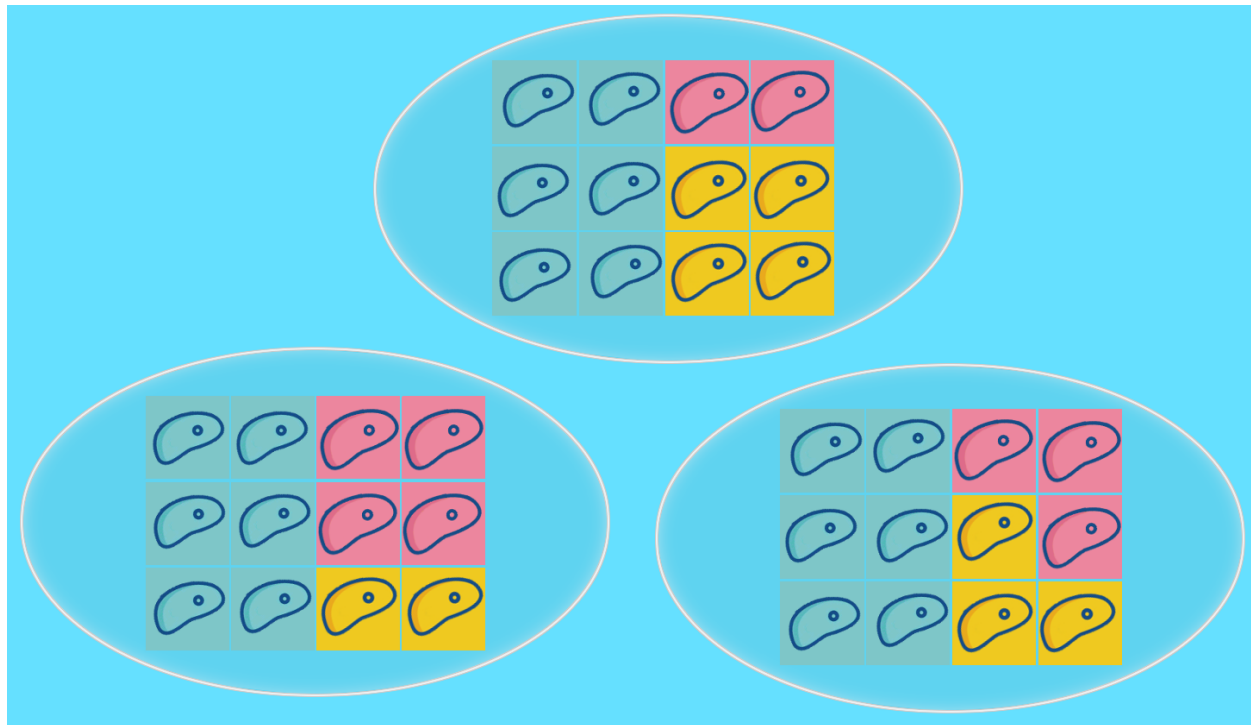
(a) Low Compositional Heterogeneity
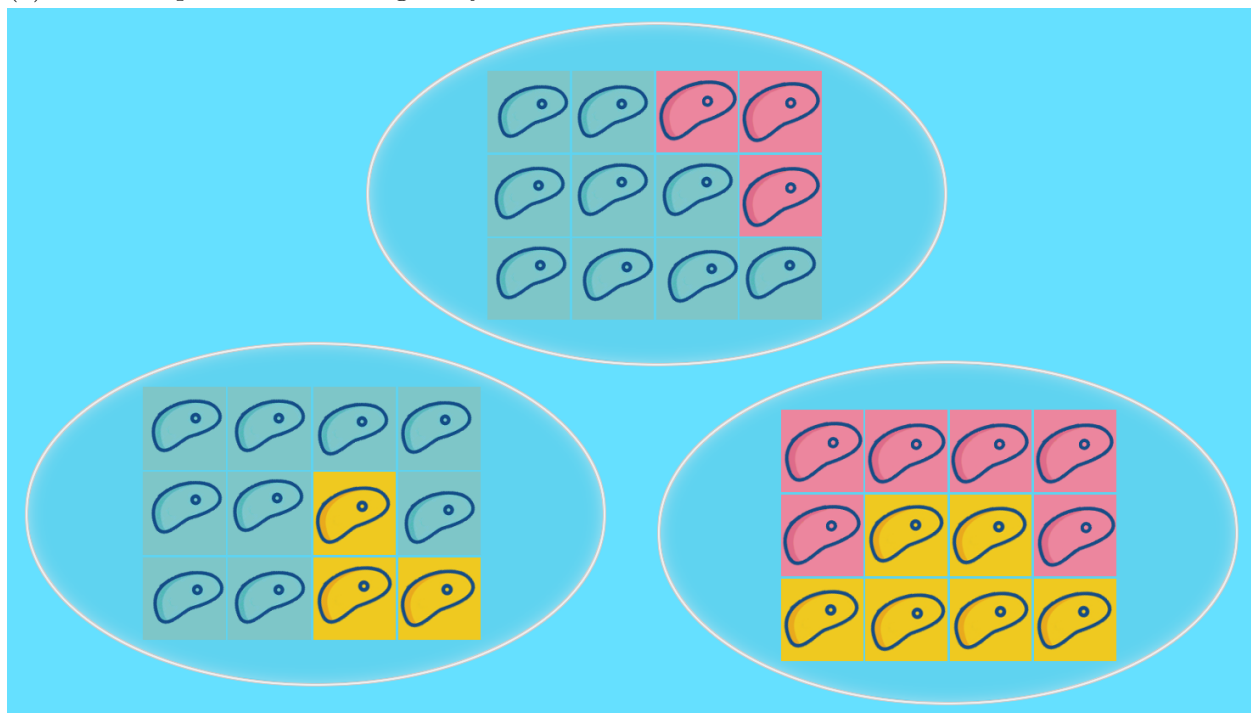
(b) High Compositional Heterogeneity

Figure 3.5: Schematic representation of varying levels of compositional heterogeneity.

"downstream" effect on variance. This should be enough by itself to make our models for the distribution of initial droplet formation more accurate.

I quantify this extra variance via the over-dispersion (defined below in section 3.4.4.1), and describe later in sections 3.6.1.3 and 3.7.1.2 how the parameters of the proposed models relate to their over-dispersion, and therefore can describe the increased variance caused by either density heterogeneity or compositional heterogeneity.

3.4.4.1 Over-Dispersion

Herein we define the over-dispersion of Y with respect to X, where Y and X are scalar random variables, as

$$\frac{\text{Var}[Y] - \text{Var}[X]}{\text{Var}[X]} . \tag{3.4.1}$$

If Y and X are random vectors, and if there exists a constant C such that

$$\frac{\text{Cov}[Y_i, Y_j] - \text{Cov}[X_i, X_j]}{\text{Cov}[X_i, X_j]} = C \tag{3.4.2}$$

for all i, j (including $i = j$), then C is herein defined to be the over-dispersion of Y with respect to X.

The over-dispersion of $\mathcal{P}_2$ with respect to $\mathcal{P}_1$, where $\mathcal{P}_2$ and $\mathcal{P}_1$ are probability distributions, is defined to be the over-dispersion of Y with respect to X, where the random variables $Y \sim \mathcal{P}_2$ and $X \sim \mathcal{P}_1$ but are otherwise arbitrary.

3.5 Working Model for Sampling Without Replacement

The hPoMu working model assumes that the number of cells of any given strain s is effectively infinite, and that therefore sampling can effectively be modeled as without replacement. However, for very rare strains the adequacy of this asymptotic approximation becomes questionable. Therefore what we need is a distribution whose definition explicitly acknowledges the finiteness of the populations of cells being sampled. In section 3.5.1 I introduce a working model which fills this gap, and then in section 3.5.2 I show that this working model extends what can be described using hPoMu alone.

3.5.1 Introduction to hTPMH

As suggested by its name, the "hierarchical truncated Poisson multivariate hypergeometric" (hTPMH) distribution uses a multivariate hypergeometric model of sampling without replacement for the categorical distribution. The multivariate hypergeometric distribution is introduced in section 3.5.1.1. The truncated Poisson distribution is introduced in section

3.5.1.2, and is a slight modification of the Poisson distribution that explicitly acknowledges that the number of cells to be sampled from is finite. Finally I combine the new information learned to define the hTPMH model in section 3.5.1.3.

The total number of cells in the sampling pool before any droplets are formed is $K =: K_0$ (cf. the notation from section 3.3.2), and the total number that belong to strain s before any droplets have formed is $K^{(s)} = f^{(s)} K =: K_0^{(s)}$. This leads us to define $\vec{\mathbf{K}} =: \vec{\mathbf{K}}_0$ as

$$\vec{\mathbf{K}} := (K^{(1)}, \dots, K^{(s)}, \dots, K^{(S)}), \quad \sum_{s=1}^{S} K^{(s)} = K. \tag{3.5.1}$$

Similarly, K_d is the total number of cells in the sampling pool **after** d droplets have been formed, whereas $K_d^{(s)}$ is the total number of cells of strain s in the sampling pool **after** d droplets have been formed. Thus **after** d droplets have been formed we also have a corresponding vector $\vec{\mathbf{K}}_d$ such that

$$\vec{\mathbf{K}}_d := (K_d^{(1)}, \dots, K_d^{(s)}, \dots, K_d^{(S)}), \quad \sum_{s=1}^{S} K_d^{(s)} = K_d. \tag{3.5.2}$$

In particular, for all $d \in [D]$ we have (by definition) that $K_d = K_0 - \sum_{\delta=1}^{d} N_\delta(0)$ and $K_d^{(s)} = K_0^{(s)} - \sum_{\delta=1}^{d} N_\delta^{(s)}(0)$.

3.5.1.1 Multivariate Hypergeometric (MH) Distribution

The PMF of the multivariate hypergeometric distribution can be written as

$$\frac{(K_{d-1} - n)!}{(K_{d-1})!} \cdot \left[\prod_{s=1}^{S} \frac{(K_{d-1}^{(s)})!}{(K_{d-1}^{(s)} - n^{(s)})!} \right] \cdot \binom{n}{n^{(1)} \cdots n^{(S)}}. \tag{3.5.3}$$

As a reminder from section 2.2.1, $n := \sum_{s=1}^{S} n^{(s)}$ (the sum of the entries of $\vec{\mathbf{n}}$).

3.5.1.2 Truncated Poisson (TP) Distribution

The mass function of the truncated Poisson distribution is

$$\frac{1}{\mathbb{P}(\mathrm{Pois}(\lambda) \leq K_{d-1})} \cdot \frac{e^{-\lambda} \lambda^n}{n!}. \tag{3.5.4}$$

Unless K_{d-1} is very small, this is effectively almost identical to the corresponding probabilities from the Poisson distribution.

3.5.1.3 hTPMH Definition

Under the hTPMH distribution, for the d'th droplet: the probability of the strain distribution vector given the number of cells

$$\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n, \vec{\mathbf{K}}_{d-1}\right)$$

is determined by a multivariate hypergeometric distribution, while the probability of the number of cells $\mathbb{P}\left(N_d(0) = n \mid \vec{\mathbf{K}}_{d-1}\right)$ is determined by a truncated Poisson distribution.

Thus for the unconditional probability:

$$
\begin{aligned}
&\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid \vec{\mathbf{K}}_{d-1}\right) \\
=\ & \mathbb{P}\left(N_d(0) = n \mid \vec{\mathbf{K}}_{d-1}\right) \cdot \mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n, \vec{\mathbf{K}}_{d-1}\right) \\
=\ & \mathbb{P}(\text{Pois}(\lambda) \le K_{d-1})^{-1} \cdot \frac{e^{-\lambda}\lambda^n}{n!} \frac{(K_{d-1} - n)!}{(K_{d-1})!} \cdot \left[\prod_{s=1}^{S} \frac{(K_{d-1}^{(s)})!}{(K_{d-1}^{(s)} - n^{(s)})!} \right] \cdot \binom{n}{n^{(1)} \cdots n^{(S)}}.
\end{aligned}
$$

(3.5.5)

3.5.2 hTPMH Generalizes hPoMu

One can show that hTPMH converges to hPoMu under appropriate limits. Hence the hPoMu working model can be considered a submodel of the hTPMH working model.

Sections 3.5.2.1, 3.5.2.2, and 3.5.2.3 give the needed context and then show that hTPMH really is an extension of hPoMu. Because hTPMH and hPoMu have nearly the same count distribution, at its core this reduces to showing how the multivariate hypergeometric distribution extends the multinomial distribution, which section 3.5.2.1 discusses.

3.5.2.1 Approximate Form of Likelihood Ratio of Multivariate Hypergeometric with respect to Multinomial and Bounds

Applying Lemma 3.1 and then algebraic manipulations (e.g. multiplying numerator and denominator by the same factor) gives

$$
\begin{aligned}
& \frac{(K_{d-1} - n)!}{(K_{d-1})!} \cdot \left[\prod_{s=1}^{S} \frac{(K_{d-1}^{(s)})!}{(K_{d-1}^{(s)} - n^{(s)})!} \right] \\
&= \exp\left(\mathscr{R}_{\text{MHg}}(\vec{\mathbf{n}}, \vec{\mathbf{K}}_{d-1}) \right) \cdot \left(\frac{K_{d-1} - n}{K_{d-1}} \right)^{K_{d-1} - n + \frac{1}{2}} \\
& \cdot \prod_{s=1}^{S} \left(\frac{K_{d-1}^{(s)}}{K_{d-1}^{(s)} - n^{(s)}} \right)^{K_{d-1}^{(s)} - n^{(s)} + \frac{1}{2}} \cdot \prod_{s=1}^{S} \left(\frac{K_{d-1}^{(s)}}{f^{(s)} K_{d-1}} \right)^{n^{(s)}} \cdot \prod_{s=1}^{S} (f^{(s)})^{n^{(s)}}.
\end{aligned}
$$

(3.5.6)

In equation 3.5.6, based on [Rob55], Lemma 1 implies that

$$\underline{\mathscr{R}}\left(\vec{\mathbf{K}}_{d-1} - \mathbf{n}, \vec{\mathbf{K}}_{d-1}\right) < \mathscr{R}_{\mathrm{MHg}}\left(\mathbf{n}, \vec{\mathbf{K}}_{d-1}\right) < \overline{\mathscr{R}}\left(\vec{\mathbf{K}}_{d-1} - \mathbf{n}, \vec{\mathbf{K}}_{d-1}\right), \tag{3.5.7}$$

with the "lower Robbins function" $\underline{\mathscr{R}}$ defined for convenience as

$$\underline{\mathscr{R}}(\vec{N}, \vec{M}) := \frac{1}{12N+1} - \sum_{s=1}^{S}\frac{1}{12N_s} - \frac{1}{12M} + \sum_{s=1}^{S}\frac{1}{12M_s+1} = -\overline{\mathscr{R}}(\vec{M}, \vec{N}), \tag{3.5.8}$$

and the "upper Robbins function" $\overline{\mathscr{R}}$ defined for convenience as

$$\overline{\mathscr{R}}(\vec{N}, \vec{M}) := \frac{1}{12N} - \sum_{s=1}^{S}\frac{1}{12N_s+1} - \frac{1}{12M+1} + \sum_{s=1}^{S}\frac{1}{12M_s} = -\underline{\mathscr{R}}(\vec{M}, \vec{N}). \tag{3.5.9}$$

Above we again used the notational the convention that, given a vector $\vec{X} = (X_1, \ldots, X_S)$, $X := \sum_{s=1}^{S} X_s$.

Equation 3.5.6 implies that the likelihood ratio of Multivariate Hypergeometric distribution with respect to the Multinomial distribution may be written

$$\exp\left(\mathscr{R}_{\mathrm{MHg}}(\mathbf{n}, \vec{\mathbf{K}}_{d-1})\right) \cdot \left(\frac{K_{d-1}-n}{K_{d-1}}\right)^{K_{d-1}-n+\frac{1}{2}} \cdot \prod_{s=1}^{S}\left(\frac{K_{d-1}^{(s)}}{K_{d-1}^{(s)}-n^{(s)}}\right)^{K_{d-1}^{(s)}-n^{(s)}+\frac{1}{2}} \cdot \prod_{s=1}^{S}\left(\frac{K_{d-1}^{(s)}}{f^{(s)}K_{d-1}}\right)^{n^{(s)}}. \tag{3.5.10}$$

In other words, multiplying the PMF of the Multinomial distribution by the quantity in 3.5.10 gives the PMF of the Multivariate Hypergeometric Distribution.

3.5.2.2 Approximate Form of Likelihood Ratio of hTPMH with respect to hPoMu and Bounds

From equation 3.5.10 it follows readily that the likelihood ratio of the hTPMH distribution with respect to the hPoMu distribution is

$$\mathbb{P}(\mathrm{Pois}(\lambda) \leq K_{d-1})^{-1} \exp\left(\mathscr{R}_{\mathrm{MHg}}(\mathbf{n}, \vec{\mathbf{K}}_{d-1})\right) \cdot \left(\frac{K_{d-1}-n}{K_{d-1}}\right)^{K_{d-1}-n+\frac{1}{2}}$$
$$\cdot \prod_{s=1}^{S}\left(\frac{K_{d-1}^{(s)}}{K_{d-1}^{(s)}-n^{(s)}}\right)^{K_{d-1}^{(s)}-n^{(s)}+\frac{1}{2}} \cdot \prod_{s=1}^{S}\left(\frac{K_{d-1}^{(s)}}{f^{(s)}K_{d-1}}\right)^{n^{(s)}}. \tag{3.5.11}$$

The factor $\mathbb{P}(\mathrm{Pois}(\lambda) \leq K_{d-1})^{-1}$ reflects how the number of cells in droplet d will be truncated Poisson distributed, not Poisson distributed. Otherwise this is the same as (3.5.10).

3.5.2.3 Sketch of Proof that hTPMH Converges in Distribution to hPoMu

Because the hTPMH and hierarchical Poisson Multinomial (hPoMu) distributions are discrete, to show that the former converges in distribution to the latter, it suffices to show that the probability mass function of the former converges to the latter.

Equation 3.5.11 gives a quantity such that, when it is multiplied with the PMF of the hPoMu distribution, the result is the PMF of the corresponding hTPMH distribution.

Fixing the values of $\vec{\mathbf{N}}_\delta(0)$ for all $\delta \in [d-1]$, from the above it follows that the quantity in 3.5.11 converging to 1 as $K \to \infty$ implies that the hTPMH distribution converges to the hPoMu distribution as $K \to \infty$.

The desired convergence follows from the Squeeze Theorem. Note that both

$$\underline{\mathscr{R}}\left(\vec{\mathbf{K}}_{d-1} - \vec{\mathbf{n}}, \vec{\mathbf{K}}_{d-1}\right) \quad \text{and} \quad \overline{\mathscr{R}}\left(\vec{\mathbf{K}}_{d-1} - \vec{\mathbf{n}}, \vec{\mathbf{K}}_{d-1}\right)$$

converge to 0 as $K \to \infty$ because of additivity of limits and that $\pm \lim\limits_{N \to \infty} \frac{1}{a_N} = 0$ for any sequence a_N such that $\lim\limits_{N \to \infty} a_N = \infty$ (due to the Archimedean property of $\mathbb{R}$).

The factor $\mathbb{P}(\text{Pois}(\lambda) \leq K_{d-1})^{-1}$ converges to 1 because of the defintion of K_{d-1} and that any valid cumulative distribution function must converge to 1 as its input goes to ∞ (or otherwise there would be nonzero probability mass "at infinity").

The last expression in (3.5.11) converges to 1 because of L'Hôspital's Rule, and the remaining factors converge to 1 because of Lemma 3.3.

We may assume without loss of generality that, when taking the limit as $K \to \infty$, we do so along a subsequence such that all $K^{(s)} = f^{(s)}K$ are in $\mathbb{N}$. This of course also guarantees that all of the $K_{d-1}^{(s)} := K^{(s)} - \sum_{\delta=1}^{d-1} N_\delta^{(s)}(0)$ are also integer-valued.

Instead of showing that the likelihood ratio converges to 1, one can also show (e.g. according to one's personal preference) that the logarithm of the likelihood ratio converges to 0, using Lemma 3.5 (the "logarithmic version" of Lemma 3.3). The "fudge function" Φ is defined for convenience in equation (3.A.3). Starting from equation (3.5.11), tedious algebra shows that the logarithm of the likelihood ratio of hTPMH with respect to hPoMu is

$$- \log\left(\mathbb{P}(\text{Pois}(\lambda) \leq K_{d-1})\right) + \sum_{s=1}^{S} \Phi(K_{d-1}^{(s)} - n^{(s)}, n^{(s)}, \frac{1}{2})$$

$$- \Phi(K_{d-1} - n, n, \frac{1}{2}) + \sum_{s=1}^{S} n^{(s)}\left(\log(K_{d-1}^{(s)}) - \log(f^{(s)}K_{d-1})\right) + \mathscr{R}_{\text{MHg}}(\vec{\mathbf{n}}, \vec{\mathbf{K}}_{d-1}).$$

$$(3.5.12)$$

The only subtlety involves recalling the definitions of $K_{d-1}^{(s)}$ and K_{d-1} so as to correctly substitute them into (3.5.12):

$$\log(K_{d-1}^{(s)}) - \log(f^{(s)}K_{d-1}) = \log\left(\frac{f^{(s)}K - \sum_{\delta=1}^{d-1} N_\delta^{(s)}(0)}{f^{(s)}K - f^{(s)}\left(\sum_{\delta=1}^{d-1} N_\delta(0)\right)}\right).$$

$$(3.5.13)$$

3.6 Working Model for Compositional Heterogeneity

The hTPMH distribution uses a categorical distribution (the multivariate hypergeometric) which does not capture over-dispersion relative to the multinomial. In fact, its categorical distribution can also be derived using homogeneity assumptions, and only relaxes the assumptions regarding effectively infinite population size. So both hPoMu and hTPMH are unable to model over-dispersion of the categorical distribution caused by compositional heterogeneity. Thus the main change we need to make is to use a categorical distribution over-dispersed relative to the multinomial distribution.

In section 3.6.1 I introduce a working model which fills this gap, and then in section 3.6.2 I show that this working model extends what can be described using hPoMu alone.

3.6.1 Introduction to hPoDM

As suggested by its name, the "hierarchical Poisson Dirichlet-Multinomial" (hPoDM) distribution fills this gap by using a new distribution for the categorical distribution, the Dirichlet-Multinomial.

In section 3.6.1.1 I introduce the Dirichlet-Multinomial distribution. In section 3.6.1.3 I explain how the Dirichlet-Multinomial distribution models over-dispersion relative to the multinomial distribution. In section 3.6.1.4 I use the new information learned to define the hPoDM family of distributions.

3.6.1.1 Dirichlet-Multinomial (DM) Distribution

The PMF for the Dirichlet-Multinomial distribution may be written

$$\frac{\Gamma(\zeta_C S)}{\Gamma(\zeta_C S + n)} \cdot \left[\prod_{s=1}^{S} \frac{\Gamma(\zeta_C S f^{(s)} + n^{(s)})}{\Gamma\left(\zeta_C S f^{(s)}\right)} \right] \cdot \binom{n}{n^{(1)} \cdots n^{(S)}}. \tag{3.6.1}$$

More typically this is parameterized in terms of $\zeta_C S$. I parameterize in terms of ζ_C instead to facilitate interpretation. Parameterized this way, $\zeta_C = 1$ always corresponds to the same concentration as the uniform distribution on the $(S-1)$-dimensional simplex, regardless of the number of strains S.

3.6.1.2 Clarification of Compositional Heterogeneity

The intended notion of "compositional heterogeneity" herein does not include irregularities corresponding to e.g. strains preferentially grouping with or distancing from each other.

First, note that the Dirichlet-Multinomial can be thought of as a hierarchical distribution, where the $f^{(s)}$ parameters of the Multinomial distribution (living on the unit simplex due to the $\sum_{s=1}^{S} f^{(s)} = 1$ constraint) are Dirichlet distributed. (Whence the name "Dirichlet-Multinomial".) Second, while the marginal distributions of the Dirichlet distribution are

anti-correlated with one another, this is in some sense due only to the normalization condition, $\sum_{s=1}^{S} f^{(s)} = 1$. It can be shown that the joint distribution of mutually independent Gamma random variables, when normalized by their sum, is Dirichlet. In particular, the Dirichlet distribution can not model a situation where e.g. particular strains s_1 and s_2 tend to attach to one another within the sampling pool.

One way to get distributions on the unit simplex able to model such preferential grouping or distancing of strains is to use correlated random variables with distributions derived from Gamma distributions and again normalize by their sum. Such an approach would most closely resemble the intuitive idea of a "correlated Dirichlet distribution". Another approach is to use the logistic-normal distribution. See [BGF01] for a study applying methodology based on the latter approach to the study of multivariate count data from ecology.

3.6.1.3 Over-Disperson of Dirichlet-Multinomial with respect to Multinomial

The over-dispersion of the Dirichlet-Multinomial distribution with concentration $\zeta_C S$, total number n, and frequencies $\vec{f}$ with respect to the corresponding Multinomial distribution with total number n and frequencies $\vec{f}$ is

$$\frac{n-1}{1+\zeta_C S}. \tag{3.6.2}$$

Thus the concentration is approximately proportional to the reciprocal of the over-dispersion. (They are asymptotically equivalent up to a constant, i.e. the limit of their ratio as $\zeta_C \to \infty$ is the constant $n-1$.)

3.6.1.4 hPoDM Definition

Under the hPoDM distribution, for the d'th droplet: the probability of the strain distribution vector given the number of cells

$$\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right)$$

is determined by a Dirichlet-Multinomial distribution, while the probability of the number of cells $\mathbb{P}(N_d(0) = n)$ is determined by a Poisson distribution.

Thus for the unconditional probability:

$$\begin{aligned}
&\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) \\
={}& \mathbb{P}(N_d(0) = n) \cdot \mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right) \\
={}& \frac{e^{-\lambda}\lambda^n}{n!} \cdot \frac{\Gamma(\zeta_C S)}{\Gamma(\zeta_C S + n)} \cdot \left[\prod_{s=1}^{S} \frac{\Gamma(\zeta_C S f^{(s)} + n^{(s)})}{\Gamma\left(\zeta_C S f^{(s)}\right)}\right] \cdot \binom{n}{n^{(1)} \cdots n^{(S)}}.
\end{aligned} \tag{3.6.3}$$

3.6.2 hPoDM Generalizes hPoMu

Just like for hTPMH, hPoDM converges to the default hPoMu model under appropriate limits. Therefore when using the hPoDM family of distributions to model the distribution of the initial formation of droplets, we "retain the same language" used by the hPoMu distribution for that task.

In sections 3.6.2.1 and 3.6.2.2 I give the context needed and then show that the hNBDM family really is an extension of hPoMu. Because hPoDM and hPoMu have the same count distribution, this reduces to showing how the Dirichlet-Multinomial distribution extends the multinomial distribution.

3.6.2.1 Approximate Form of Likelihood Ratio of DM with respect to Multinomial and of hPoDM with respect to hPoMu and Bounds

Using Lemma 3.4 we get that

$$
\frac{\Gamma(\zeta_C S)}{\Gamma(\zeta_C S + n)} \cdot \left[\prod_{s=1}^{S} \frac{\Gamma(\zeta_C S f^{(s)} + n^{(s)})}{\Gamma(\zeta_C S f^{(s)})} \right]
$$
$$
= \exp\left(\mathscr{R}_{\mathrm{DM}}(\vec{\mathbf{n}}, \zeta_C S \vec{f}) \right) \cdot \left(\frac{\zeta_C S}{\zeta_C S + n} \right)^{\zeta_C S + n - \frac{1}{2}}
$$
$$
\cdot \left[\prod_{s=1}^{S} \left(\frac{\zeta_C S f^{(s)} + n^{(s)}}{\zeta_C S f^{(s)}} \right)^{\zeta_C S f^{(s)} + n^{(s)} - \frac{1}{2}} \right] \cdot \prod_{s=1}^{S} (f^{(s)})^{n^{(s)}},
$$
(3.6.4)

where (with the upper $\overline{\mathscr{R}}$ and lower $\underline{\mathscr{R}}$ functions as defined before)

$$
\underline{\mathscr{R}}(\zeta_C S \vec{f}, \zeta_C S \vec{f} + \vec{\mathbf{n}}) < \mathscr{R}_{\mathrm{DM}}(\vec{\mathbf{n}}, \zeta_C S \vec{f}) < \overline{\mathscr{R}}(\zeta_C S \vec{f}, \zeta_C S \vec{f} + \vec{\mathbf{n}}).
$$
(3.6.5)

Unfortunately the bounds in (3.6.5) appear to be tight in general only for larger values of ζ_C. (They become tighter as $\zeta_C \to \infty$; cf. section 3.6.2.2.)

Thus the likelihood ratio of the Dirichlet-Multinomial distribution with respect to the Multinomial distribution, which is the same as the likelihood ratio of the hPoDM distribution with respect to the hPoMu distribution, based on equation 3.6.4 above equals

$$
\exp\left(\mathscr{R}_{\mathrm{DM}}(\vec{\mathbf{n}}, \zeta_C S \vec{f}) \right) \cdot \left(\frac{\zeta_C S}{\zeta_C S + n} \right)^{\zeta_C S + n - \frac{1}{2}} \cdot \left[\prod_{s=1}^{S} \left(\frac{\zeta_C S f^{(s)} + n^{(s)}}{\zeta_C S f^{(s)}} \right)^{\zeta_C S f^{(s)} + n^{(s)} - \frac{1}{2}} \right].
$$
(3.6.6)

3.6.2.2 Sketch of Proof that hPoDM Converges in Distribution to hPoMu

Again, it suffices to show that the probability mass function of hPoDM converges to that of hPoMu. Cf. section 3.5.2.3.

That the expression in (3.6.6) approaches 1 as $\zeta_C \to \infty$ again follows from the Squeeze Theorem and Lemma 3.3. (As before, that the expressions involving the upper and lower Robbins functions approach 0 is an elementary consequence of the Archimedean property of the real numbers.) Thus the hPoDM distribution can be seen to converge to the hPoMu distribution as $\zeta_C \to \infty$.

Because the likelihood ratio of the Dirichlet-Multinomial distribution with respect to the Multinomial distribution is the same as the likelihood ratio of the hPoDM distribution with respect to the hPoMu distribution, it follows that the above also shows how the Dirichlet-Multinomial distribution converges to the Multinomial distribution as $\zeta_C \to \infty$.

Instead of showing that the likelihood ratio converges to 1, one can also show (e.g. according to one's personal preference) that the logarithm of the likelihood ratio converges to 0, using Lemma 3.5 (the "logarithmic version" of Lemma 3.3). The "fudge function" Φ is defined for convenience in equation (3.A.3). Then, starting from (3.6.6), tedious algebra shows that the logarithm of the likelihood ratio of (hPo)DM with respect to (hPo)Mu is

$$\sum_{s=1}^{S} \Phi(\zeta_C S f^{(s)}, n^{(s)}, n^{(s)} - \frac{1}{2}) - \Phi(\zeta_C S, n, n - \frac{1}{2}) + \mathscr{R}_{\mathrm{DM}}(\vec{\mathbf{n}}, \zeta_C S \vec{f}). \tag{3.6.7}$$

3.7   Working Model for Density Heterogeneity

The hPoDM working model has the same count distribution as hPoMu, and hence is unable to model over-dispersion of the count distribution caused by density heterogeneity. Thus the main change we need to make is from using the Poisson distribution as the count distribution to using a distribution over-dispersed relative to the Poisson distribution. Section 3.7.1 introduces a working model that fills this gap, and then section 3.7.2 shows that this working model can also be considered an extension of hPoMu.

3.7.1   Introduction to hNBDM

As suggested by its name, the "hierarchical Negative Binomial Dirichlet-Multinomial" (abbreviated as hNBDM) working model fills this gap by using a new distribution for the count distribution, the negative binomial. Section 3.7.1.1 introduces the negative binomial distribution. Section 3.7.1.2 explains how the negative binomial distribution models over-dispersion relative to the Poisson distribution. Section 3.7.1.3 defines the hNBDM working model.

3.7.1.1   Negative Binomial (NB) Distribution

One may write the PMF of the negative binomial distribution as

$$\frac{\Gamma(\zeta_D S + n)}{\Gamma(\zeta_D S)} \cdot \frac{(\zeta_D S)^{\zeta_D S}}{(\zeta_D S + \lambda)^{\zeta_D S + n}} \cdot \frac{\lambda^n}{n!} \tag{3.7.1}$$

where λ denotes the expected value of the distribution.

More typically this is parameterized in terms of $\zeta_D S$ instead of ζ_D. (Indeed in principle it should not even be necessary to specify a number of strains S in order to define this distribution.) I parameterize in terms of ζ_D to be compatible with the parameterization of the Dirichlet-Multinomial I use in section 3.6.1.1. See section 3.7.1.3 for how this pays off later by simplifying the math.

The parameter ζ_D is herein called the "density concentration". Using the negative binomial distribution for modelling density heterogeneity has precedent in the literature. For macroecological studies, the parameter λ "has been defined as the density of organisms in the area of interest" [WFY84]. Herein λ will be interpreted as the *average* cell density in the sampling pool. Previous work has also claimed that ζ_D should be able to capture variability in the density of organisms: "The practical use of $[\zeta_D S]$ ... in ecological studies of aggregation requires caution because $[\zeta_D S]$ is usually density-dependent" [CP89].

If a Poisson distribution with rate λ corresponds to a homogeneous cell density λ throughout the sampling pool, a negative binomial distribution with rate λ is interpreted to have the same *average* cell density. Larger values of density concentration ζ_D correspond to *smaller* heterogeneity around the average density, and thus more closely resemble the Poisson case (which is the limit as $\zeta_D \to \infty$), cf. figure 3.4a. Similarly, smaller values of density concentration ζ_D correspond to *larger* heterogeneity around the average density, cf. figure 3.4b.

3.7.1.2   Over-Disperson of Negative Binomial with respect to Poisson

The over-dispersion of the negative binomial distribution with concentration $\zeta_D S$ and mean λ with respect to the Poisson distribution with mean λ is

$$\frac{\lambda}{\zeta_D S} . \tag{3.7.2}$$

Thus the concentration is proportional to the reciprocal of the over-dispersion.

3.7.1.3   hNBDM Definition

For hNBDM we set $\zeta_D = \zeta_C =: \zeta$. So for the d'th droplet, the strain distribution vector $\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right)$ given the number of cells is Dirichlet-Multinomial distributed with parameter ζ, and the number of cells $\mathbb{P}(N_d(0) = n)$ is Negative Binomial distributed with parameter ζ. Thus the PMF defining the unconditional probability is

$$
\begin{aligned}
&\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) \\
&= \; \mathbb{P}(N_d(0) = n)\cdot\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right) \\
&\frac{\Gamma(\zeta S + n)}{\Gamma(\zeta S)}\cdot\frac{(\zeta S)^{\zeta S}}{(\zeta S + \lambda)^{\zeta S + n}}\cdot\frac{\lambda^n}{n!}\frac{\Gamma(\zeta S)}{\Gamma(\zeta S + n)}\cdot\left[\prod_{s=1}^{S}\frac{\Gamma(\zeta S f^{(s)} + n^{(s)})}{\Gamma\left(\zeta S f^{(s)}\right)}\right]\cdot\binom{n}{n^{(1)}\cdots n^{(S)}}
\end{aligned}
\tag{3.7.3}
$$

which via tedious algebraic manipulations can be seen to equal

$$\prod_{s=1}^{S}\left[\frac{\Gamma(\zeta Sf^{(s)}+n^{(s)})}{\Gamma(\zeta Sf^{(s)})}\cdot\frac{(\zeta Sf^{(s)})^{\zeta Sf^{(s)}}}{(\zeta Sf^{(s)}+f^{(s)}\lambda)^{\zeta Sf^{(s)}+n^{(s)}}}\cdot\frac{(f^{(s)}\lambda)^{n^{(s)}}}{(n^{(s)})!}\right].$$
(3.7.4)

The marginal distributions of the hNBDM distribution correspond to mutually independent Negative Binomial random variables whose parameters have been rescaled by the $f^{(s)}$.

This makes the relationship of the Negative Binomial distribution to the Dirichlet-Multinomial a direct analogue of the relationship of the Poisson distribution to the Multinomial Distribution. Unsurprisingly these observations have precedent in the literature. Cf. for example Theorem 1 of [Zho18].

3.7.2   hNBDM Generalizes hPoMu

Just like for hTPMH and hPoDM, hNBDM converges to the default hPoMu model under appropriate limits. Therefore when using the hNBDM family of distributions to model the distribution of the initial formation of droplets, we "retain the same language" used by the hPoMu distribution for that task.

In sections 3.7.2.1 and 3.7.2.2 I give the context needed and then show that, as a count distribution, the negative binomial distribution truly is a generalization of the Poisson. In sections 3.7.2.3 and 3.7.2.4 I give the context needed and then show that the hNBDM family really is an extension of hPoMu.

3.7.2.1   Approximate Form of Likelihood Ratio of Negative Binomial with respect to Poisson and Bounds

Applying Lemma 3.4, one has

$$\frac{\Gamma(\zeta_D S+n)}{\Gamma(\zeta_D S)}=e^{-n}\cdot\left(\frac{\zeta_D S+n}{\zeta_D S}\right)^{\zeta_D S+n-\frac{1}{2}}\cdot(\zeta_D S)^n\cdot\exp\left(\mathscr{R}_{\mathrm{NB}}(\zeta_D S,n)\right),$$
(3.7.5)

where $\mathscr{R}_{\mathrm{NB}}(\zeta_D S,n)$ satisfies

$$\frac{1}{12(\zeta_D S+n)+1}-\frac{1}{12\zeta_D S}<\mathscr{R}_{\mathrm{NB}}(\zeta_D S,n)<\frac{1}{12(\zeta_D S+n)}-\frac{1}{12\zeta_D S+1}.$$
(3.7.6)

It follows from equations (3.7.1) and (3.7.5) that the likelihood ratio of the negative binomial distribution (with mean λ) with respect to the Poisson distribution (with mean λ) is

$$e^{-n}\cdot\left(\frac{\zeta_D S+n}{\zeta_D S}\right)^{\zeta_D S+n-\frac{1}{2}}\cdot\left(\frac{\zeta_D S}{\zeta_D S+\lambda}\right)^{\zeta_D S+n}\cdot e^{\lambda}\cdot\exp\left(\mathscr{R}_{\mathrm{NB}}(\zeta_D S,n)\right).$$
(3.7.7)

3.7.2.2 Sketch of Proof that NB Converges in Distribution to Poisson

It follows nearly immediately from Lemma 3.3 that the above expression (3.7.7) converges to 1 as $\zeta_D \to \infty$. (Again, both of the bounds for $\mathscr{R}_{\text{NB}}(\zeta_D S, n)$ from (3.7.6) approach 0 for trivial reasons, allowing us to apply the Squeeze Theorem.) In particular, it follows that the negative binomial distribution converges to the Poisson distribution as $\zeta_D \to \infty$.

3.7.2.3 Approximate Form of Likelihood Ratio of hNBDM with respect to hPoMu and Bounds

Using either of the expressions (3.7.3) or (3.7.4) for the PMF of the hNBDM distribution given above, more tedious algebraic manipulations (including application of Lemma 3.4) allow one to show that the likelihood ratio of the hNBDM distribution relative to the corresponding hPoMu distribution is

$$\left[\prod_{s=1}^{S} e^{-n^{(s)}} \cdot \left(\frac{\zeta S f^{(s)} + n^{(s)}}{\zeta S f^{(s)}} \right)^{\zeta S f^{(s)} + n^{(s)} - \frac{1}{2}} \cdot \left(\frac{\zeta S f^{(s)}}{\zeta S f^{(s)} + f^{(s)}\lambda} \right)^{\zeta S f^{(s)} + n^{(s)}} \cdot e^{f^{(s)}\lambda} \right] \cdot \mathscr{R}_{\text{hNBDM}}(\zeta S \vec{f}, \vec{n}),$$

(3.7.8)

where

$$\mathscr{R}_{\text{hNBDM}}(\zeta S \vec{f}, \mathbf{n}) = \mathscr{R}_{\text{NB}}(\zeta S, n) + \mathscr{R}_{\text{DM}}(\zeta S \vec{f}, \mathbf{n}) = \sum_{s=1}^{S} \mathscr{R}_{\text{NB}}(\zeta S f^{(s)}, n^{(s)}). \qquad (3.7.9)$$

Observe that adding the bounds from (3.6.5) and (3.7.6) gives the following bounds for $\mathscr{R}_{\text{hNBDM}}(\zeta S \vec{f}, \vec{n})$:

$$\sum_{s=1}^{S} \left[\frac{1}{12(\zeta S f^{(s)} + n^{(s)}) + 1} - \frac{1}{12\zeta S f^{(s)}} \right] + \left(\frac{1}{12\zeta S + 1} - \frac{1}{12\zeta S} \right) + \left(\frac{1}{12(\zeta S + n) + 1} - \frac{1}{12(\zeta S + n)} \right)$$

$$= \quad \underline{\mathscr{R}}(\zeta S \vec{f}, \zeta S \vec{f} + \vec{n}) + \frac{1}{12(\zeta S + n) + 1} - \frac{1}{12\zeta S}$$

$$< \quad \mathscr{R}_{\text{DM}}(\zeta S \vec{f}, \vec{n}) + \mathscr{R}_{\text{NB}}(\zeta S, n) = \mathscr{R}_{\text{hNBDM}}(\zeta S \vec{f}, \vec{n}) \qquad (3.7.10)$$

$$< \quad \overline{\mathscr{R}}(\zeta S \vec{f}, \zeta S \vec{f} + \vec{n}) + \frac{1}{12(\zeta S + n)} - \frac{1}{12\zeta S + 1}$$

$$= \sum_{s=1}^{S} \left[\frac{1}{12(\zeta S f^{(s)} + n^{(s)})} - \frac{1}{12\zeta S f^{(s)} + 1} \right] + \left(\frac{1}{12\zeta S} - \frac{1}{12\zeta S + 1} \right) + \left(\frac{1}{12(\zeta S + n)} - \frac{1}{12(\zeta S + n) + 1} \right).$$

Compared to the above, observe that summing (over s) the bounds from (3.7.6) based

on (3.7.1) actually gives strictly tighter bounds:

$$\sum_{s=1}^{S} \left[\frac{1}{12(\zeta S f^{(s)} + n^{(s)}) + 1} - \frac{1}{12 \zeta S f^{(s)}} \right]$$

$$< \sum_{s=1}^{S} \mathscr{R}_{\mathrm{NB}}(\zeta S f^{(s)}, n^{(s)}) = \mathscr{R}_{\mathrm{hNBDM}}(\zeta S \vec{f}, \vec{\mathbf{n}}) \qquad (3.7.11)$$

$$< \sum_{s=1}^{S} \left[\frac{1}{12(\zeta S f^{(s)} + n^{(s)})} - \frac{1}{12 \zeta S f^{(s)} + 1} \right].$$

The additional terms from the bounds in (3.7.10) not present in the bounds from (3.7.11) are the result of twice applying Lemma 3.4 unnecessarily, once to $\frac{\Gamma(\zeta S + n)}{\Gamma(\zeta S)}$ and once to $\frac{\Gamma(\zeta S)}{\Gamma(\zeta S + n)}$. Because they cancel each other out, they obviously do not need to be bounded. Therefore the bounds do not contradict one another.

3.7.2.4 Sketch of Proof that hNBDM Converges in Distribution to hPoMu

Again, it suffices to show that the probability mass function of hNBDM converges to that of hPoMu. Cf. section 3.5.2.3.

That (3.7.8) converges to 1 as $\zeta \to \infty$ quickly follows from Lemma 3.3. This convergence implies that the hNBDM distribution converges to the hPoMu distribution as $\zeta \to \infty$.

Equivalently, the convergence of the hNBDM distribution to the hPoMu distribution as $\zeta \to \infty$ follows from the convergence of (3.7.7) to 1 as $\zeta \to \infty$ (replacing ζ, n, and λ with $\zeta f^{(s)}$, $n^{(s)}$, and $f^{(s)} \lambda$ respectively). Because (3.7.8) is the product of (3.7.7) and (3.6.6), it also equivalently follows from the convergence of both (3.7.7) and (3.6.6) to 1 as $\zeta \to \infty$.

Instead of showing that the likelihood ratio converges to 1, one can also show (e.g. according to one's personal preference) that the logarithm of the likelihood ratio converges to 0, using Lemma 3.5 (the "logarithmic version" of Lemma 3.3). The "fudge function" Φ is defined for convenience in equation (3.A.3). Then, starting from (3.7.8), tedious algebra shows that the logarithm of the likelihood ratio of hNBDM with respect to hPoMu is

$$\lambda - n + \sum_{s=1}^{S} \Phi(\zeta S f^{(s)}, n^{(s)}, n^{(s)} - \frac{1}{2}) - \sum_{s=1}^{S} \Phi(\zeta S f^{(s)}, f^{(s)} \lambda, n^{(s)}) + \mathscr{R}_{\mathrm{hNBDM}}(\zeta S \vec{f}, \vec{\mathbf{n}}). \quad (3.7.12)$$

3.8 Working Model for Arbitrary Combinations of Heterogeneities

While the hNBDM family of distributions is useful for modelling density heterogeneity, it does have the limitation that the value of density heterogeneity is assumed to be coupled with corresponding value of compositional heterogeneity. By relaxing this restriction we can

generalize the hNBDM family to the "generalized hNBDM" or "ghNBDM" family of distributions. This makes it possible to model arbitrary combinations of density and compositional heterogeneities.

3.8.1   ghNBDM Family Definition

Under a distribution in the ghNBDM family, for the d'th droplet: the probability for the strain distribution vector given the number of cells $\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right)$ is described by a Dirichlet-Multinomial distribution with concentration parameter ζ_C, while the probability for the number of cells $\mathbb{P}(N_d(0) = n)$ is described by a Negative binomial distribution with concentration parameter ζ_D. Thus for the unconditional probability:

$$
\begin{aligned}
&\mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}}\right) \\
=\ & \mathbb{P}(N_d(0) = n) \cdot \mathbb{P}\left(\vec{\mathbf{N}}_d(0) = \vec{\mathbf{n}} \mid N_d(0) = n\right) \\
=\ & \frac{\Gamma(\zeta_D S + n)}{\Gamma(\zeta_D S)} \cdot \frac{(\zeta_D S)^{\zeta_D S}}{(\zeta_D S + \lambda)^{\zeta_D S + n}} \cdot \frac{\lambda^n}{n!} \cdot \frac{\Gamma(\zeta_C S)}{\Gamma(\zeta_C S + n)} \cdot \left[\prod_{s=1}^{S} \frac{\Gamma(\zeta_C S f^{(s)} + n^{(s)})}{\Gamma\left(\zeta_C S f^{(s)}\right)}\right] \cdot \binom{n}{n^{(1)} \cdots n^{(S)}} \cdot
\end{aligned}
$$
(3.8.1)

Unlike for the hNBDM family, for arbitrary members of the ghNBDM family one may have that $\zeta_C \neq \zeta_D$. In general the marginal distributions of members of the ghNBDM family are *not* mutually independent (and in fact have non-zero cross-covariance) and need not be negative binomial distributed.

The hNBDM family is the special case of the ghNBDM family when $\zeta_C = \zeta_D$. The hPoDM and hPoMu distributions emerge as limiting cases of the ghNBDM family, or are bona fide members if we allow "∞" as a valid concentration parameter value. Cf. figure 3.6. The case where ζ_C is allowed to approach infinity while ζ_D remains finite corresponds to the hierarchical Negative Binomial Multinomial (hNBMu) distribution.
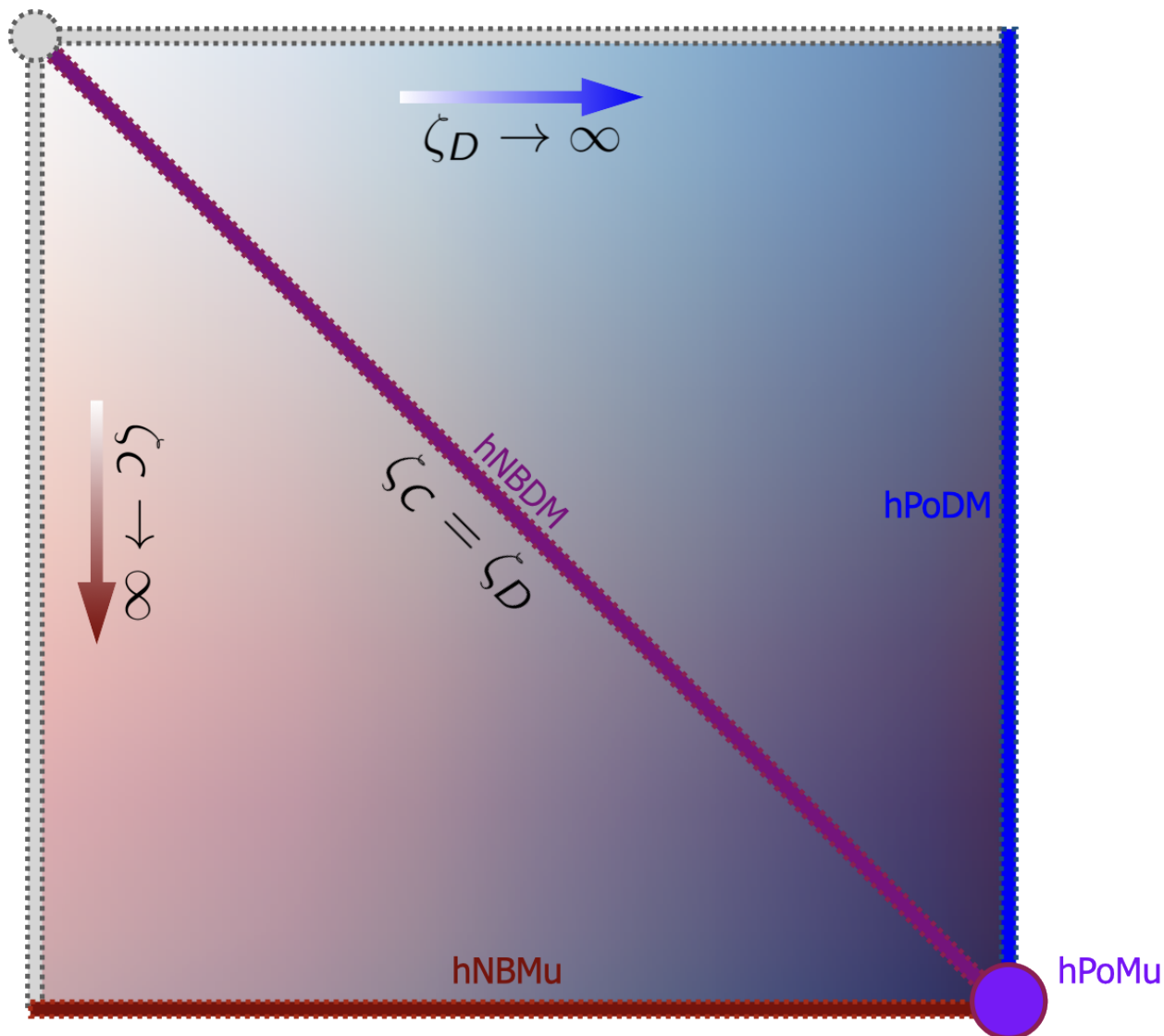
Figure 3.6: Schematic of the ghNBDM family of distributions, assuming a fixed λ and vector of population strain frequencies $\vec{f}$.

3.8.2   Over-Dispersions for these Distributions

Since the non-diagonal covariances of the hPoMu distribution are 0, it of course does not make sense to speak of over-dispersion relative to them using the definition (3.4.2). However, we may still sensibly speak of the over-dispersion of their marginals with respect to the corresponding marginals of the hPoMu distribution (which unsurprisingly we will now do).

The over-dispersion of the s'th marginal of the ghNBDM distribution with respect to the

s'th marginal of the hPoMu distribution is

$$\frac{[(1 + \zeta_D S) + f^{(s)}(\zeta_C S - \zeta_D S)]}{(1 + \zeta_C S)(\zeta_D S)} \lambda \,. \tag{3.8.2}$$

Observe how this value depends on the particular s. Moreover, the expression (3.8.2) above:

- reduces to (3.8.3) when $\zeta_D = \zeta_C$,

- approaches (3.8.4) when $\zeta_D \to \infty$,

- approaches (3.8.5) when $\zeta_C \to \infty$,

- approaches 0 when both $\zeta_D \to \infty$ and $\zeta_C \to \infty$.

The last corresponds to hPoMu, of course. Details about the remaining special cases (hNBDM, hPoDM, hNBMu) are given below.

3.8.2.1   hNBDM Over-dispersion

The over-dispersion of the s'th marginal of the hNBDM distribution with respect to the s'th marginal of the hPoMu distribution is

$$\frac{\lambda}{\zeta S} \,, \tag{3.8.3}$$

which interestingly does *not* depend on the particular strain s, and moreover equals the over-dispersion of a Negative Binomial distribution with respect to a Poisson distribution, cf. (3.7.2).

The latter is to be expected given that the marginals of the hPoMu distribution are $\text{Pois}(f^{(s)}\lambda)$ distributed and the marginals of the hNBDM distribution are $\text{NB}(\zeta S f^{(s)}, f^{(s)}\lambda)$ distributed. Note that (3.8.3) also tends to 0 as $\zeta \to \infty$ as it should. Finally observe how the over-dispersion (3.8.3) is inversely proportional to the concentration parameter ζ, just like (3.7.2) which it equals.

3.8.2.2   hPoDM Over-dispersion

The over-dispersion of the s'th marginal of the hPoDM distribution with respect to the s'th marginal of the hPoMu distribution is

$$\frac{1 - f^{(s)}}{1 + \zeta_C S} \lambda \,. \tag{3.8.4}$$

Observe how this value depends on the particular s. And of course (3.8.4) tends to 0 as $\zeta_C \to \infty$. Further observe how the over-dispersion (3.8.4) is (asymptotically) proportional to the inverse of the concentration parameter ζ_C, just like (3.6.2).

3.8.2.3 hNBMu Over-dispersion

The over-dispersion of the s'th marginal of the hNBMu distribution with respect to the s'th marginal of the hPoMu distribution is

$$\frac{f^{(s)}\lambda}{\zeta_D S} \,.\tag{3.8.5}$$

Observe how this value depends on the particular s. And of course (3.8.5) tends to 0 as $\zeta_D \to \infty$. Further observe how the over-dispersion (3.8.5) is proportional to the inverse of the concentration parameter ζ_D, just like (3.6.2).

3.9 Conclusion

Findings and Contributions   We have identified

- a default working model we can use to for the initial droplet formation,

- key implicit assumptions made by that default working model,

- concrete ways those assumptions could fail to be true in practice, and

- working models generalizing the default working model which can account for any of those failed assumptions.

Practical Implications   We can now start to ask how realistic the default working model is and what changes might need to be made to use a more realistic working model instead. Being able to use a more realistic working model means that the data-generating process for MOREI can be simulated more accurately. Being able to simulate the data-generating process for MOREI more accurately will ultimately help to identify best-performing methods for characterizing microbial interactions based on MOREI data, cf. part III.

Next Steps and Open Questions   In chapter 4 we compare the default model with several of these alternative models to see whether there are noticeable differences. In chapter 5 we look at the effects of these alternative models on the throughput of data useful for characterizing microbial interactions. Finally in chapter 6 we propose and investigate the performance of estimators which will allow us to ascertain which of these models is most realistic in practice. We want to know how much the default working model differs from other, potentially more realistic working models and how to quantify any such differences.

Chapter 3. Appendices

Appendix 3.A Results for Bounding Approximate Likelihood Ratios

Lemma 3.1. *Using Stirling's Approximation as given in [Rob55],*

$$
\frac{1}{12(n+m)+1} - \frac{1}{12n}
$$
$$
< \log \left(\left(\frac{(n+m)!}{n!} \right) \middle/ \left(e^{-m} \left(\frac{n+m}{n} \right)^{n+m+\frac{1}{2}} n^m \right) \right) \tag{3.A.1}
$$
$$
< \frac{1}{12(n+m)} - \frac{1}{12n+1} \, .
$$

    Proof: This follows basically immediately from the bounds in [Rob55] via simple but tedious algebraic manipulations. $\square$

    Heuristically, this says that $\frac{(n+m)!}{n!} \approx n^m$ when n is large. This is because the expression $\left(\frac{n+m}{n} \right)^{n+m+\frac{1}{2}}$ approaches e^m as $n \to \infty$ by Lemma 3.3.

Lemma 3.2. *Given $\alpha, c, y_1, y_2 \in \mathbb{R}$ (in particular c is constant w.r.t. x):*

$$
\lim_{x \to \infty} \left(\frac{\alpha x + y_1}{\alpha x + y_2} \right)^c = 1 \, .
$$

Proof: By properties of limits, since $f(x) = x^c$ is continuous, this equals

$$
\left[\lim_{x \to \infty} \left(\frac{\alpha x + y_1}{\alpha x + y_2} \right) \right]^c ,
$$

which by L'Hôspital's Rule equals $1^c = 1$. $\square$

Lemma 3.3. *Given $\alpha, c, y_1, y_2 \in \mathbb{R}$, $\alpha > 0$, one has*

$$
\lim_{x \to \infty} \left(\frac{\alpha x + y_1}{\alpha x + y_2} \right)^{\alpha x + c} = e^{y_1} e^{-y_2} \, .
$$

Proof: By properties of limits, the LHS of the above equals

$$\left[\lim_{x\to\infty}\left(\frac{\alpha x + y_1}{\alpha x + y_2}\right)^{\alpha x}\right]\left[\lim_{x\to\infty}\left(\frac{\alpha x + y_1}{\alpha x + y_2}\right)^{c}\right] = \lim_{x\to\infty}\left(\frac{\alpha x + y_1}{\alpha x + y_2}\right)^{\alpha x},$$

where the second equality uses Lemma 3.2. Rearranging further:

$$= \lim_{x\to\infty}\left[\left(\left(\frac{x + \frac{y_1}{\alpha}}{x}\right)^{x}\right)^{\alpha}\left(\left(\frac{x + \frac{y_2}{\alpha}}{x}\right)^{x}\right)^{-\alpha}\right],$$

from which the result more or less immediately follows from the properties of limits and the definition of the exponential function. *(In particular because $f_1(x) = x^{\alpha}$ and $f_2(x) = x^{-\alpha}$ are continuous.)* $\square$

Lemma 3.4. *Using the generalization of the bounds on Stirling's approximation to the entire Γ function as given in Theorem 5 of [Gor94]*

$$\frac{1}{12(x+\beta)+1} - \frac{1}{12x}$$
$$< \log\left(\left(\frac{\Gamma(x+\beta)}{\Gamma(x)}\right)\middle/\left(e^{-\beta}\left(\frac{x+\beta}{x}\right)^{x+\beta-\frac{1}{2}}x^{\beta}\right)\right) \tag{3.A.2}$$
$$< \frac{1}{12(x+\beta)} - \frac{1}{12x+1}.$$

Proof: Lemma 3.4 follows almost immediately from Theorem 5 of [Gor94] via tedious algebraic manipulations. $\square$

Heuristically, Lemma 3.4 can be interpreted as saying that $\frac{\Gamma(x+\beta)}{\Gamma(x)} \approx x^{\beta}$ for large x. Again, Lemma 3.3 implies that $\left(\frac{x+\beta}{x}\right)^{x+\beta-\frac{1}{2}}$ approaches e^{β} as $x \to \infty$.

The "fudge function" Φ is defined as

$$\Phi(x, y, z) := (x + z)\left[\log(x + y) - \log(x)\right]. \tag{3.A.3}$$

Lemma 3.5. *Given $y, z \in \mathbb{R}$ which are constant with respect to $x \in \mathbb{R}$,*

$$\lim_{x\to\infty}\Phi(x, y, z) = y.$$

Proof: This is a corollary of Lemma 3.3. Specifically, because the logarithm is continuous and continuous functions commute with limits,

$$\lim_{x\to\infty}\Phi(x, y, z) = \log\left(\lim_{x\to\infty}\exp(\Phi(x, y, z))\right) = \log\left(\lim_{x\to\infty}\left(\frac{x+y}{x}\right)^{x+z}\right)$$

$$= \log(e^y) = y,$$

where the penultimate equality was an application of 3.3. $\square$

3.A.1 Comments about Robbins-like Bounds

The bounds on $n!$ given in [Rob55] follow from those for the Γ function given in Theorem 5 of [Gor94] using the identity $n! = n\Gamma(n) = n(n-1)!$ (for all $n \in \mathbb{N}$). They are still slightly different though, because for $n \in \mathbb{N}$ the expression from [Gor94] gives $\Gamma(n) = \frac{n!}{n} = (n-1)! \sim \sqrt{2\pi}n^{n-\frac{1}{2}}e^{-n}$, whereas *directly*[12] applying the expression from [Rob55] to $(n-1)$ gives $(n-1)! \sim \sqrt{2\pi}(n-1)^{n-\frac{1}{2}}e^{-(n-1)}$. Although these expressions are technically distinct, it follows from Lemma 3.3 that they are asymptotically equivalent ($\sim$, the limit of their ratio approaches 1 as $n \to \infty$). Therefore they are consistent with one another.

Appendix 3.B Cumulants of Hierarchical Count-Categorical Distributions

The following are consequences of the law of total covariance (of which the law of total variance is a special case) and in turn from the law of total expectation.

Lemma 3.6. *If $\vec{\mathbf{X}}$ is such that $\vec{\mathbf{X}}|N \sim \mathrm{Mult}(N, \vec{f})$ or $\vec{\mathbf{X}}|N \sim \mathrm{DirMult}(\zeta_C S, N, \vec{f})$, then for all $s \in [S]$:*

$$\mathbb{E}[X_s] = f^{(s)}\mathbb{E}[N] \,.$$

Lemma 3.7. *If $\vec{\mathbf{X}}$ is such that $\vec{\mathbf{X}}|N \sim \mathrm{Mult}(N, \vec{f})$, then for all $s \in [S]$:*

$$\mathrm{Var}[X_s] = (f^{(s)})^2 \mathrm{Var}[N] + f^{(s)}(1 - f^{(s)})\mathbb{E}[N] \,. \tag{3.B.1}$$

Lemma 3.8. *If $\vec{\mathbf{X}}$ is such that $\vec{\mathbf{X}}|N \sim \mathrm{DirMult}(\zeta_C S, N, \vec{f})$, then for all $s \in [S]$:*

$$\mathrm{Var}[X_s] = \left[\frac{f^{(s)}(1 - f^{(s)})}{1 + \zeta_C S} + (f^{(s)})^2\right] \mathrm{Var}[N] + \frac{f s(1 - f^{(s)})}{1 + \zeta_C S}\mathbb{E}[N](\mathbb{E}[N] + \zeta_C S) \,. \tag{3.B.2}$$

Observe how (using L'Hôspital's rule) equation (3.B.2) approaches (3.B.1) as $\zeta_C \to \infty$.

Lemma 3.9. *If $\vec{\mathbf{X}}$ is such that $\vec{\mathbf{X}}|N \sim \mathrm{Mult}(N, \vec{f})$, then for all $s_1, s_2 \in [S]$ ($s_1 \neq s_2$):*

$$\mathrm{Cov}[X_{s_1}, X_{s_2}] = f^{(s_1)}f^{(s_2)}\left(\mathrm{Var}[N] - \mathbb{E}[N]\right) \,. \tag{3.B.3}$$

Lemma 3.10. *If $\vec{\mathbf{X}}$ is such that $\vec{\mathbf{X}}|N \sim \mathrm{DirMult}(\zeta_C S, N, \vec{f})$, then for all $s_1, s_2 \in [S]$ ($s_1 \neq s_2$):*

$$\mathrm{Cov}[X_{s_1}, X_{s_2}] = f^{(s_1)}f^{(s_2)}\left(\mathrm{Var}[N] - \frac{1}{1 + \zeta_C S}\left[\zeta_C S\mathbb{E}[N] + \mathbb{E}[N^2]\right]\right) \,. \tag{3.B.4}$$

Observe how (using L'Hôspital's rule) equation (3.B.4) approaches (3.B.3) as $\zeta_C \to \infty$.

---

[12]As opposed to indirectly, by dividing the expression for $n!$ by n.

3.B.1 Variance and Covariance of ghNBDM

From Lemma 3.8, for the ghNBDM distributions, for each $s \in [S]$:

$$\mathrm{Var}[N^{(s)}(0)] = f^{(s)}\lambda + \frac{[f^{(s)}(1 + \zeta_D S) + (f^{(s)})^2(\zeta_C S - \zeta_D S)]}{(1 + \zeta_C S)(\zeta_D S)}\lambda^2 \,. \tag{3.B.5}$$

Note how (3.B.5) reduces to (3.B.9) when $\zeta_D = \zeta_C$. Similarly, (3.B.5) approaches (3.B.11) as $\zeta_D \to \infty$, and (3.B.5) approaches (3.B.13) as $\zeta_C \to \infty$, both conclusions following from L'Hôspital's rule.

From Lemma 3.10, for the ghNBDM distributions, for each $s_1, s_2 \in [S]$ ($s_1 \neq s_2$):

$$\mathrm{Cov}[N^{(s_1)}(0), N^{(s_2)}(0)] = f^{(s_1)}f^{(s_2)}\left[\frac{\zeta_C S - \zeta_D S}{(\zeta_D S)(1 + \zeta_C S)}\right]\lambda^2 \,. \tag{3.B.6}$$

Note how (3.B.6) reduces to (3.B.10) when $\zeta_D = \zeta_C$. Similarly, (3.B.6) approaches (3.B.12) as $\zeta_D \to \infty$, and (3.B.6) approaches (3.B.14) as $\zeta_C \to \infty$, both conclusions again following from L'Hôspital's rule. Information about special cases (hPoMu, hNBDM, hPoDM, hNBMu) is given below.

3.B.1.1 Variance and Covariance of hPoMu

For the hPoMu distribution, for each $s \in [S]$:

$$\mathrm{Var}[N^{(s)}(0)] = f^{(s)}\lambda \,, \tag{3.B.7}$$

and for all $s_1, s_2 \in [S]$ ($s_1 \neq s_2$):

$$\mathrm{Cov}[N^{(s_1)}(0), N^{(s_2)}(0)] = 0 \,. \tag{3.B.8}$$

This follows either from applying Lemmas 3.7 and 3.9 directly, or recalling that the marginals of the hPoMu distribution are mutually independent $\mathrm{Pois}(f^{(s)}\lambda)$ random variables.

3.B.1.2 Variance and Covariance of hNBDM

For the hNBDM distribution, for each $s \in [S]$:

$$\mathrm{Var}[N^{(s)}(0)] = f^{(s)}\lambda + \frac{f^{(s)}}{\zeta S}\lambda^2 \,, \tag{3.B.9}$$

and for all $s_1, s_2 \in [S]$ ($s_1 \neq s_2$):

$$\mathrm{Cov}[N^{(s_1)}(0), N^{(s_2)}(0)] = 0. \tag{3.B.10}$$

These follow from applying Lemmas 3.8 and 3.10 directly, or recalling that the marginals of the hNBDM distribution are mutually independent $\mathrm{NB}(\zeta S f^{(s)}, f^{(s)}\lambda)$ random variables. Notice also how (3.B.9) approaches (3.B.7), as $\zeta \to \infty$, as it should. Of course (3.B.10) already equals (3.B.8).

3.B.1.3 Variance and Covariance of hPoDM

For the hPoDM distribution, for each $s \in [S]$:

$$\mathrm{Var}[N^{(s)}(0)] = f^{(s)}\lambda + \frac{f^{(s)}(1 - f^{(s)})}{1 + \zeta_C S}\lambda^2 , \tag{3.B.11}$$

and for all $s_1, s_2 \in [S]$ $(s_1 \neq s_2)$:

$$\mathrm{Cov}[N^{(s_1)}(0), N^{(s_2)}(0)] = -\frac{f^{(s_1)}f^{(s_2)}}{1 + \zeta_C S}\lambda^2 . \tag{3.B.12}$$

These follow from applying Lemmas 3.8 and 3.10. Notice also how (3.B.11) approaches (3.B.7), and (3.B.12) approaches (3.B.8), as $\zeta_C \to \infty$, as it should.

3.B.1.4 Variance and Covariance of hNBMu

From Lemma 3.7, for the hNBMu distribution, for each $s \in [S]$:

$$\mathrm{Var}[N^{(s)}(0)] = f^{(s)}\lambda + \frac{(f^{(s)})^2}{\zeta_D S}\lambda^2 . \tag{3.B.13}$$

As expected, (3.B.13) approaches (3.B.7) as $\zeta_D \to \infty$.
   From Lemma 3.9, for the hNBMu distribution, for each $s_1, s_2 \in [S]$ $(s_1 \neq s_2)$:

$$\mathrm{Cov}[N^{(s_1)}(0), N^{(s_2)}(0)] = \frac{f^{(s_1)}f^{(s_2)}}{\zeta_D S}\lambda^2 . \tag{3.B.14}$$

Again, as expected, (3.B.14) approaches (3.B.8) as $\zeta_D \to \infty$.

Chapter 4

Log-Likelihood Ratio Statistics for Goodness of Fit

Herein I show that log likelihood ratios can be used to determine whether failure of assumptions from section 3.3.3 causes substantially new behavior. See sections 4.2, 4.3.2, and 4.4. Failure of the sampling without replacement assumption has negligible effects in practice, but failures of the other assumptions could be important. See sections 4.4 and 4.5.

Section 4.1 provides context for the chapter. Section 4.2 explains some of the technical details of the comparisons. Section 4.3 explains the implementation details of the comparisons. Section 4.4 describes and explains the comparisons. Section 4.5 interprets the results from section 4.4, explaining what they mean for modelling the initial formation of droplets. Section 4.6 summarizes the findings and concludes with implications and next steps.

Chapter 4 Contents

4.1   Background and Significance

The overall scientific problem remains how to design the experiment so that enough data is collected from a single experimental run to adequately power statistical inference, without additional costly and time-consuming experimental runs being required. Cf. section 2.3.2. As argued in section 2.3, this amounts to knowing the joint distribution of the strain counts well enough to predict the probabilities associated with any possible combination of strains. In the previous chapter 3, I identified a simple working model (hPoMu) for the initial formation of droplets which practitioners are likely to assume is unquestionably accurate. I explained the assumptions required to justify such a simple working model, and presented working models resulting from relaxing some of the most unreasonable of those assumptions.

Broader field   The problem considered in this chapter falls under the general field of "model selection". See [HTF09, chapter 7] for a more widely accessible introduction to the general theory of model selection, or [BBM99] for a highly sophisticated discussion thereof. Cf. the discussion from the introduction to Part II. Herein we only need to consider the subfield of model selection corresponding to so-called "goodness of fit" tests, in particular those using the classic likelihood ratio statistic. For a basic introduction, tailored for non-statisticians, on the use of the likelihood ratio statistic for model selection with ecological count data, see [LBG10]. For a more sophisticated discussion of the likelihood ratio statistic, see [Kee10, chapter 17] or [Vaa98, chapter 16] for information such as a general definition, or the statement and proof of the result about its asymptotic χ^2 distribution.

Specific problem   Herein I investigate whether the failure of any of the implicit hPoMu assumptions from section 3.3.3 causes substantially new behavior. As explained previously, failure of any of these assumptions can be associated with a new working model that generalizes hPoMu. Therefore, if data generated by a given such generalized working model fits

hPoMu well, it seems reasonable to say that the generalized working model, and thus also the corresponding assumption failure, is not causing substantially new behavior.

On the other hand, given two working models which are both adequate for describing the initial droplet formation process, the simpler working model will be easier to understand. Therefore, given two working models with the same descriptive power, we should choose the simpler working model. It follows that justifying the use of a more complicated working model over a simpler working model requires showing that the more complicated working model has substantially more descriptive power than the simpler working model.

Therefore, to justify potentially using any of the other, more complicated working models in place of hPoMu, we must first ascertain whether any of the other working models have substantially more descriptive power than hPoMu. Thus we want to ascertain how well hPoMu can describe data generated by the other working models. Only those working models which do not always produce data that can be described well by hPoMu have substantially more descriptive power than hPoMu. Any other working models are not as useful.

In statistical terms, we know that a working model as simple as hPoMu will inevitably lead to biased estimators. So we want to investigate whether the failure of any of these assumptions can cause estimates derived from hPoMu to become biased enough that they are no longer useful. Thinking of the estimates produced by each of these working models as an algorithm, in the targeted learning roadmap [vdLR11] we want to investigate whether we can rely on a single "default" algorithm based on questionable assumptions.

Particular approach In this chapter I describe the information needed to implement goodness of fit tests based on the log-likelihood ratio statistic that compare any one of these more complicated working models against the "default" working model (hPoMu). After generating data from any of these working models whose assumptions are more relaxed than those of hPoMu, if these tests demonstrate that the resulting data are significantly different than what would be expected under the null model (hPoMu), then this suggests that the predictions made by hPoMu are inadequate. We could conclude that we may need to pay attention to whether the corresponding assumption of hPoMu is violated in practice. In contrast, if the tests demonstrate no significant difference between the resulting data and what would be expected under the null model (hPoMu), then this suggests we do not need to care whether the corresponding assumption of hPoMu is violated in practice.

Comparing the fit of two models for a given dataset by comparing the values of their likelihoods is a straightforward idea. In turn the (arguably) simplest way to make that comparison is to consider (the logarithm of) the ratio of those likelihoods. Because this is exploratory and preliminary work, there does not seem to be any compelling reason to use approaches that are more sophisticated than the simplest possible.

For problems similar to these, either the Wald statistic or the Rao score statistic is often used in place of the likelihood ratio statistic. However, in addition to being less straightforward, these statistics are also "equivalent" (in a sense that can be made precise) to the likelihood ratio statistic. See [Vaa98, section 16.2] or [Kee10, section 17.4] for discussion

of the relationship of the Wald, Rao score, and likelihood ratio statistics.

Similarly, the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) methods can both (roughly) be thought of as more sophisticated versions of the likelihood ratio statistic. See, for example, [BA02], in particular [BA02, section 6.9.3]. Thus the same preference for simplicity described above is also a reason for not using AIC or BIC.

The "null model" in this case, hPoMu, is (effectively by design) nested within all of these generalized working models against which it is being compared. Nevertheless, as explained in [LBG10], the assumption that the null model is nested within the alternative model is not necessary in general to make comparisons using the likelihood ratio statistic. So failure of nesting in other contexts is not a reason to avoid using the likelihood ratio statistic.

In comparing the performances of "Negative Binomial Factor Analysis" (related to the hNBDM model) and "Poisson Factor Analysis" (related to the hPoMu model), it can be argued that [Zho18, section 6] effectively addresses a problem similar to the one considered in this chapter. However, the underlying questions being asked therein are substantially different from those questions being asked herein. Hence, there seems to be no reason to expect that the comparison from [Zho18, section 6] would map (well) to this problem.

4.2 Preliminaries

Below I give the formulae for the log likelihood ratios comparing more complex working models to hPoMu. These exact formulae for the log likelihood ratios do not invoke any version of Stirling's approximation. I used these formulae as described in section 4.3.2.

4.2.1 hTPMH

The likelihood ratio of the hTPMH distribution with respect to the corresponding hPoMu distribution is

$$\mathbb{P}(\text{Pois}(\lambda) \leq K_{d-1})^{-1} \cdot \frac{(K_{d-1} - n)!}{(K_{d-1})!} \cdot \left[\prod_{s=1}^{S} \frac{(K_{d-1}^{(s)})!}{(K_{d-1}^{(s)} - n^{(s)})!} \right] \cdot \left[\prod_{s=1}^{S} (f^{(s)})^{n^{(s)}} \right]^{-1}. \tag{4.2.1}$$

Therefore the log likelihood ratio is

$$-\log\left(\mathbb{P}(\text{Pois}(\lambda) \leq K_{d-1})\right) + \log\left((K_{d-1} - n)!\right) - \log\left((K_{d-1})!\right) + \\ \sum_{s=1}^{S} \left[\log\left((K_{d-1}^{(s)})!\right) - \log\left((K_{d-1}^{(s)} - n^{(s)})!\right) \right] - \sum_{s=1}^{S} n^{(s)} \log(f^{(s)}). \tag{4.2.2}$$

Using the identity $\Gamma(n+1) = n!$, the above can be rewritten using the log-gamma function[1] (which is directly implemented in SciPy [VGO$^+$20]):

$$-\log\left(\mathbb{P}(\text{Pois}(\lambda) \leq K_{d-1})\right) + \log\left(\Gamma(K_{d-1} - n + 1)\right) - \log\left(\Gamma(K_{d-1} + 1)\right) +$$

$$\sum_{s=1}^{S}\left[\log\left(\Gamma(K_{d-1}^{(s)} + 1)\right) - \log\left(\Gamma(K_{d-1}^{(s)} - n^{(s)} + 1)\right)\right] - \sum_{s=1}^{S} n^{(s)}\log(f^{(s)}). \tag{4.2.3}$$

As an aside, because the above expression equals (3.5.12) from section 3.5.2.3, one can derive an explicit formula for $\mathscr{R}_{\text{MHg}}(\vec{\mathbf{n}}, \vec{\mathbf{K}}_{d-1})$. It is not very elegant.

For all observed values of K_{d-1}, $\mathbb{P}(\text{Pois}(\lambda) \leq K_{d-1})^{-1}$ was indistinguishable from 1 up to numerical precision. (Indeed for $\lambda = 2$ that probability appears to be indistinguishable from 1 up to numerical precision already for $K_{d-1} \geq 20$, values many orders of magnitude smaller than any observed value.) Therefore the contribution of the logarithm of this term was treated as if it was exactly 0 in the code implementation and not explicitly included.

4.2.2   hPoDM

The likelihood ratio of the hPoDM distribution with respect to the corresponding hPoMu distribution is

$$\frac{\Gamma(\zeta_C S)}{\Gamma(\zeta_C S + n)} \cdot \left[\prod_{s=1}^{S} \frac{\Gamma(\zeta_C S f^{(s)} + n^{(s)})}{\Gamma\left(\zeta_C S f^{(s)}\right)}\right] \cdot \left[\prod_{s=1}^{S}(f^{(s)})^{n^{(s)}}\right]^{-1}. \tag{4.2.4}$$

Therefore the log likelihood ratio is

$$\log\left(\Gamma(\zeta_C S)\right) - \log\left(\Gamma(\zeta_C S + n)\right) + \sum_{s=1}^{S}\left[\log\left(\Gamma(\zeta_C S f^{(s)} + n^{(s)})\right) - \log\left(\Gamma\left(\zeta_C S f^{(s)}\right)\right)\right] - \sum_{s=1}^{S} n^{(s)}\log(f^{(s)}).$$
$$\tag{4.2.5}$$

As an aside, because the above expression equals (3.6.7) from section 3.6.2.2, one can derive an explicit formula for $\mathscr{R}_{\text{DM}}(\vec{\mathbf{n}}, \zeta_C S\vec{f})$. Again, it is not elegant.

4.2.3   hNBDM

The likelihood ratio of the hNBDM distribution with respect to the corresponding hPoMu distribution is

$$\prod_{s=1}^{S}\left[\frac{\Gamma(\zeta S f^{(s)} + n^{(s)})}{\Gamma(\zeta S f^{(s)})} \cdot \frac{(\zeta S f^{(s)})^{\zeta S f^{(s)}}}{(\zeta S f^{(s)} + f^{(s)}\lambda)^{\zeta S f^{(s)} + n^{(s)}}}\right] \cdot \left[\prod_{s=1}^{S} e^{-f^{(s)}\lambda}\right]^{-1}. \tag{4.2.6}$$

---

[1] Which is most likely more numerically stable and computationally tractable than directly computing the logarithm of the factorial.

Therefore the log likelihood ratio is

$$\sum_{s=1}^{S} \left[\log \left(\Gamma(\zeta S f^{(s)} + n^{(s)}) \right) - \log \left(\Gamma(\zeta S f^{(s)}) \right) + \zeta S f^{(s)} \log \left(\zeta S f^{(s)} \right) \right.$$

$$\left. - (\zeta S f^{(s)} + n^{(s)}) \log \left(\zeta S f^{(s)} + f^{(s)} \lambda \right) \right] + \sum_{s=1}^{S} f^{(s)} \lambda \,. \tag{4.2.7}$$

As an aside, because the above expression equals (3.7.12) from section 3.7.2.4, one can derive an explicit formula for $\mathscr{R}_{\mathrm{hNBDM}}(\zeta S \vec{f}, \vec{\mathbf{n}})$. Again, it is not elegant.

4.3   Methods

Section 4.3.1 explains which distributions were simulated, why, and how. Section 4.3.2 explains how logarithms of the likelihood ratios were computed for five of the distributions. I used NumPy [HMvdW+20] version 1.20.2 and SciPy [VGO+20] version 1.6.2 for computations, and Matplotlib [Hun07] version 3.4.1 and/or Seaborn [Was21] version 0.11.1 for plots. Complete implementation details can be found in the code at the relevant GitLab repository. See https://gitlab.com/krinsman/droplets.

4.3.1   Simulated Distributions

To test how much the assumptions of hPoMu may be violated with hPoMu remaining an adequate model, I simulated five distributions. The first distribution, the hierarchical truncated Poisson Multivariate Hypergeometric (hTPMH) accounts for sampling without replacement (section 4.3.1.1), removing the sampling with replacement approximation of hPoMu. The remaining four distributions removed the homogeneity assumption of hPoMu by modelling density and compositional heterogeneities via over-dispersion (sections 4.3.1.2 and 4.3.1.3). Details of the simulations common to all distributions are explained in section 4.3.1.4.

4.3.1.1   Simulating Sampling without Replacement

I used an initial total population size of $500,000,000$ (500 million) cells for defining the parameter K of hTPMH. This is substantially less than the number of cells ($\sim 10^{10}$) which would be sampled from in practice. The NumPy version of the multivariate hypergeometric distribution does not scale beyond a total population size of 10^9.

4.3.1.2   Simulating Compositional Heterogeneity Only

The first two distributions model only compositional heterogeneity and have no density heterogeneity. hPoDM (cf. section 3.6) with compositional concentration $\zeta_C = 100$ has the lowest compositional heterogeneity, followed by hPoDM with $\zeta_C = 1$.

4.3.1.3 Simulating Compositional and Density Heterogeneities

The final two distributions model both compositional and density heterogeneity, with both assumed to correspond to equal over-dispersion. hNBDM (cf. section 3.7) with concentration $\zeta := \zeta_C = \zeta_D = 100$ has the lowest heterogeneities, followed by hNBDM with $\zeta = 1$.

4.3.1.4 Simulation Implementation Details

For all simulations and distributions, the same "simulated community" (or "SimCom") of 91 strains[2] was used. These correspond to 90 strains distributed across 9 distinct relative abundances or population frequencies: 10 strains for each of .01%, .02%, .05%, .1%, .2%, .5%, 1%, 2%, and 5%. The 91st strain was a "remainder" strain with abundance $\approx 12\%$. All distributions also used the same value of the rate parameter, $\lambda = 2$. This equals the expected number of cells for each droplet for all distributions except[3] technically for hTPMH.

Each of the distributions was statistically independently simulated 500 times, with each simulation having 15 million statistically independently simulated droplets, except for hTPMH, for which the statistical dependence structure of successive droplets is inherently Markov. (This corresponds to 5 batches, each with 3 million droplets.)

4.3.2 Log Likelihood Ratios

I derived explicit formulae for the logarithms of the likelihood ratios relative to hPoMu for the hTPMH, hPoDM, and hNBDM families of distributions (cf. section 4.2). I did not attempt to derive analogous formulae for the hExhPoDM and hExhNBDM families of distributions to avoid making mistakes when attempting to compute anti-derivatives of expressions involving the (logarithm of the) Gamma function. (Cf. section 5.3.1.)

Then, for each of the five distributions (hTPMH, hPoDM $\zeta_C = 100$, hNBDM $\zeta = 100$, hPoDM $\zeta_C = 1$, hNBDM $\zeta = 1$) with explicit formulae for the logarithm of their likelihood ratios with respect to hPoMu, of the 500 simulations I arbitrarily chose one (the 18th) and evaluated the respective formula on all 15 million droplets. For each of the five distributions, I also collected the results of doing so into gluttonous groups[4] (which by definition overlap) and computed the geometric mean of the observed likelihood ratios for all groups.

---

[2]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.

[3]The expectation of a truncated Poisson distribution will be less than that of the corresponding Poisson distribution. In this context any deficit compared to $\lambda = 2$ is negligible in practice, because the support was always truncated to a value no smaller than ≈ 400 million.

[4]See section 5.3.4 for clarification of the specific meaning of the term "gluttonous groups". See also footnote 6 of section 5.3.4 regarding the groups defining the diagonals of the heatmaps.

4.4 Results

Section 4.4.1 discusses evidence indicating that sampling with replacement is the most innocuous of the assumptions behind hPoMu. Section 4.4.2 discusses evidence indicating that ghNBDM distributions with low (but nonzero) heterogeneity may be adequately modeled by hPoMu. Finally, section 4.4.3 discusses evidence indicating that ghNBDM distributions with only moderate heterogeneity can be easily distinguished from hPoMu.

4.4.1 Sampling without Replacement

All available evidence suggests that, even after $15,000,000$ droplets have been formed, the hTPMH distribution for the chosen value of K is extremely similar to the corresponding hPoMu distribution.

4.4.1.1 Distribution of Likelihood Ratios w.r.t. hPoMu

When plotted on the same scale as that used for the other distributions, the histogram for hTPMH of observed likelihood ratios relative to hPoMu seen in figure 4.1a has no right tail whatsoever. This indicates few, if any, of the droplets were substantially "better explained" using the hTPMH distribution compared to using hPoMu.

4.4.1.2 Typical Likelihood Ratios as a Function of Strain Frequency

The heat map of typical values for hTPMH of the likelihood ratio across gluttonous groups showin in figure 4.1b also shows, even at worst, negligible differences from hPoMu. The values are slightly smaller for the least abundant strains, reflecting how the sampling with approximation is weakest for those strains. Even so, the values are still very close to 1 and well within the range observed for other distributions.

4.4.2 Very Similar Distributions

All available evidence suggests that the hPoDM with $\zeta_C = 100$ and hNBDM with $\zeta = 100$ distributions are very similar to the corresponding hPoMu distribution. Nevertheless, they also appear to differ slightly more than the hTPMH distribution.

4.4.2.1 Distributions of Likelihood Ratios w.r.t. hPoMu

Unlike what occurs for hTPMH, figures 4.2a and 4.2b respectively do show the distributions of observed likelihood ratios for hPoDM with $\zeta_C = 100$ and hNBDM with $\zeta = 100$ having right tails. However, these right tails are small.
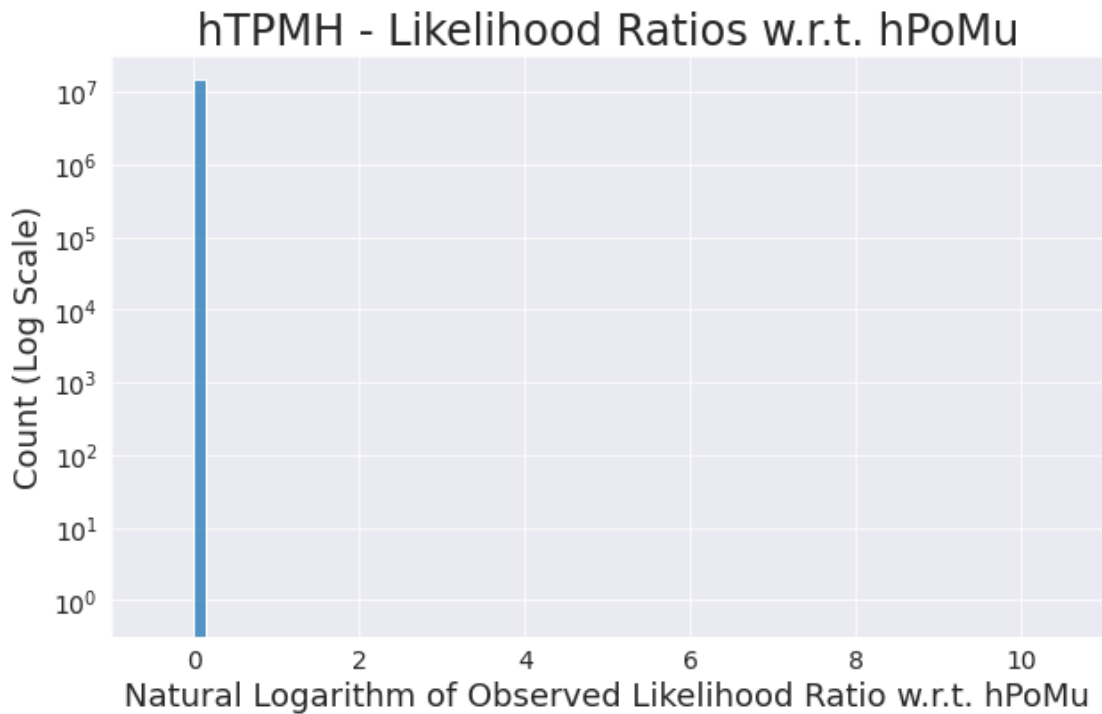
(a) Likelihood ratio histogram for hTPMH.
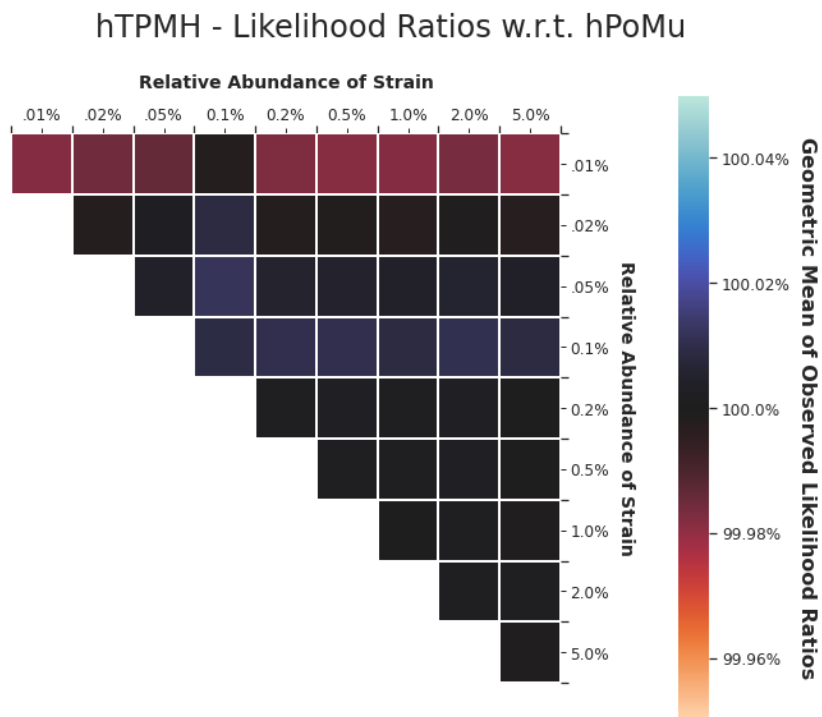
(b) Grouped likelihood ratios for hTPMH.

Figure 4.1: The likelihood ratio with respect to hPoMu never substantially differs from 1.
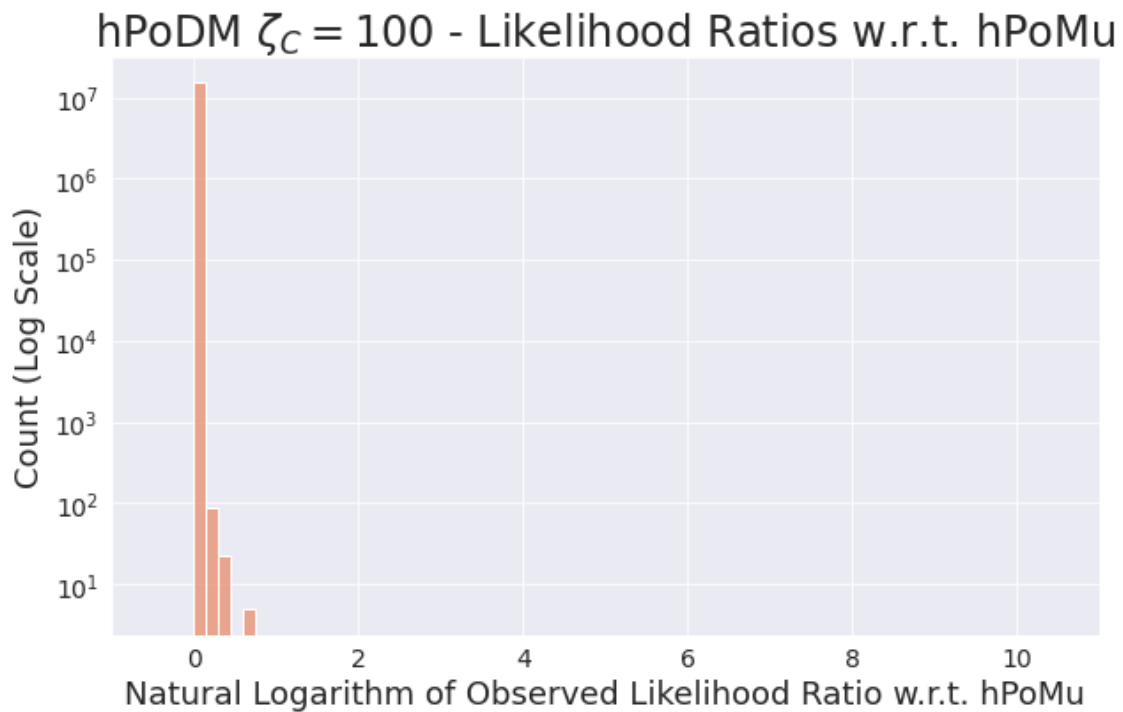
(a) Compare with figure 4.1a.
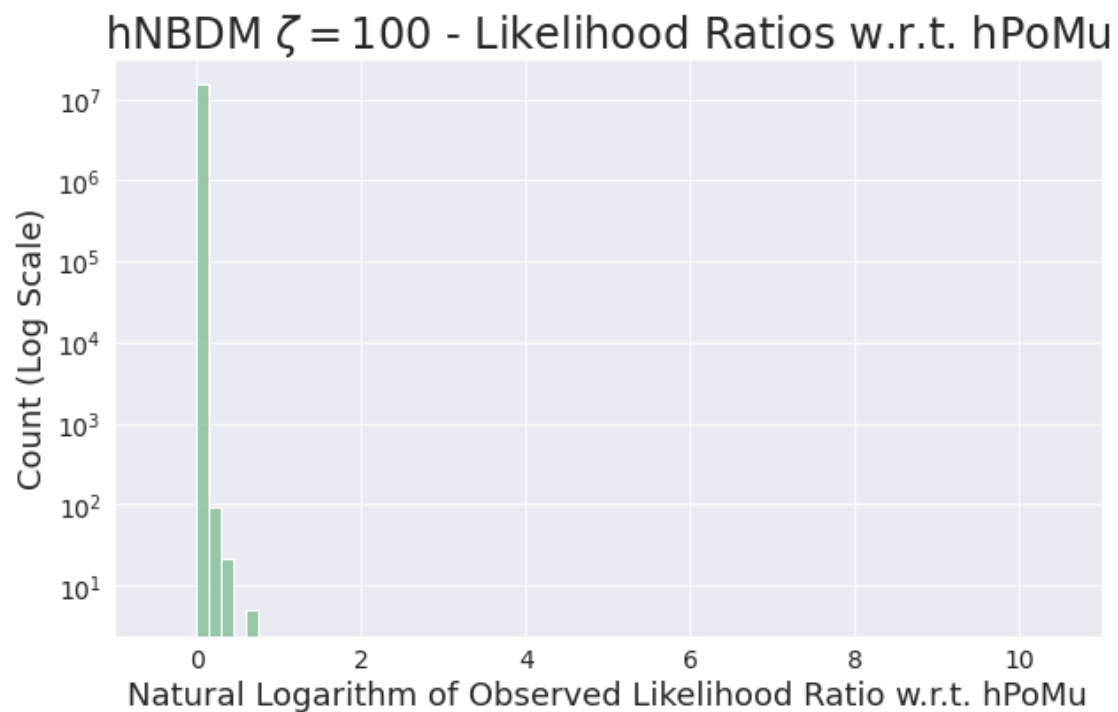
(b) Compare with figure 4.1a.

Figure 4.2: Likelihood ratio histograms for very similar distributions.

4.4.2.2 Typical Likelihood Ratios as a Function of Strain Frequency

Similarly, while figures 4.3a and 4.3b also show greater differences relative to hPoMu compared to those observed for the hTPMH distribution, the differences still trend very small. (That the geometric means across the gluttonous groups tend to be slightly *less* than 1, rather than slightly greater, most likely reflects, in addition to the overlapping nature of the gluttonous groups, the depletion of droplets with 3 or more strains relative to hPoMu that will be observed in these distributions in chapter 5.)

4.4.3 Slightly Different Distributions

All available evidence suggests that the hPoDM with $\zeta_C = 1$ and hNBDM with $\zeta = 1$ distributions are (at least) slightly different from the corresponding hPoMu distribution.

4.4.3.1 Distributions of Likelihood Ratios w.r.t. hPoMu

Figures 4.4a and 4.4b show very large right tails for the distributions of observed likelihood ratios for both the hPoDM ($\zeta_C = 1$) and the hNBDM ($\zeta = 1$). This shows that many droplets were substantially "better explained" using the true simulation distribution rather than hPoMu, something which did not occur for any of the less heterogeneous distributions. It is unclear whether either large right tail is "larger" than the other.

4.4.3.2 Typical Likelihood Ratios as a Function of Strain Frequency

Finally, figures 4.5a and 4.5b show typical values for the likelihood ratios in each gluttonous group for hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$ that consistently "float above" 1, in clear contrast to the "hovering around 1" trend observed in figures 4.1b, 4.3a, and 4.3b that correspond to less heterogeneous distributions. One also sees that the values in figure 4.5b trend slightly higher than those in figure 4.5a, reflecting again how the hNBDM family of distributions incorporates both kind of heterogeneities whereas the hPoDM family does not.

4.5 Discussion

The results show the following. The sampling with replacement assumption of hPoMu is completely harmless for adequately modelling the data. Depending on our target parameters, and which values of heterogeneities are realistic, the homogeneity assumptions of hPoMu could potentially inadequately model the data. Density and compositional heterogeneities are the only failure of the hPoMu assumptions with the potential to be important in practice.

Section 4.5.1 discusses the taxonomy of distributions identified in section 4.4. Section 4.5.2 discusses the two most similar groups, which are either very difficult or impossible to distinguish from hPoMu. Section 4.5.3 discusses the distributions which can be easily distinguished from hPoMu, but whose differences from hPoMu may not be *practically* important.
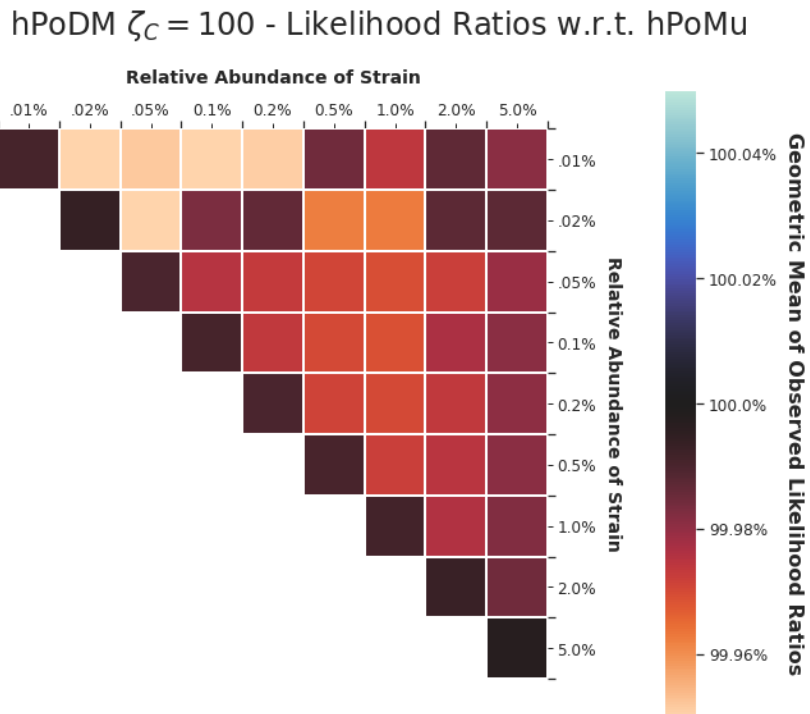
(a) Compare with figure 4.1b.
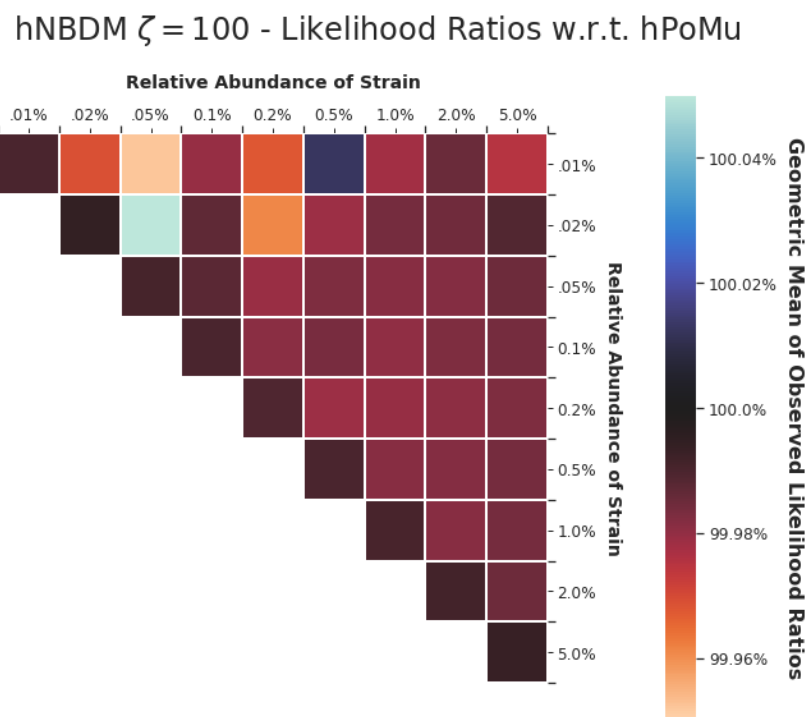
(b) Compare with figure 4.1b.

Figure 4.3: Grouped likelihood ratios for very similar distributions.
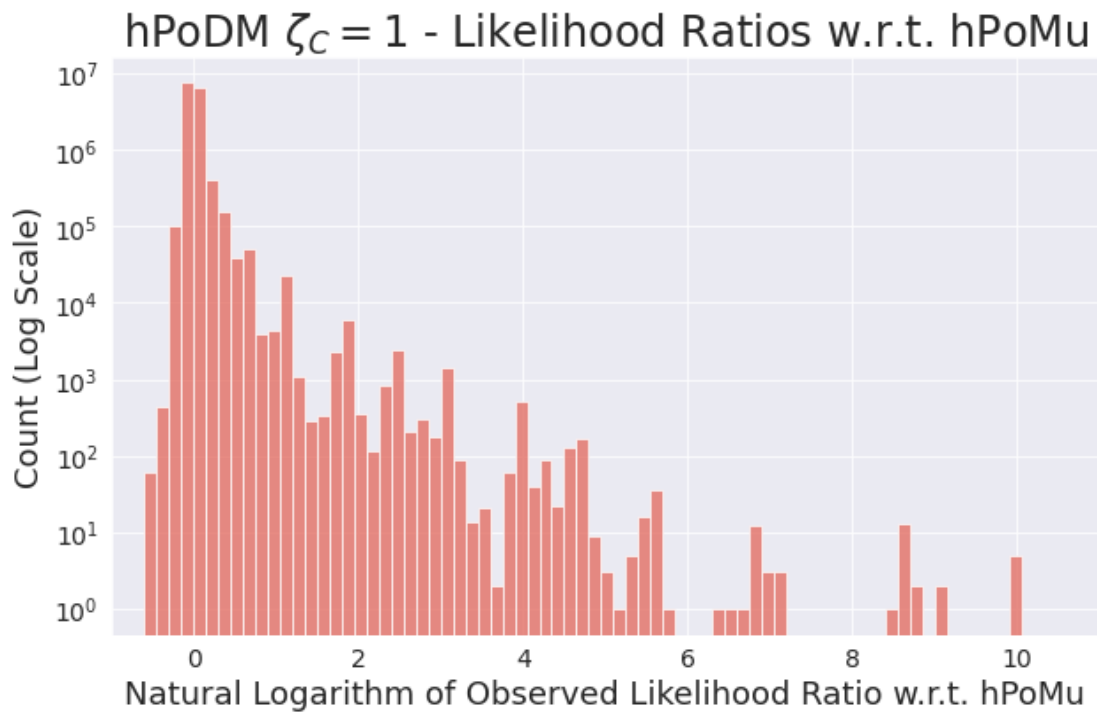
(a) Compare with figure 4.1a as well as with figure 4.2a.
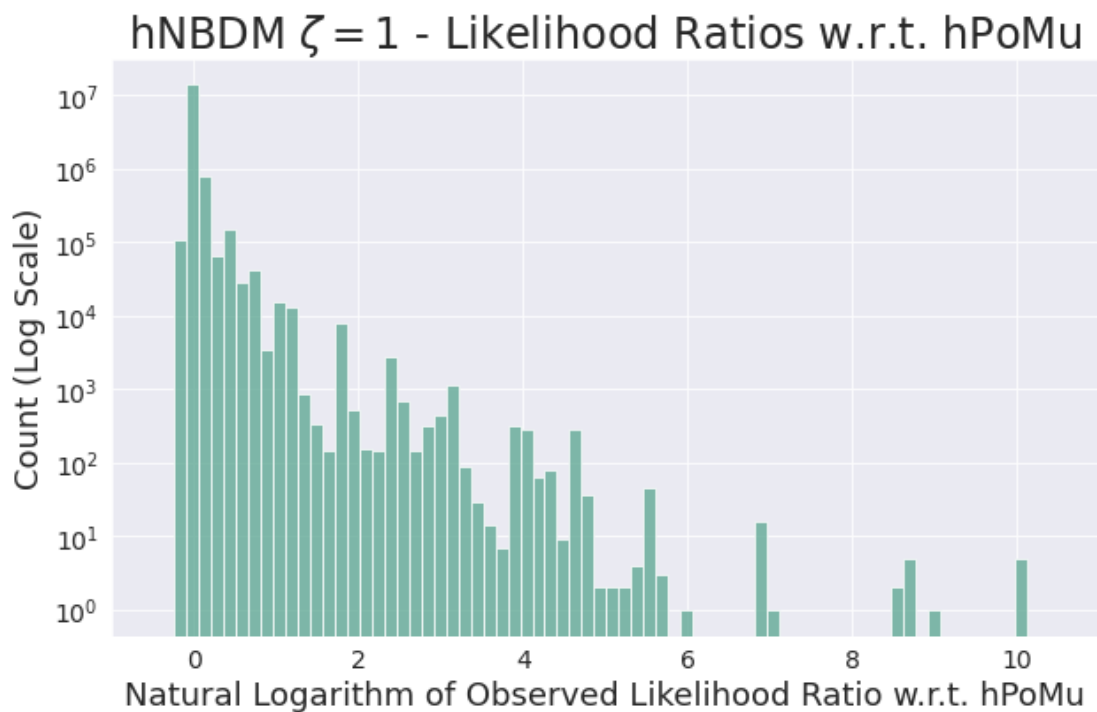
(b) Compare with figure 4.1a as well as with figure 4.2b.

Figure 4.4: Likelihood ratio histograms for slightly similar distributions.
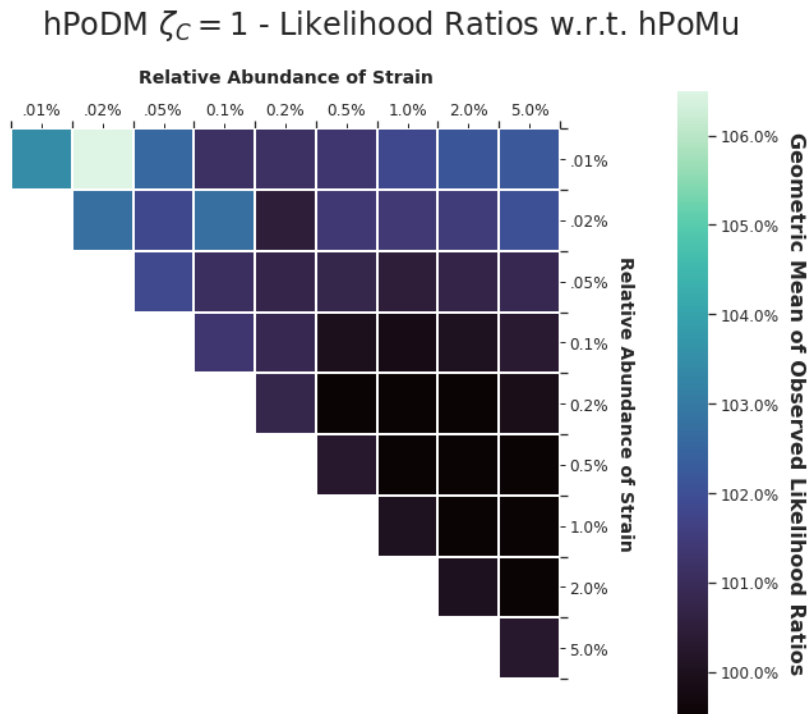
(a) Compare with figure 4.1b as well as with figure 4.3a.
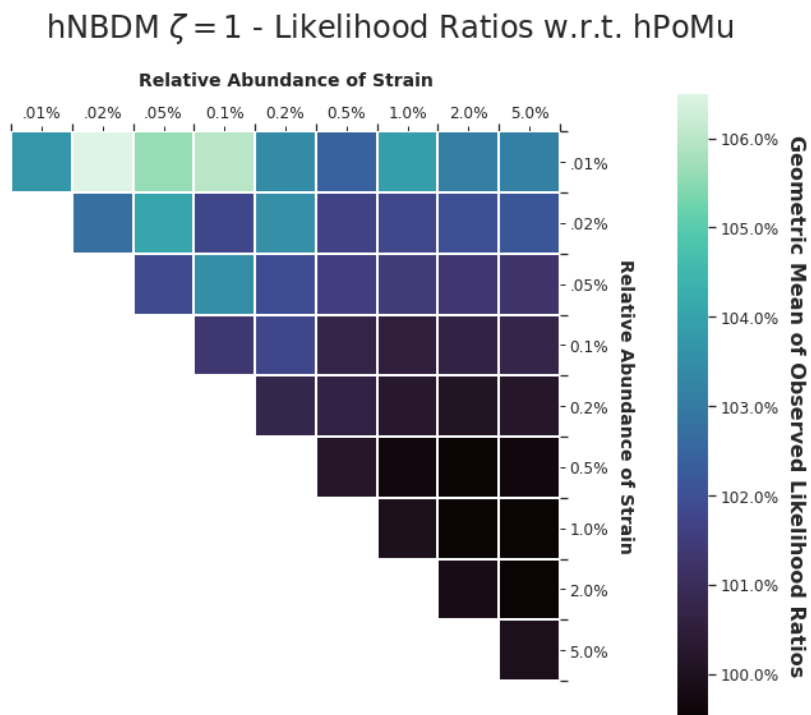
(b) Compare with figure 4.1b as well as with figure 4.3b.

Figure 4.5: Grouped likelihood ratios for slightly similar distributions.

Section 4.6 reviews the next steps to ascertain which differences from hPoMu might affect our practical goal of being able to infer microbial interactions.

4.5.1 Overview of Groups of Distributions

The simulated distributions can be grouped into three levels of assumption violation.

For the first two levels, the first consisting of hTPMH and the second consisting of hPoDM $\zeta_C = 100$ and hNBDM $\zeta = 100$, the data produced is difficult to distinguish from data produced by the hPoMu distribution. Therefore hPoMu is most likely an adequate model and approximation for data generated by distributions from these two levels.

The third level consists of hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, which produce data that can be relatively easily distinguished from data produced by hPoMu.

4.5.2 Very Similar Distributions

While hTPMH, hPoDM $\zeta_C = 100$, and hNBDM $\zeta = 100$ are all very similar to hPoMu, the above results suggest that hTPMH is even more similar to hPoMu than the other two.

Moreover, the simulations actually *overstate* how different data generated by hTPMH should be in practice from data generated by hPoMu. I showed in section 3.5.2.3 that hTPMH converges in distribution to hPoMu as the size of the total population of cells being sampled from increases. While the simulation assumed a population size of 500 **m**illion, or 5×10^8, cells, a real MOREI experiment is likely to sample from a population containing upwards of 10 **b**illion, or 1×10^{10}, cells, more than twenty times as many.

Thus, to the extent we consider that data generated by hPoDM $\zeta_C = 100$ and hNBDM $\zeta = 100$ are difficult to distinguish from data generated by hPoMu, data that would be generated by hTPMH in practice should be considered effectively indistinguishable from data generated by hPoMu. In particular, even for extremely low abundance strains, these results show that the sampling with replacement approximation is completely irrelevant in practice to accurate modelling of the data generated by MOREI.

4.5.3 Slightly Different Distributions

The results show that there are evident differences between data generated by distributions in the third level, hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, and data generated by hPoMu. Yet this still is not enough to be certain that there are *practically* significant differences. (Setting aside how to be more precise about what is a "practically significant" difference.)

For example, the long right tails from figures 4.4a and 4.4b are only easily visible when plotting the number of counts on a log scale. The total fraction of droplets represented by those long right tails is actually relatively small. It's unclear whether the existence of a small faction of droplets which are *much* better explained by the true models than by hPoMu is an important demerit against hPoMu when most droplets are nearly equally well explained

by both. Focusing on "typical" values of the likelihood ratios as depicted in figures 4.5a and 4.5b, the advantage of the true models over hPoMu is usually *at most* around 6%.

On the other hand, the "typical" likelihood ratio values in figures figures 4.5a and 4.5b also show a distinct upwards trend as the abundance of one of the involved strains decreases. As explained in section 1.5, quantifying the extent of MOREI's ability to advance the state of the art requires focusing on the *least* abundant strains. Therefore, inasmuch as the discrepancies of these models compared to hPoMu may be particularly important for the least abundant strains, the discrepancies may also be important overall.

Thus the results in this chapter seem to point both ways about whether hPoMu is an adequate working model for data generated by these distributions.

4.6 Conclusion

Findings and Contributions I demonstrated that log likelihood ratios can be used to determine whether the failure of assumptions from section 3.3.3 causes substantially new behavior. Failure of the sampling without replacement assumption turns out to have negligible effects in practice, but failures of the other assumptions could be important.

Practical Implications The results from this chapter clarify that we can safely ignore concerns regarding the sampling without replacement assumption of hPoMu when choosing a (working) model for the initial formation of droplets. The practical implications of the results from this chapter are otherwise limited to the extent that they do not unambiguously indicate whether or not we can ignore concerns regarding the homogeneity assumptions of hPoMu. These results are useful, however, for at least raising the possibility that violations of the homogeneity assumptions could have important effects on our ability to predict the numbers of droplets available for making inferences about given microbial interactions.

Next Steps and Open Questions Although these results do suggest that data generated by certain distributions is unlikely to be adequately modeled by hPoMu, we need to be more precise to substantiate such a claim. We need to decide precisely which aspects of the data we want to be model accurately, to target, or otherwise we have no way to measure whether hPoMu is an adequate model for the data. Only experiments measuring those targeted aspects of the data will give results which could definitively determine whether or when hPoMu is an adequate model. This is chapter 5.

Moreover, even if hPoMu is inaccurate for those targeted aspects of the data, it could still be an adequate working model as long as its predictions are conservative compared to the truth, rather than overly optimistic. To make that assessment, we need to know what values of heterogeneity might be realistic in practice. That in turn requires having estimators for the heterogeneities to use on empirical data. We fill that gap later in chapter 6.

Chapter 5

Goodness of Fit Using Contingency Tables of the Targeted Estimands

Effects of Heterogeneities on Data Throughput

Herein I confirm that more severe failures of the assumptions from section 3.3.3 cause more severe discrepancies with the predictions derived from hPoMu. See for example sections 5.4.1.3, 5.4.1.4, 5.4.2.1. The nature of the effect depends on the chosen grouping (picky or gluttonous) of droplets defining the targeted estimands. See sections 5.4.2.3 and 5.5.3. Effects of the failure of the sampling with replacement assumption are further confirmed to be negligible in practice. See for example sections 5.4.1.1 or 5.4.2.1.

Section 5.2 explains some of the technical details of the analyses. Section 5.3 explains implementation details of the analyses. Section 5.4 explains the results of the analyses. Section 5.5 explains how the results are relevant to modelling the initial formation of droplets.

Chapter 5 Contents

5.1   Background and Significance

When planning the previous chapter 4, I had anticipated the effect of the homogeneity assumptions to be negligible. This was true of the effect of the sampling with replacement assumption, which was investigated earlier. Similarity of underlying distributions $\mathcal{P}_1$ and $\mathcal{P}_2$ should be enough to imply similarity of target parameters $\boldsymbol{\Psi}(\mathcal{P}_1)$ and $\boldsymbol{\Psi}(\mathcal{P}_2)$. The log-likelihood ratio statistics provide evidence of similarity of the underlying distributions $\mathcal{P}_1$ and $\mathcal{P}_2$. Hence, if the anticipated findings of the previous chapter 4 had been valid, then the log-likelihood ratio statistics alone would have been sufficient to demonstrate them.

The questions here are the same as those of the previous chapter 4. In practice we only care about the *relevant* portion $\mathcal{Q}_O$ of the probability distribution of the observed data that affects our target parameters $\boldsymbol{\Psi}(\mathcal{P}_O)$. (Cf. the overview of these ideas in [vdLR11, section 1.5].) Again, in the previous chapter 4 we found evidence that violations of the homogeneity assumptions could cause the *entire* underlying probability distribution of the observed data $\mathcal{P}_O$ to substantially differ from that under hPoMu. Yet even if two probability distributions $\mathcal{P}_1$ and $\mathcal{P}_2$ substantially differ, that does not necessarily mean that the *relevant* portions $\mathcal{Q}_1$ and $\mathcal{Q}_2$ also substantially differ. Hence, as mentioned before in section 4.6, the evidence from chapter 4 is insufficient to say whether violations of the homogeneity assumptions lead to practically significant differences, ones that affect $\mathcal{Q}_O$ / $\boldsymbol{\Psi}(\mathcal{P}_O)$. Thus the goal for this chapter is to look for and implement goodness of fit procedures that speak more directly to comparisons of target parameters $\boldsymbol{\Psi}(\mathcal{P})$ and relevant portions $\mathcal{Q}$ of probability distributions.

Broader field   As was the case for the previous chapter 4, the problem of this chapter also belongs to the general field of model selection, and to the subfield of goodness of fit tests in particular. Again, the more accessible [HTF09, chapter 7] or the highly sophisticated [BBM99] are recommended as references for the general theory of model selection. Cf. the discussion earlier from the introduction to Part II or from section 4.1. While the previous chapter 4 discusses goodness of fit tests using the likelihood ratio statistic, this chapter focuses on goodness of fit tests for contingency tables that use the χ^2-divergence [CZPA09, p. 57] statistic. (Cf. section 5.2.2 for terminology.) See the first two chapters of [HCBNM02], or the introduction of [Joh04], for reviews of the classical procedure for using the χ^2-divergence statistic ("Pearson" χ^2 statistic) in goodness of fit tests for contingency tables. The article [WSCR03] discusses "non-classical" use of the χ^2-divergence statistic for goodness of fit via Monte Carlo approximations of null distributions, and applications thereof to ecology. The review [Fie11] discusses the use of classical χ^2 procedure as well as other, more sophisticated methods for analyzing contingency tables (for goodness of fit and for other reasons).

Specific problem   Herein "practically significant" differences correspond to the target parameters defined in section 2.3.4, the "data throughput" of the experiment. These target parameters allow us to predict how much data (measured in numbers of droplets) will be produced that can be used to infer given microbial interactions, cf. again section 2.3.2. Thus,

in this chapter I investigate how failures of the implicit hPoMu assumptions, in particular failures of the homogeneity assumptions, affect these target parameters.

As long as the average number of cells per droplet is not too large, the number of droplets available to serve as controls will usually be much larger than the number of droplets available to serve as treatments (i.e. droplets where strains co-occur). Cf. the heuristic argument in section 2.C. Hence, if an estimation method is ever too starved of data to make inferences about a given microbial interaction, most likely the cause will be too few treatment droplets. In terms of "data throughput", bottlenecks that occur will usually be caused by insufficiently many treatment droplets. Therefore the *main* goal of this chapter is to understand effects on those target parameters corresponding specifically to the treatment droplets.

These target parameters, and estimators thereof, can be grouped meaningfully into contingency tables. (Cf. sections 5.2.3 and 5.2.4.) Therefore we can use goodness of fit procedures designed for contingency tables to accomplish our goal, to find and implement goodness of fit procedures that speak more directly to comparisons of these target parameters.

The paper [WSCR03] discusses goodness of fit procedures for ecological simulations. This includes ecological models that produce count data that can be grouped into a contingency table, but the authors also show how the same methodology can be applied more generally. The approach is particularly useful because it can be applied in situations that violate the assumptions of the "classical" Pearson goodness of fit procedure for contingency tables.

The paper [Zho18] studies mixed-membership models involving distributions similar to those in this work. The analysis of mixed-membership models is known to be related to the analysis of contingency tables, see for example the review article [Fie11]. The key idea relevant here appears to be that multivariate count data is closely related to multivariate categorical data. For finite samples, observed multivariate count data can always be converted into multivariate categorical data via truncation. In fact, the procedures in e.g. section 5.2.3.1 can be thought of as "collapsing" or "marginalizing" a sparse "multi-dimensional" contingency table with 2^S cells ("exclusive and nonempty subsets" [Zho18]) into a "one-dimensional" contingency table whose number of cells is at most $O(S^2)$. So although the connection with the problems and ideas in [Zho18] is certainly currently somewhat hazy, it also seems reasonable to expect that the connection could be clarified in future work.

Particular approach Under the assumptions of the classical goodness of fit test for contingency tables developed by Pearson [HCBNM02], the asymptotic distribution of the χ^2-divergence [CZPA09, p. 57] statistic is the Pearson χ^2 distribution [WSCR03, p. 53]. Because this is only an asymptotic result, for finite samples the sampling distribution of the χ^2-divergence statistic may actually substantially differ from the χ^2-distribution, even under the null hypothesis. Cf. section 5.3.2. Therefore I computed a Monte Carlo approximation of the sampling distribution of the χ^2-divergence statistic and used that to compute p-values. Cf. [WSCR03, section 3] for a fairly detailed explanation of Monte Carlo hypothesis testing.

Details of the classical approach, which assumes that the asymptotic distribution is a valid approximation, again can be found in [HCBNM02]. That assumption seemed unrealistic here,

cf. again section 5.3.2. Bayesian versions of this procedure also exist, cf. [Joh04]. However such an approach seemed unnecessarily complicated for exploratory work like this.

5.2 Preliminaries

In this section I clarify details needed to analyze the data throughput predicted by each model of initial droplet formation. Section 5.2.1 gives precise definitions of the statistics that are relevant for studying data throughput (following the argument given in section 5.C). Section 5.2.2 defines the divergence used to compare distributions. Section 5.2.3 clarifies the details of the hypothesis test comparing distributions to hPoMu using this divergence. In addition to the "picky" and "global" way of grouping of droplets by strain[1] for hypothesis testing defined already in section 5.2.3, section 5.2.4 defines a "gluttonous", "local", or "pairwise" way of grouping droplets by strain for hypothesis testing.

5.2.1 Definition of Throughput

Previously we have been looking at distributions of numbers of cells in a given droplet (of any strain or of a given strain). As argued in section 5.C, I believe that understanding the distributions of numbers of *strains* in a given droplet is the most relevant for predicting data throughput. In section 5.2.1.1 I clarify how to define these as statistics computed from the numbers of cells. Then sections 5.2.1.2 and 5.2.1.3 give two different possible definitions of throughput. This section is a more detailed version of section 2.3.4.

5.2.1.1 Strain Count Distribution

The notation $\mathbb{I}\{A\}$ denotes the indicator function for the event A.

For a given strain s, the **strain presence** RV (random variable):

$$\mathbb{S}^{(s)}(0) := \mathbb{I}\{N^{(s)}(0) \geq 1\} = \mathbb{I}\{N^{(s)}(0) > 0\} . \tag{5.2.1}$$

indicates whether any cells of strain s belong to the droplet. It takes values in $\{0, 1\}$ like any indicator random variable. Specifically, it equals 1 if strain s is present in the droplet and 0 if strain s is absent.

Similarly, for a given strain s, the **strain absence** RV:

$$\overline{\mathbb{S}}^{(s)}(0) := 1 - \mathbb{S}^{(s)}(0) = \mathbb{I}\{N^{(s)}(0) = 0\} , \tag{5.2.2}$$

takes values in $\{0, 1\}$ like any indicator random variable, equalling 1 if strain s is absent in the droplet and 0 if strain s is present.

---

[1]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.

For a given droplet, the joint distribution of the strain presence RVs for all strains corresponds to a binary random vector:

$$\vec{\mathbb{S}}(0) := (\mathbb{S}^{(1)}(0), \ldots, \mathbb{S}^{(S)}(0)) . \tag{5.2.3}$$

Its values are in $\bigtimes_{s=1}^{S}\{0, 1\}$ (the S-fold Cartesian product of $\{0, 1\}$ with itself).

For a given droplet, its **strain count** $\mathbb{S}(0)$ is the number of strains present in the droplet. The strain count equals the sum of the strain presence RVs:

$$\mathbb{S}(0) := \sum_{s \in [S]} \mathbb{S}^{(s)}(0) = \sum_{s \in [S]} \mathbb{I}\{N^{(s)}(0) \geq 1\} . \tag{5.2.4}$$

It is an RV taking values in $\{0\} \cup [S]$.

Another useful notion is the **support** of a (non-negative) vector $\vec{v} \in \mathbb{R}_{\geq 0}^{S}$:

$$\operatorname{supp}(\vec{v}) := \{s \in [S] : [\vec{v}]_s > 0\} \subseteq [S] . \tag{5.2.5}$$

It follows directly from the definitions that $\operatorname{supp}(\vec{N}(0)) = \operatorname{supp}(\vec{\mathbb{S}}(0))$ always, corresponding to the set of $s \in [S]$ such that $\mathbb{S}^{(s)}(0) = 1$. Moreover we also always have[2] that $\mathbb{S}(0) = |\operatorname{supp}(\vec{N}(0))| = |\operatorname{supp}(\vec{\mathbb{S}}(0))|$.

As argued in section 5.C, I believe that the notion of "throughput" that is important for this problem corresponds to $\vec{\mathbb{S}}(0)$ and not (directly[3]) to $\vec{N}(0)$. While choosing $\lambda \approx 2$ is intended to make $\vec{N}(0)$ as similar to $\vec{\mathbb{S}}(0)$ as possible, they are definitely not the same, something which e.g. the multiple representatives problem from section 5.C makes clear.

Even when the first moments of $\vec{N}(0)$ remain the same, changes to the second moments of $\vec{N}(0)$ can still cause changes in the expected numbers of droplets serving as treatments or controls. A heuristic way to understand this is to observe that the expected data throughput corresponds to the second (or higher order) moments of $\vec{\mathbb{S}}(0)$; see sections 5.2.1.2 and 5.2.1.3. Although for any distribution belonging to the ghNBDM working model (see section 3.8.1) the expected cell counts for each strain are the same as those under hPoMu, the choice of distribution from the ghNBDM working model still affects the predicted data throughput.

5.2.1.2 Gluttonous Definition of Data Throughput

Recall from section 1.3, or the introduction to Part I, that the treatment group for inferring the effect of strain s_1 on the growth of strain s_2 corresponds to droplets where s_1 and s_2 co-occur. The set of all droplets satisfying this condition is

$$\begin{aligned}
\mathcal{T}_g^{(s_1, s_2)} &:= \left\{d \in [D] : \mathbb{S}_d^{(s_1)}(0) \cdot \mathbb{S}_d^{(s_2)}(0) = 1\right\} \\
&= \left\{d \in [D] : \mathbb{I}\left\{N_d^{(s_1)}(0) \geq 1 \ \& \ N_d^{(s_2)}(0) \geq 1\right\} = 1\right\} \tag{5.2.6} \\
&= \left\{d \in [D] : \operatorname{supp}(\vec{\mathbb{S}}_d(0)) \supseteq \{s_1, s_2\}\right\} .
\end{aligned}$$

---

[2] The ***cardinality*** $|\mathcal{S}|$ of a finite set $\mathcal{S}$ is the number of elements of the set. E.g. $|[S]| = S$.

[3] Only indirectly as mediated via $\vec{\mathbb{S}}(0)$, which of course can be computed from $\vec{N}(0)$.
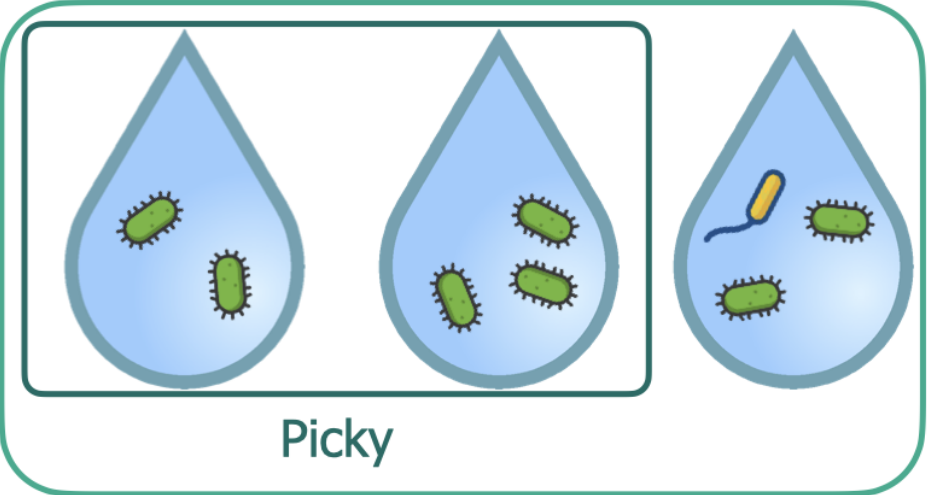
Figure 5.1: Gluttonous vs. Picky Definitions of Data Throughput. Cf. text.

This is called the "**gluttonous**" **treatment** group because *all* droplets where s_1 and s_2 co-occur are included. Cf. figure 5.1.

Recall also from section 1.3, or the introduction to Part I, that the control group for inferring the effect of strain s_1 on the growth of strain s_2 corresponds to droplets where s_2 occurs but s_1 does not. The set of all droplets satisfying this condition is

$$
\begin{aligned}
\mathcal{C}_g^{(s_1,s_2)} &:= \left\{ d \in [D] : \overline{\mathbb{S}}_d^{(s_1)}(0) \cdot \mathbb{S}_d^{(s_2)}(0) = 1 \right\} \\
&= \left\{ d \in [D] : \mathbb{I}\left\{ N_d^{(s_1)}(0) = 0 \ \& \ N_d^{(s_2)}(0) \geq 1 \right\} = 1 \right\} \\
&= \left\{ d \in [D] : \mathrm{supp}(\vec{\mathbb{S}}_d(0)) \supseteq \{s_2\}, \ s_1 \notin \mathrm{supp}(\vec{\mathbb{S}}_d(0)) \right\} .
\end{aligned}
\tag{5.2.7}
$$

This is called the "**gluttonous**" **control** group because *all* droplets where s_2 occurs but s_1 does not are included. Cf. figure 5.1.

These definitions place no constraints on the presence or absence of strains $s \in [S]$ besides s_1 and s_2. This can be favorable e.g. when studying very rare strains, for which there may be no or very few droplets containing s_1 and s_2 only, but some containing s_1 and s_2 along with other strains. These definitions give the best chance of avoiding "data starvation".

From equation (5.2.6) we get (assuming the droplets are identically distributed) that the expected number of droplets for the gluttonous treatment group for inferring the effect of strain s_2 on the growth of s_1 is

$$
\mathbb{E}\left[\left|\mathcal{T}_g^{(s_1,s_2)}\right|\right] = D \cdot \mathbb{E}\left[\mathbb{S}^{(s_1)}(0) \ \mathbb{S}^{(s_2)}(0)\right] .
\tag{5.2.8}
$$

Similarly, from equation (5.2.7) we get (assuming the droplets are identically distributed) that the expected number of droplets for the corresponding gluttonous control group for inferring the effect of strain s_2 on the growth of s_1 is

$$
\mathbb{E}\left[\left|\mathcal{C}_g^{(s_1,s_2)}\right|\right] = D \cdot \mathbb{E}\left[\overline{\mathbb{S}}^{(s_1)}(0) \cdot \mathbb{S}^{(s_2)}(0)\right] = D \cdot \mathbb{E}\left[\mathbb{S}^{(s_2)}(0) - \mathbb{S}^{(s_1)}(0) \cdot \mathbb{S}^{(s_2)}(0)\right] .
\tag{5.2.9}
$$

As claimed in section 5.2.1.1, these expressions for the expected gluttonous data throughput involve the second moments of $\vec{\mathbb{S}}(0)$.

5.2.1.3 Picky Definition of Data Throughput

Comparing figures 5.1 and 1.2, we see that the above "gluttonous" definitions correspond to more possible combinations of strains than we probably would have included when studying the effect of strain s_1 on the growth of strain s_2 by manually plating cells. This motivates the following definitions.

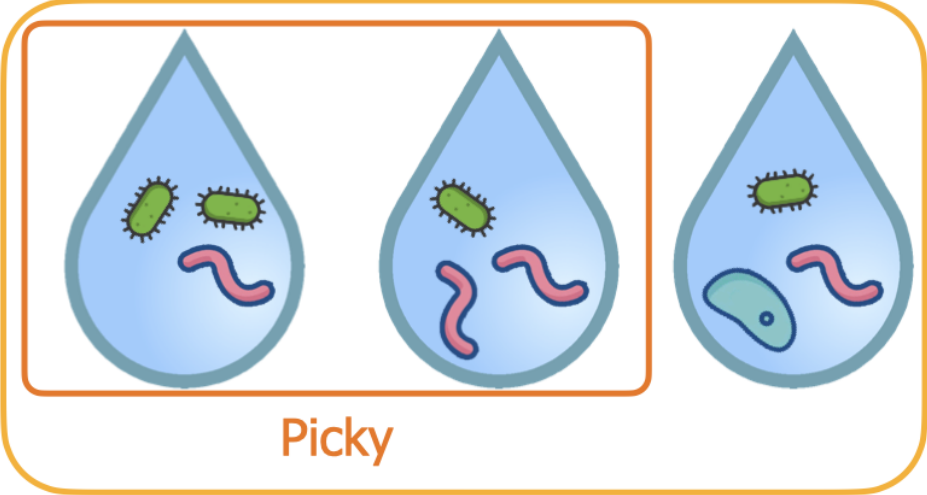The set of droplets where s_1 and s_2 co-occur in a way corresponding to the combination of strains we would use if we were manually plating cells is

$$
\begin{aligned}
\mathcal{T}_p^{(s_1,s_2)} &:= \left\{ d \in [D] : \mathbb{S}_d^{(s_1)}(0) \cdot \mathbb{S}_d^{(s_2)}(0) \prod_{\sigma \notin \{s_1,s_2\}} \overline{\mathbb{S}}_d^{(\sigma)}(0) = 1 \right\} \\
&= \left\{ d \in [D] : \mathbb{I}\left\{ N_d^{(s_1)}(0) \geq 1 \ \& \ N_d^{(s_2)}(0) \geq 1 \ \& \ \forall \sigma \notin \{s_1, s_2\}, \ N_d^{(\sigma)}(0) = 0 \right\} = 1 \right\} \\
&= \left\{ d \in [D] : \mathrm{supp}(\vec{\mathbb{S}}_d(0)) = \{s_1, s_2\} \right\} .
\end{aligned}
$$
(5.2.10)

This is called the "**picky**" **treatment** group because droplets where s_1 and s_2 co-occur are included *only when all other strains are absent.* Cf. figure 5.1.

The set of droplets where *only s_2* occurs, corresponding to the combination of strains we would use if we were manually plating cells, is

$$
\begin{aligned}
\mathcal{C}_p^{(s_1,s_2)} &:= \left\{ d \in [D] : \overline{\mathbb{S}}_d^{(s_1)}(0) \cdot \mathbb{S}_d^{(s_2)}(0) \prod_{\sigma \notin \{s_1,s_2\}} \overline{\mathbb{S}}_d^{(\sigma)}(0) = 1 \right\} \\
&= \left\{ d \in [D] : \mathbb{I}\left\{ N_d^{(s_1)}(0) = 0 \ \& \ N_d^{(s_2)}(0) \geq 1 \ \& \ \forall \sigma \notin \{s_1, s_2\}, \ N_d^{(\sigma)}(0) = 0 \right\} = 1 \right\} \\
&= \left\{ d \in [D] : \mathrm{supp}(\vec{\mathbb{S}}_d(0)) = \{s_2\} \right\} .
\end{aligned}
$$
(5.2.11)

This is called the "**picky**" **control** group because droplets where s_2 occurs but s_1 does not occur belong *only when all other strains are absent.* It is not enough for s_1 alone to be absent. Cf. figure 5.1.

Unlike the gluttonous groups, the picky groups *do* place constraints on the presence or absence of strains $s \in [S]$ besides s_1 and s_2. For both the treatments and the controls, for "picky" groups all strains besides s_1 or s_2 must have zero counts. Every picky group is by definition a subset of its gluttonous counterpart.

The picky definitions can be favorable when we have plenty of droplets from which to make estimates. They reduce the possibility of confounding effects on the growth of strain s_2 that could be caused by strains that are not strain s_1. On the other hand, in instances where there are only very few or no droplets without "extra" strains, the picky definitions could lead to "data starvation".

From equation (5.2.10) we get (assuming the droplets are identically distributed) that the expected number of droplets for the picky treatment group for inferring the effect of strain s_2 on the growth of s_1 is

$$
\mathbb{E}\left[\left| \mathcal{T}_p^{(s_1,s_2)} \right| \right] = D \cdot \mathbb{E}\left[\mathbb{S}^{(s_1)}(0) \cdot \mathbb{S}^{(s_2)}(0) \prod_{\sigma \notin \{s_1,s_2\}} \left(1 - \mathbb{S}^{(\sigma)}(0) \right) \right] .
$$
(5.2.12)

Similarly, from equation (5.2.11) we get (assuming the droplets are identically distributed) that the expected number of droplets for the corresponding picky control group for inferring the effect of strain s_2 on the growth of s_1 is

$$\mathbb{E}\left[\left|\mathcal{C}_p^{(s_1,s_2)}\right|\right] = D \cdot \mathbb{E}\left[\left(1 - \mathbb{S}^{(s_1)}(0)\right) \cdot \mathbb{S}^{(s_2)}(0) \prod_{\sigma \notin \{s_1,s_2\}} \left(1 - \mathbb{S}^{(\sigma)}(0)\right)\right]. \tag{5.2.13}$$

As claimed in section 5.2.1.1, these expressions for the expected picky data throughput involve the higher order moments of $\vec{\mathbb{S}}(0)$.

5.2.2   Pearson Categorical Divergence

The Pearson categorical divergence (more often referred to as the "Pearson χ^2 divergence" or "Pearson χ^2 statistic") of $\vec{\mathbf{v}} = (v^{(1)}, \ldots, v^{(I)}) \in \mathbb{R}_{\geq 0}$ relative to $\vec{\mathbf{w}} = (w^{(1)}, \ldots, w^{(I)}) \in \mathbb{R}_{\geq 0}$ is defined[CZPA09, p. 57] as

$$\sum_{i=1}^{I} \frac{(v^{(i)} - w^{(i)})^2}{w^{(i)}} , \tag{5.2.14}$$

where both $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ belong to the unit simplex (entries sum to 1 and are all ≥ 0), and $\vec{\mathbf{w}}$ belongs to the interior of the unit simplex (entries are all > 0). This definition is not symmetric. In general the divergence of $\vec{\mathbf{w}}$ relative to $\vec{\mathbf{v}}$ differs from that of $\vec{\mathbf{v}}$ relative to $\vec{\mathbf{w}}$.

Herein I call this the "Pearson categorical divergence" to avoid confusion with either

(a) the "Pearson divergence", $1 - \rho_{\vec{x}\vec{y}}$, where $\rho_{\vec{x}\vec{y}}$ is the Pearson correlation of $\vec{x}$ and $\vec{y}$, or

(b) the "Pearson χ^2 distribution"[4], the asymptotic sampling distribution of this statistic under the null distribution of "Pearson-style" goodness of fit tests.

In particular, it would sound confusing to say (in section 5.3.2) that the sampling distribution of the "χ^2 divergences" is not a "χ^2 distribution".

5.2.3   Global Picky Goodness of Fit Test

The strains present in a droplet determine the microbial interactions it can help characterize. Therefore we want to group the droplets according to the strains that are present in each droplet. Different models for the initial formation of droplets make different predictions about which such groups of droplets will be most common. A hypothesis test whose null distribution is a sampling distribution derived from the hPoMu working model compares the predictions made by other (working) models with those made by hPoMu. Rejecting the null hypothesis corresponds to asserting that hPoMu is unable to describe the observed data, so the true model that generated the data has more "descriptive power", cf. section 4.1.

---

[4]Technically a parameterized family of distributions and not a single distribution.

The definition of the hypothesis test is split into several parts. Section 5.2.3.1 defines how the droplets are grouped according to strains for the hypothesis test. Section 5.2.3.2 gives probabilities for each group under the null hypothesis (hPoMu), while section 5.2.3.3 clarifies the expected counts. Section 5.2.3.4 defines the observed counts used in defining the goodness of fit test. Section 5.2.3.5 explains why some groups are redundant. Finally, section 5.2.3.6 uses the conclusions of the previous sections to define the test statistic.

5.2.3.1   Categories

There are technically exponentially many, 2^S, potential combinations of strains that could occur in a given droplet. All of these could be used in defining mutually exclusive categories for a goodness of fit test. However, for both reasons of computational feasibility and data availability[5], herein only combinations with two or fewer strains are given distinct categories, with all combinations of three or more strains lumped into a category.

Thus the number of categories for the "global" goodness of fit test is

$$1 + S + \binom{S}{2} + 1 \,. \tag{5.2.15}$$

This is one category for all droplets with zero strains (empty droplets), S categories for each of the picky control groups (droplets with exactly one strain), $\binom{S}{2}$ categories for each of the picky treatment groups (droplets with exactly two strains), and one category for all droplets with three or more ("multiple") strains. The groups being "picky" and not "gluttonous" ensures that categories do not overlap. This makes it feasible to define a single "global" test.

Cf. section 5.2.1 for more details about "picky" and "gluttonous" groups.

---

[5] Most combinations of three or more strains are highly unlikely and will be represented by few or no data points. Cf. the related discussion in sections 1.2, 2.3.5, and 7.7.3.

5.2.3.2 Probabilities

The probabilities associated with each category are

$$
\begin{aligned}
p_{\mathcal{G}}^{(\emptyset)} &:= \mathbb{P}(N(0) = 0) = e^{-\lambda}, \\
p_{\mathcal{G}}^{(s)} &:= \mathbb{P}(N^{(s)}(0) \geq 1, \forall \sigma \neq s : N^{(\sigma)}(0) = 0) \\
&= \mathbb{P}(N^{(s)}(0) \geq 1) \cdot \prod_{\sigma \neq s} \mathbb{P}(N^{(\sigma)}(0) = 0) \\
&= (1 - e^{-f^{(s)}\lambda}) \cdot e^{-(1 - f^{(s)})\lambda}, \\
p_{\mathcal{G}}^{(s_1, s_2)} &:= \mathbb{P}(N^{(s_1)}(0) \geq 1, N^{(s_2)}(0) \geq 1, \forall \sigma \notin \{s_1, s_2\} : N^{(\sigma)}(0) = 0) \\
&= \mathbb{P}(N^{(s_1)}(0) \geq 1) \cdot \mathbb{P}(N^{(s_2)}(0) \geq 1) \cdot \prod_{\sigma \notin \{s_1, s_2\}} \mathbb{P}(N^{(\sigma)}(0) = 0) \\
&= (1 - e^{-f^{(s_1)}\lambda})(1 - e^{-f^{(s_2)}\lambda}) \cdot e^{-(1 - (f^{(s_1)} + f^{(s_2)}))\lambda}, \\
p_{\mathcal{G}}^{(3+)} &:= 1 - p_{\mathcal{G}}^{(\emptyset)} - \sum_{\sigma=1}^{S} p_{\mathcal{G}}^{(\sigma)} - \sum_{\sigma_1=1}^{S} \sum_{\sigma_2 > \sigma_1} p_{\mathcal{G}}^{(\sigma_1, \sigma_2)}.
\end{aligned}
\tag{5.2.16}
$$

5.2.3.3 Expected Counts

The expected counts equal the total number of droplets D multiplied by the corresponding probabilities.

$$
\begin{aligned}
M_{\mathcal{G}}^{(\emptyset)} &:= D \cdot p_{\mathcal{G}}^{(\emptyset)}, \\
M_{\mathcal{G}}^{(s)} &:= D \cdot p_{\mathcal{G}}^{(s)}, \\
M_{\mathcal{G}}^{(s_1, s_2)} &:= D \cdot p_{\mathcal{G}}^{(s_1, s_2)}, \\
M_{\mathcal{G}}^{(3+)} &:= D \cdot p_{\mathcal{G}}^{(3+)}.
\end{aligned}
\tag{5.2.17}
$$

5.2.3.4   Observed Counts

The observed counts for the global goodness of fit test are

$$\hat{M}_{\mathcal{G}}^{(\emptyset)} := \sum_{d \in [D]} \mathbb{I}\{N_d(0) = 0\} \,,$$

$$\hat{M}_{\mathcal{G}}^{(s)} := \sum_{d \in [D]} \mathbb{I}\left\{N_d^{(s)}(0) \geq 1\right\} \cdot \prod_{\sigma \neq s} \mathbb{I}\left\{N_d^{(\sigma)}(0) = 0\right\} \,,$$

$$\hat{M}_{\mathcal{G}}^{(s_1, s_2)} := \sum_{d \in [D]} \left[\mathbb{I}\left\{N_d^{(s_1)}(0) \geq 1\right\} \cdot \mathbb{I}\left\{N_d^{(s_2)}(0) \geq 1\right\} \right. \tag{5.2.18}$$

$$\left. \cdot \prod_{\sigma \notin \{s_1, s_2\}} \mathbb{I}\left\{N_d^{(\sigma)}(0) = 0\right\} \right] \,,$$

$$\hat{M}_{\mathcal{G}}^{(3+)} := D - \hat{M}_{\mathcal{G}}^{(\emptyset)} - \sum_{\sigma=1}^{S} \hat{M}_{\mathcal{G}}^{(\sigma)} - \sum_{\sigma_1=1}^{S} \sum_{\sigma_2 > \sigma_1} \hat{M}_{\mathcal{G}}^{(\sigma_1, \sigma_2)} \,.$$

5.2.3.5   Symmetry of Definitions

Observe how, for any given pair of strains s_1, s_2:

$$\begin{aligned} p_{\mathcal{G}}^{(s_1, s_2)} &= p_{\mathcal{G}}^{(s_2, s_1)}, \\ M_{\mathcal{G}}^{(s_1, s_2)} &= M_{\mathcal{G}}^{(s_2, s_1)}, \\ \hat{M}_{\mathcal{G}}^{(s_1, s_2)} &= \hat{M}_{\mathcal{G}}^{(s_2, s_1)}. \end{aligned} \tag{5.2.19}$$

Thus for any given pair of strains s_1, s_2, only the left hand sides of (5.2.19) are considered, avoiding redundancy and overlapping categories.

5.2.3.6   Test Statistic

The test statistic is the Pearson categorical divergence of

$$D^{-1} \cdot \left(\hat{M}_{\mathcal{G}}^{(\emptyset)}, \hat{M}_{\mathcal{G}}^{(1)}, \dots, \hat{M}_{\mathcal{G}}^{(S)}, \hat{M}_{\mathcal{G}}^{(1,2)}, \dots, \hat{M}_{\mathcal{G}}^{(S-1,S)}, \hat{M}_{\mathcal{G}}^{(3+)} \right)$$

relative to   (5.2.20)

$$D^{-1} \cdot \left(M_{\mathcal{G}}^{(\emptyset)}, M_{\mathcal{G}}^{(1)}, \dots, M_{\mathcal{G}}^{(S)}, M_{\mathcal{G}}^{(1,2)}, \dots, M_{\mathcal{G}}^{(S-1,S)}, M_{\mathcal{G}}^{(3+)} \right) \,.$$

We can get an asymptotic p-value in the standard way, via the survival function of the χ^2 distribution with $\left(1 + S + \binom{S}{2} + 1\right) - 1$ degrees of freedom.

5.2.4 Pairwise Gluttonous Goodness of Fit Tests

In section 5.2.3 droplets were grouped according to strain in a non-overlapping way. This made it possible to compare other distributions to hPoMu using a single hypothesis test, but for practical reasons it also required lumping together all droplets with three or more strains. However, if we are willing to perform multiple hypothesis tests, we can better utilize the information contained in those droplets. By performing a separate hypothesis test for each pair of strains, we can include more droplets in both our "treatment" and "control" groups.

The definition of the hypothesis test corresponding to each pair of strains is split into several parts. Section 5.2.4.1 defines how the droplets are grouped. Section 5.2.4.2 gives probabilities for each group under the null hypothesis (hPoMu), while section 5.2.4.3 clarifies the expected counts. Section 5.2.4.4 defines the observed counts used in defining the goodness of fit test. Section 5.2.4.5 explains why some of the groups are redundant. Finally, section 5.2.4.6 combines the information from the previous sections to define the test statistic.

5.2.4.1 Categories

Note that the above definitions correspond to "gluttonous" groups, unlike the "global" test defined before which corresponds to "picky" groups. The "gluttony" means that the categories for distinct combinations of strains can overlap. This is what necessitates separate tests for each combination of strains, i.e. why they are "pairwise".

5.2.4.2 Probabilities

The probabilities $p_{0,0}^{(s_1,s_2)}, p_{1,0}^{(s_1,s_2)}, p_{1,1}^{(s_1,s_2)}, p_{0,1}^{(s_1,s_2)}$, for the pairwise tests are

$$
\begin{aligned}
p_{0,0}^{(s_1,s_2)} &:= \mathbb{P}_{\mathrm{hPoMu}(\lambda,\vec{f})}\big(N^{(s_1)}(0) = 0, N^{(s_2)}(0) = 0\big) \\
&= e^{-f^{(s_1)}\lambda} \cdot e^{-f^{(s_2)}\lambda}, \\
p_{1,0}^{(s_1,s_2)} &:= \mathbb{P}_{\mathrm{hPoMu}(\lambda,\vec{f})}\big(N^{(s_1)}(0) \geq 1, N^{(s_2)}(0) = 0\big) \\
&= \big(1 - e^{-f^{(s_1)}\lambda}\big) \cdot \cdot e^{-f^{(s_2)}\lambda}, \\
p_{1,1}^{(s_1,s_2)} &:= \mathbb{P}_{\mathrm{hPoMu}(\lambda,\vec{f})}\big(N^{(s_1)}(0) \geq 1, N^{(s_2)}(0) \geq 1\big) \\
&= \big(1 - e^{-f^{(s_1)}\lambda}\big) \cdot \big(1 - e^{-f^{(s_2)}\lambda}\big), \\
p_{0,1}^{(s_1,s_2)} &:= \mathbb{P}_{\mathrm{hPoMu}(\lambda,\vec{f})}\big(N^{(s_1)}(0) = 0, N^{(s_2)}(0) \geq 1\big) \\
&= e^{-f^{(s_1)}\lambda} \cdot \big(1 - e^{-f^{(s_2)}\lambda}\big).
\end{aligned}
\tag{5.2.21}
$$

Note the implicit dependence on an assumed λ in the definitions of $p_{0,0}^{(s_1,s_2)}$, $p_{1,0}^{(s_1,s_2)}$, $p_{1,1}^{(s_1,s_2)}$, $p_{0,1}^{(s_1,s_2)}$. Thus the scientist needs an estimate, or some a priori knowledge, of the value of λ to use these tests.

5.2.4.3 Expected Counts

The expected counts for the pairwise tests are

$$
\begin{aligned}
M_{0,0}^{(s_1,s_2)} &:= D \cdot p_{0,0}^{(s_1,s_2)}, \\
M_{1,0}^{(s_1,s_2)} &:= D \cdot p_{1,0}^{(s_1,s_2)}, \\
M_{1,1}^{(s_1,s_2)} &:= D \cdot p_{1,1}^{(s_1,s_2)}, \\
M_{0,1}^{(s_1,s_2)} &:= D \cdot p_{0,1}^{(s_1,s_2)}.
\end{aligned}
\tag{5.2.22}
$$

5.2.4.4 Observed Counts

The observed counts are defined as

$$
\begin{aligned}
\hat{M}_{0,0}^{(s_1,s_2)} &:= \left| \left\{ d : N_d^{(s_1)}(0) = 0, N_d^{(s_2)}(0) = 0 \right\} \right|, \\
\hat{M}_{1,0}^{(s_1,s_2)} &:= \left| \left\{ d : N_d^{(s_1)}(0) \geq 1, N_d^{(s_2)}(0) = 0 \right\} \right|, \\
\hat{M}_{1,1}^{(s_1,s_2)} &:= \left| \left\{ d : N_d^{(s_1)}(0) \geq 1, N_d^{(s_2)}(0) \geq 1 \right\} \right|, \\
\hat{M}_{0,1}^{(s_1,s_2)} &:= \left| \left\{ d : N_d^{(s_1)}(0) = 0, N_d^{(s_2)}(0) \geq 1 \right\} \right|.
\end{aligned}
\tag{5.2.23}
$$

5.2.4.5 Symmetry of Definitions

Of course,

$$
\begin{aligned}
\hat{M}_{1,0}^{(s_1,s_2)} &= \hat{M}_{0,1}^{(s_2,s_1)}, & p_{1,0}^{(s_1,s_2)} &= p_{0,1}^{(s_2,s_1)}, \\
\hat{M}_{0,0}^{(s_1,s_2)} &= \hat{M}_{0,0}^{(s_2,s_1)}, & p_{0,0}^{(s_1,s_2)} &= p_{0,0}^{(s_2,s_1)}, \\
\hat{M}_{1,1}^{(s_1,s_2)} &= \hat{M}_{1,1}^{(s_2,s_1)}, & p_{1,1}^{(s_1,s_2)} &= p_{1,1}^{(s_2,s_1)}.
\end{aligned}
\tag{5.2.24}
$$

Therefore the results of the test for (s_2, s_1) will always be the same as those for (s_1, s_2), and so at most one of the two tests needs to be performed.

5.2.4.6 Test Statistic

The test statistic is the Pearson categorical divergence of

$$
D^{-1} \cdot \left(\hat{M}_{0,0}^{(s_1,s_2)}, \hat{M}_{1,0}^{(s_1,s_2)}, \hat{M}_{1,1}^{(s_1,s_2)}, \hat{M}_{0,1}^{(s_1,s_2)} \right)
$$
$$
\text{relative to} \quad D^{-1} \cdot \left(M_{0,0}^{(s_1,s_2)}, M_{1,0}^{(s_1,s_2)}, M_{1,1}^{(s_1,s_2)}, M_{0,1}^{(s_1,s_2)} \right),
\tag{5.2.25}
$$

namely

$$
\sum_{i=0}^{1} \sum_{j=0}^{1} \frac{\left(\hat{M}_{i,j}^{(s_1,s_2)} - M_{i,j}^{(s_1,s_2)} \right)^2}{M_{i,j}^{(s_1,s_2)}}.
\tag{5.2.26}
$$

One can then get an asymptotic *p*-value for this test in the standard way, via the survival function of the χ^2 distribution with 3 degrees of freedom.

5.3   Methods

Section 5.3.1 explains which distributions were simulated, why, and how. Section 5.3.2 explains how categorical divergences and corresponding approximate p-values were computed. Section 5.3.3 explains how percentages and percent changes for droplets with a given number of strains were computed. Section 5.3.4 explains how I computed the percent changes for gluttonous groups. Section 5.3.5 explains how I computed the percent changes for picky groups. I plotted results using Matplotlib [Hun07] version 3.4.1 and Seaborn [Was21] version 0.11.1. The bar plot was made using the *microbiome* [LS19], *phyloseq* [MH13], and *ggplot2* [Wic16] packages for the R programming language. Complete implementation details can be found in the code at the relevant GitLab repository. See `https://gitlab.com/krinsman/droplets`.

5.3.1   Simulated Distributions

To test how much the assumptions of hPoMu may be violated with hPoMu remaining an adequate working model, I simulated seven distributions. Below, I describe the two new distributions beyond the five distributions that were already described in section 4.3.1. Details of the simulations common to all seven distributions were given already in section 4.3.1.4.

5.3.1.1   Simulating High Compositional Heterogeneity Only

This distribution was a hierarchical Exponential hPoDM model (hExhPoDM), with ζ_C distributed as an Exponential(1) random variable.

5.3.1.2   Simulating High Compositional and Density Heterogeneities

This distribution was a hierarchical Exponential hNBDM model (hExhNBDM), and ζ was distributed as an Exponential(1) random variable.

5.3.2   Computing p-values and χ^2 Divergences

Using the categories and expected counts defined for the "global" goodness of fit hypothesis test of the null hPoMu using "picky" groups defined in section 5.2.1.3, I computed Pearson categorical ("χ^2", cf. section 5.2.2) divergence values and χ^2-approximated p-values for all 500 simulations of all 7 distributions.

Despite that each simulation corresponded to an extremely large number of droplets, the expected counts for some categories were very small (< 5), making the asymptotic approximation provided by the χ^2 assumption possibly inadequate. To account for this, I generated 10^9 (one **b**illion) independent replicates of the Multinomial distribution corresponding to the number of droplets per simulation and the probabilities for each category in the goodness of fit test. From this, I generated 10^9 replicates from the sampling distribution under the null of the Pearson categorical divergence statistic. I then used this Monte Carlo distribution of Pearson categorical divergence statistics to compute "Monte Carlo p-values" for all 500

simulations from all 7 distributions. Given the very large number of replicates, I expect these p-values to be more accurate than those from the χ^2 asymptotic approximation. Even so the results using the χ^2 approximation were similar, cf. section 5.A.

5.3.3   Numbers of Droplets with n Strains

Using the relationships:

$$
\mathbb{E}[|\{d \in [D] : \mathbb{S}_d(0) = n\}|] = D \cdot \mathbb{P}(\mathbb{S}(0) = n)
$$

$$
= D \cdot \left[\sum_{\mathcal{S} \subseteq [S] : |\mathcal{S}| = n} \mathbb{P}(\mathrm{supp}(\vec{\mathbb{S}}(0)) = \mathcal{S}) \right], \tag{5.3.1}
$$

I computed via brute force these quantities for $n \in [5]$ for hPoMu:

$$
D \cdot \left[\sum_{\mathcal{S} \subseteq [S] : |\mathcal{S}| = n} e^{-\sum_{s \notin \mathcal{S}} f^{(s)} \lambda} \cdot \prod_{s \in \mathcal{S}} \left(1 - e^{-f^{(s)} \lambda} \right) \right]
$$

$$
= D \cdot \left[\sum_{\mathcal{S} \subseteq [S] : |\mathcal{S}| = n} e^{-(1 - \sum_{s \in \mathcal{S}} f^{(s)}) \lambda} \cdot \prod_{s \in \mathcal{S}} \left(1 - e^{-f^{(s)} \lambda} \right) \right]. \tag{5.3.2}
$$

The value for 6+ strains is just D minus the sum of the values for all $n \in [5]$.

For each of the seven distributions, I stratified and then counted the droplets from each of the 500 simulations according to how many strains were present. I then computed the average values for each distribution as the arithmetic mean over the 500 simulations. (By the law of large numbers the arithmetic mean is consistent for the true expectation and the only reasonable choice of measure of "central tendency" in this context.) I then computed the average percent change for each distribution by subtracting the value expected under hPoMu from the mean observed value, and then dividing by the expected value.

5.3.4   Gluttonous Groups

Using the formulae (5.2.21) and (5.2.22), I computed the expected count for each treatment gluttonous group[6] under hPoMu. For each of the seven distributions, for each of the 500 simulations using NumPy [HMvdW+20] version 1.20.2 I computed via brute force the number of occurrences of each of the gluttonous groups, and then reported the arithmetic means over all 500 simulations. I again computed the average percent change as the observed value minus the value expected under hPoMu divided by the value expected under hPoMu.

---

[6] For the gluttonous "controls", i.e. corresponding to the diagonals of the heatmaps, I actually used the sets $\{d \in [D] : N^{(s)}(0) \geq 1\}$, which is a superset of all gluttonous controls for experiments measuring effects on the growth of strain s. The expected size of these sets is $D \cdot \mathbb{E}[\mathbb{S}^{(s)}(0)] = D \cdot \mathbb{P}(N^{(s)}(0) \geq 1)$, which under hPoMu equals $D \cdot (1 - e^{-f^{(s)} \lambda})$.

Initially I considered "filtering" differences depending on whether they were "significant" according to the (asymptotic) p-values of the hypothesis tests defined in section 5.2.4. However, I found that doing so made the results more difficult to interpret by obscuring overall trends. The inferred overall trends were qualitatively the same in either case. Moreover, because each gluttonous pair corresponds to a distinct hypothesis test, and the gluttonous groups overlap, this leads to a difficult multiple comparisons problem. One has to not only decide how to account for multiple testing between simulations but also "within" simulations. Each of the multiple testing correction schemes I tried seemed objectionable in some way. Given all of the above, I decided to report the "raw" average percent change without "filtering" for "significance".

5.3.5 Picky Groups

Using the formulae (5.2.16) and (5.2.17), cf. also the equations (5.2.12) and (5.2.13), I computed the expected count for each picky group under hPoMu. Note that, unlike for gluttonous treatment groups, picky treatment groups do not overlap. Moreover, unlike for gluttonous control groups, the picky control group for investigating the effect of strain s_1 on the rate of growth of strain s_2 is the same regardless of what s_1 is. So the values shown on the diagonals of the heat maps really correspond to the actual picky control groups. The procedure was then otherwise exactly the same as for the gluttonous groups. I did not "filter" according to "significance".

5.4 Results

Section 5.4.1 examines the results of the global χ^2 goodness of fit tests computed from the picky contingency tables. Section 5.4.2 directly examines both the picky and gluttonous contingency tables.

5.4.1 Global χ^2 Tests

Section 5.4.1.1 discusses evidence indicating that sampling with replacement is the most innocuous of the assumptions behind hPoMu. Section 5.4.1.2 discusses evidence indicating that ghNBDM distributions with low (but nonzero) heterogeneity may be adequately modeled by hPoMu. Section 5.4.1.3 discusses evidence indicating that ghNBDM distributions with only moderate heterogeneity can be easily distinguished from hPoMu. Finally, section 5.4.1.4 discusses evidence indicating that even further heterogeneity can make differences from hPoMu extremely obvious.

5.4.1.1 Sampling without Replacement

All available evidence suggests that, even after $15,000,000$ droplets have been formed, the hTPMH distribution for the chosen value of K is extremely similar to the corresponding
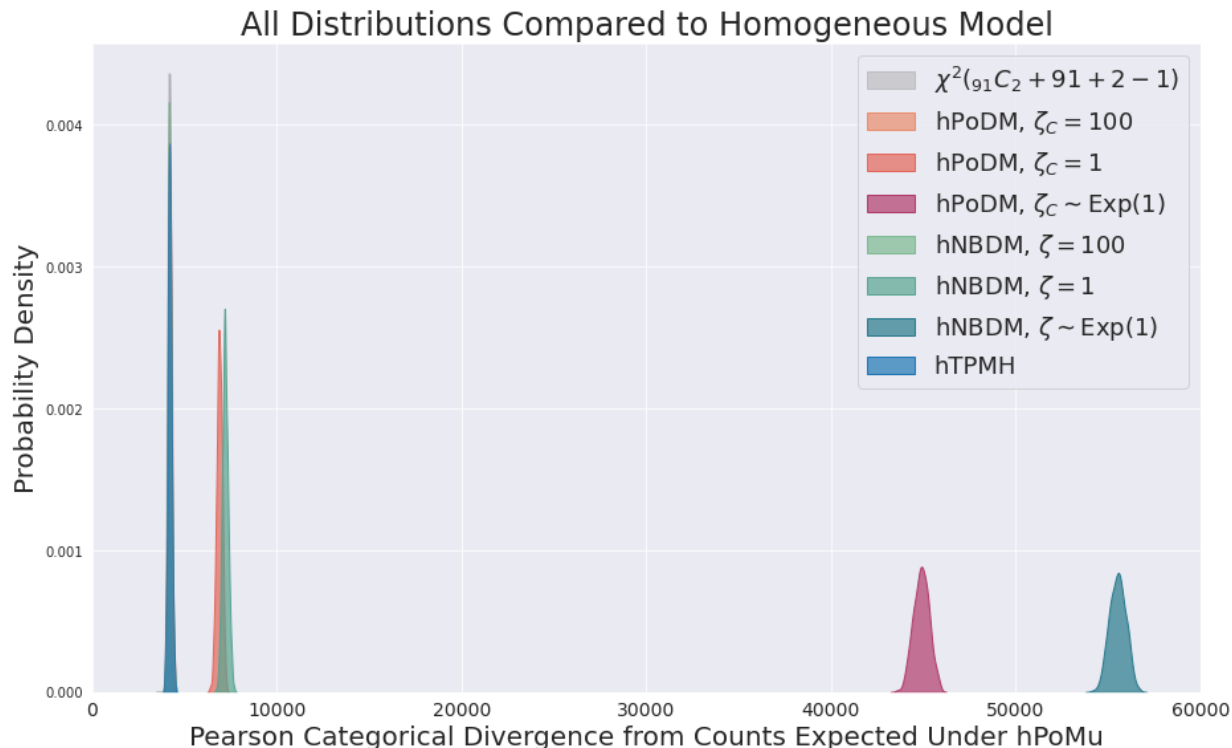
Figure 5.2

hPoMu distribution.

Goodness of Fit p-Values Are Approximately Uniformly Distributed   For the hTPMH distribution, figure 5.3 shows how the distribution of (Monte Carlo) p-values over 500 simulations is approximately uniform, as it would be if we had actually sampled from the true hPoMu null distribution instead.

Distribution of Pearson Categorical Divergences for Sampling without Replacement   Moreover, as seen clearly in figure 5.4 (and less so in figure 5.2), the observed distribution of Pearson categorical divergences from the hTPMH distribution overlaps strongly with approximately what would have been expected under the hPoMu null distribution.

5.4.1.2   Very Similar Distributions

All available evidence suggests that the hPoDM with $\zeta_C = 100$ and hNBDM with $\zeta = 100$ distributions are very similar to the corresponding hPoMu distribution. Nevertheless, they also appear to differ slightly more than the hTPMH distribution.
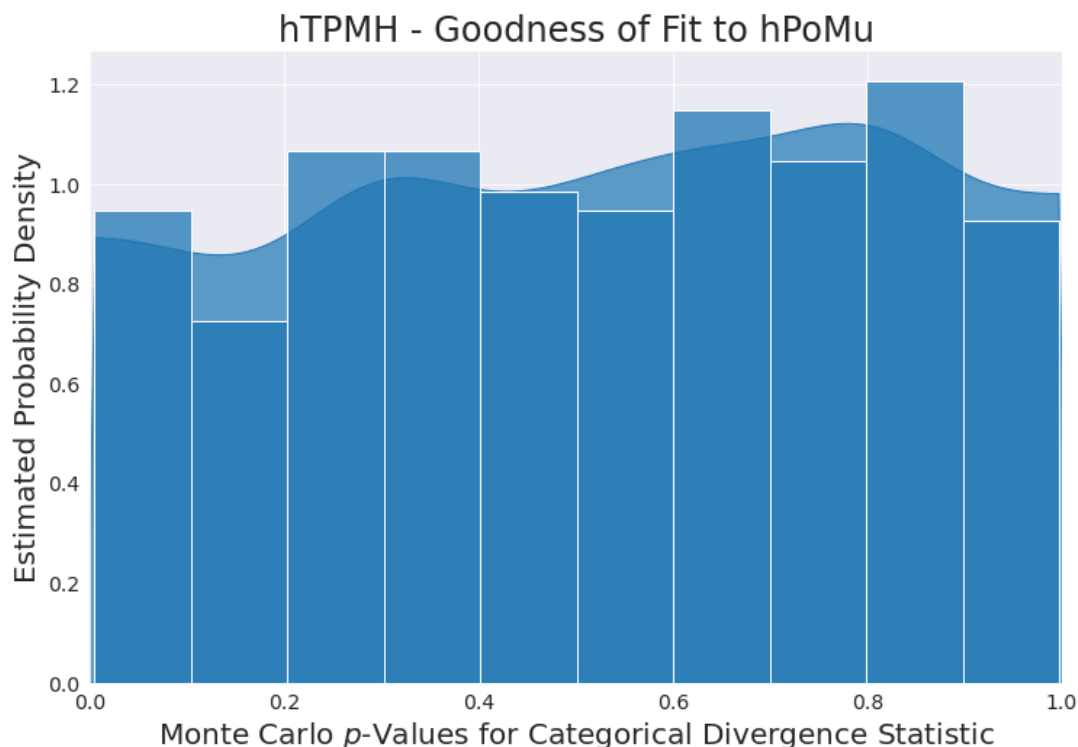
Figure 5.3: Monte Carlo p-value distribution for sampling without replacement.

Goodness of Fit p-Values Are Roughly Uniformly Distributed For the hPoDM distribution with $\zeta_C = 100$ and the hNBDM distribution with $\zeta = 100$, figures 5.5a and 5.5b show respectively how their distributions of (Monte Carlo) p-values over 500 simulations is roughly similar to the uniform distribution.

The p-value distribution for hNBDM with $\zeta = 100$ in figure 5.5b appears to fit the uniform distribution worse than the corresponding distribution for hPoDM with $\zeta_C = 100$ in figure 5.5a. This could possibly be a reflection of how the hNBDM family incorporates both kind of heterogeneities whereas the hPoDM family does not. However, it could also possibly be a reflection merely of the relatively small ($n = 500$) sample size.

Distributions of Pearson Categorical Divergences for Very Similar Distributions Furthermore, figure 5.4 indicates how the distribution of their Pearson categorical divergence statistics also overlaps well with roughly what would be expected under the null distribution.

The overlap for hNBDM $\zeta = 100$ seen in figure 5.4 does not seem to be substantially worse (nor better) than that observed for hPoDM $\zeta_C = 100$, in contrast to the extent of difference that might have been anticipated after comparing figures 5.5a and 5.5b.
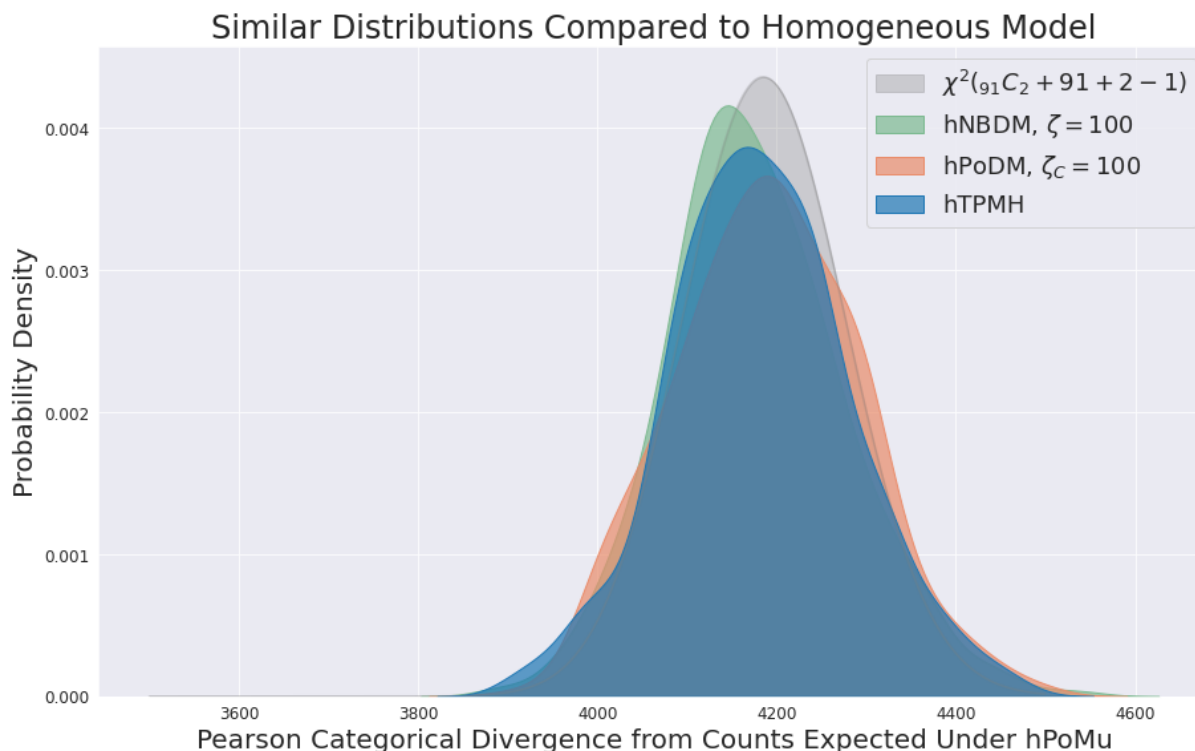
Figure 5.4

5.4.1.3 Slightly Different Distributions

All available evidence suggests that the hPoDM with $\zeta_C = 1$ and hNBDM with $\zeta = 1$ distributions are (at least) slightly different from the corresponding hPoMu distribution.

Goodness of Fit p-Values Are All Effectively Zero For the hPoDM distribution with $\zeta_C = 1$ and the hNBDM distribution with $\zeta = 1$, all of the approximate p-values were indistinguishable from 0 up to numerical precision. Therefore their distributions of approximate p-values, being a "single point spike" or "Dirac delta", were not plotted.

Distributions of Pearson Categorical Divergences for Slightly Different Distributions The reason behind these low p-values is fairly obvious from either figure 5.2 or 5.6: their distributions of observed Pearson categorical divergences overlap almost not at all with what would be anticipated under the hPoMu null distribution. Also, the distribution of Pearson categorical divergences for hNBDM with $\zeta = 1$ is slightly further to the right than that of hPoDM with $\zeta_C = 1$, reflecting how the former incorporates both density and compositional heterogeneity while the latter incorporates only compositional heterogeneity.
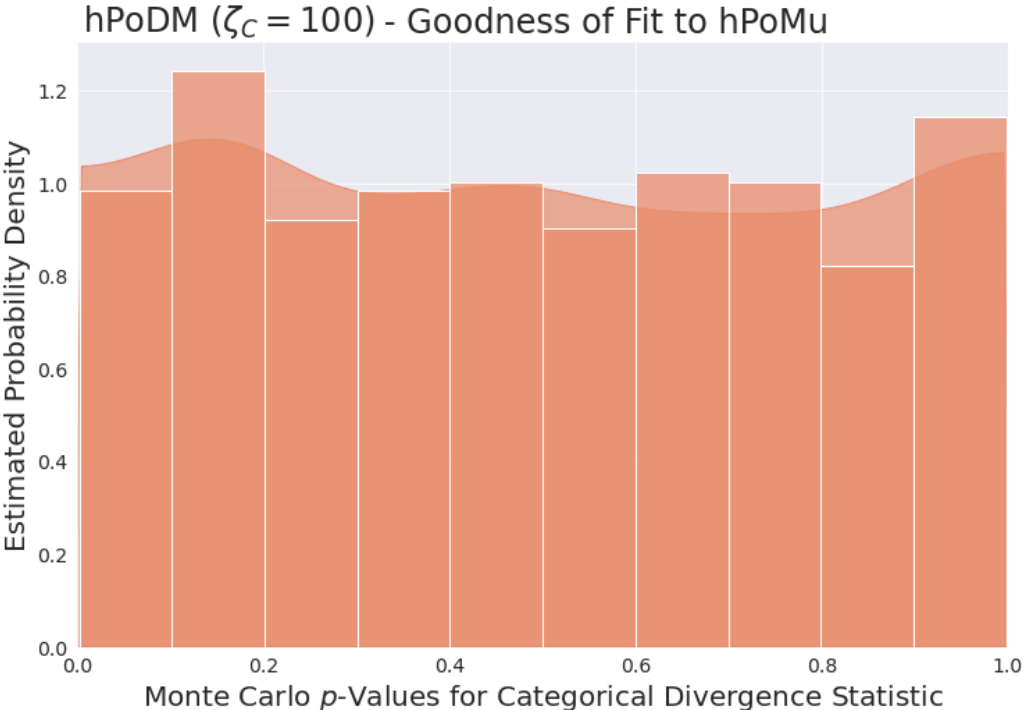
(a)
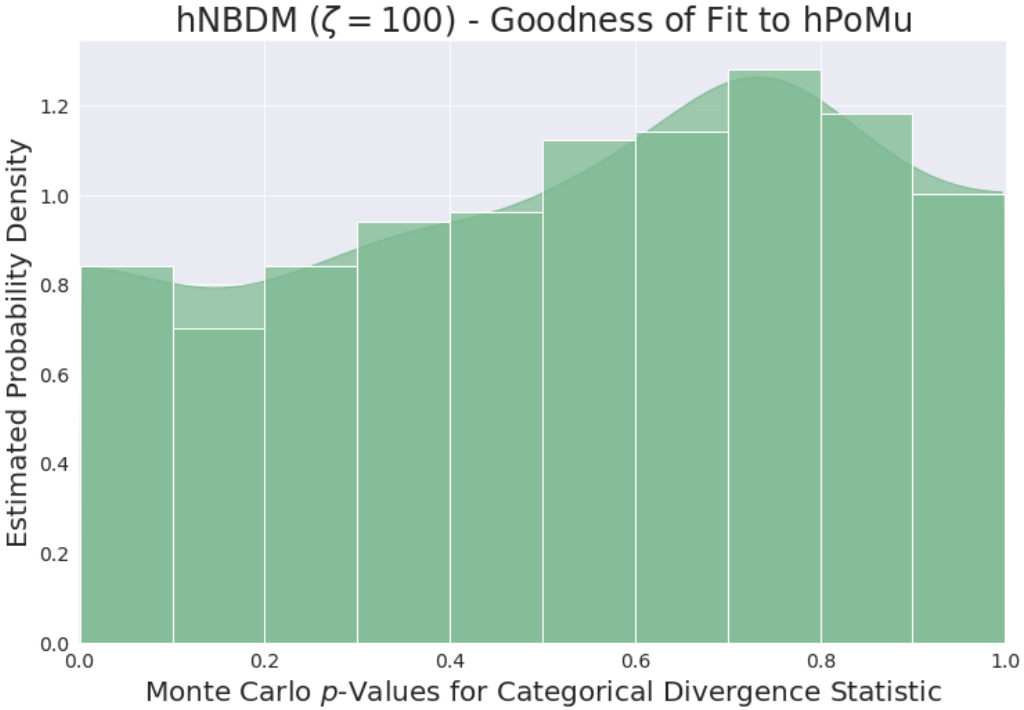
(b)

Figure 5.5: Monte Carlo goodness of fit p-value distributions for very similar distributions.
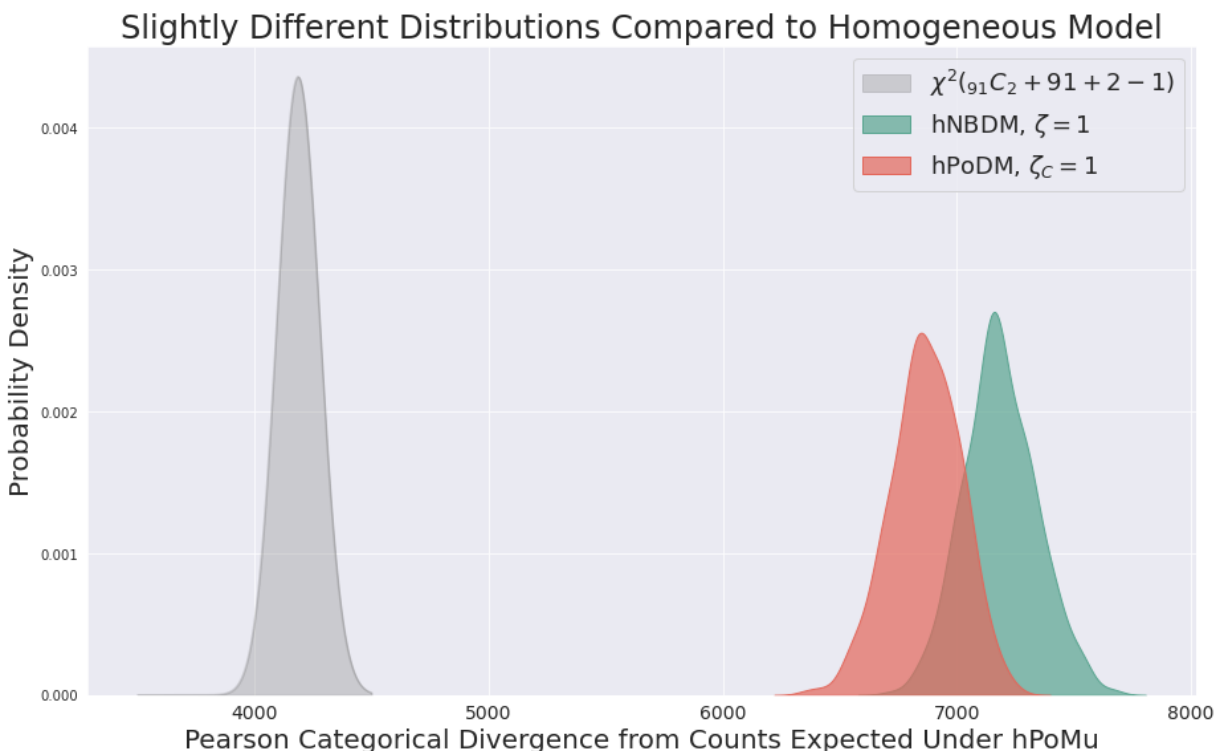
Figure 5.6

5.4.1.4   Very Different Distributions

All available evidence suggests that the hExhPoDM with $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM with $\mathbb{E}[\zeta] = 1$ distributions are very different from the corresponding hPoMu distribution.

Goodness of Fit p-Values Are All Effectively Zero   For the hExhPoDM with $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM with $\mathbb{E}[\zeta] = 1$ distributions, all of the (approximate) p-values were again indistinguishable from 0 up to numerical precision, as was also the case for hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$. The approximate p-values were again not plotted for the same reasons as for hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$.

Distributions of Pearson Categorical Divergences for Very Different Distributions   The reason for these small p-values is also the same as for hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$: as can be seen clearly in either figure 5.2 or figure 5.7, the peaks of their distributions of Pearson categorical divergences completely fail to overlap with the range of values that would be anticipated under the hPoMu null distribution. Both peaks are also *much* further to the right than anything observed for any of the other five distributions, indicating that the hExhPoDM $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM $\mathbb{E}[\zeta] = 1$ distributions are (by far) the most heterogeneous of the seven distributions simulated.
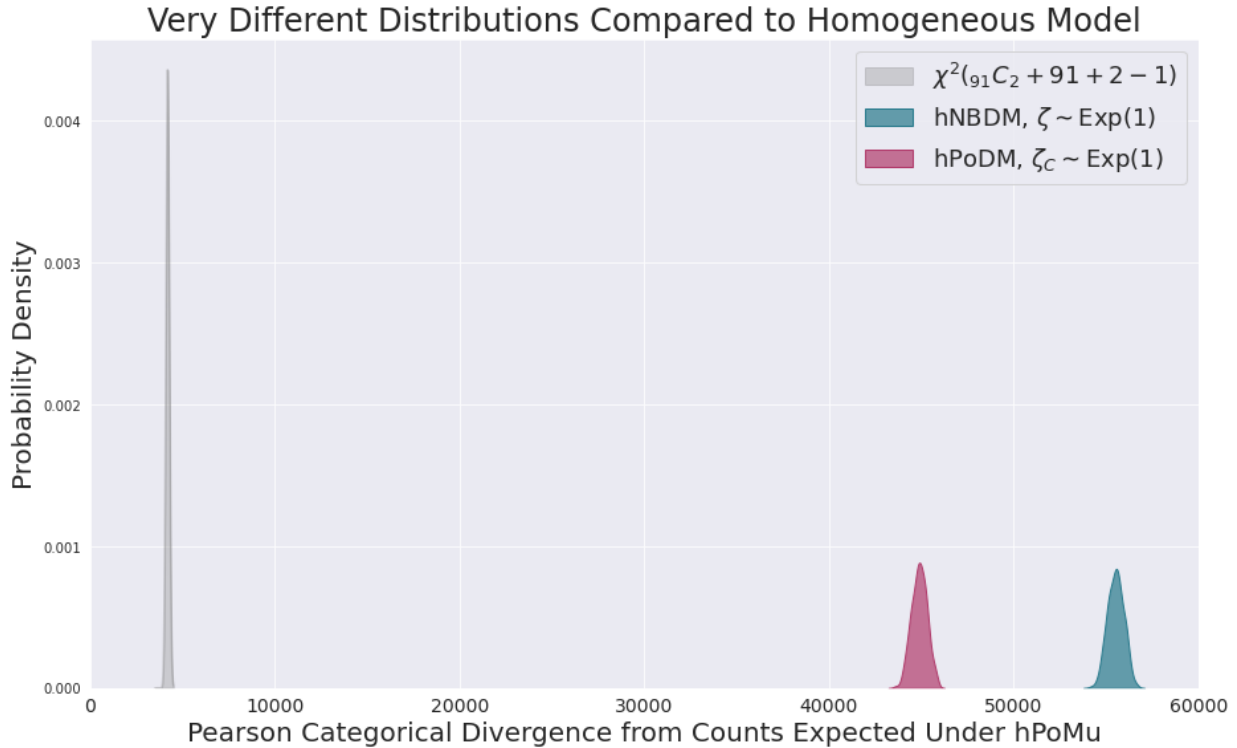
Figure 5.7

Analogous to what occurs for hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, the peak for hExhNBDM $\mathbb{E}[\zeta] = 1$ is further to the right than the peak for hExhPoDM $\mathbb{E}[\zeta_C] = 1$, reflecting how the former incorporates both kind of heterogeneities whereas the latter does not. (Thus hExhNBDM $\mathbb{E}[\zeta] = 1$ can be considered the single most heterogeneous distribution of the seven that were simulated.) However, in contrast to the situation observed for hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, the peaks for hExhNBDM $\mathbb{E}[\zeta] = 1$ and hExhPoDM $\mathbb{E}[\zeta_C] = 1$ do not overlap at all, with the peak for hExhNBDM $\mathbb{E}[\zeta] = 1$ being greatly further to the right than that for hExhPoDM $\mathbb{E}[\zeta_C] = 1$.

5.4.2   Direct Examination of Contingency Tables

Section 5.4.2.1 explains how the observed effects on data throughput support the "taxonomy" of distributions previously proposed in section 4.4. Section 5.4.2.2 explains the evidence indicating that increased heterogeneity decreases the average number of droplets with 3 or more strains. Section 5.4.2.3 explains the effect of increased heterogeneity on data throughput for droplets with 2 or fewer strains, with section 5.4.2.3 explaining what happens for compositional heterogeneity without density heterogeneity, and section 5.4.2.3 explaining what happens when both types of heterogeneity are present.

(a)
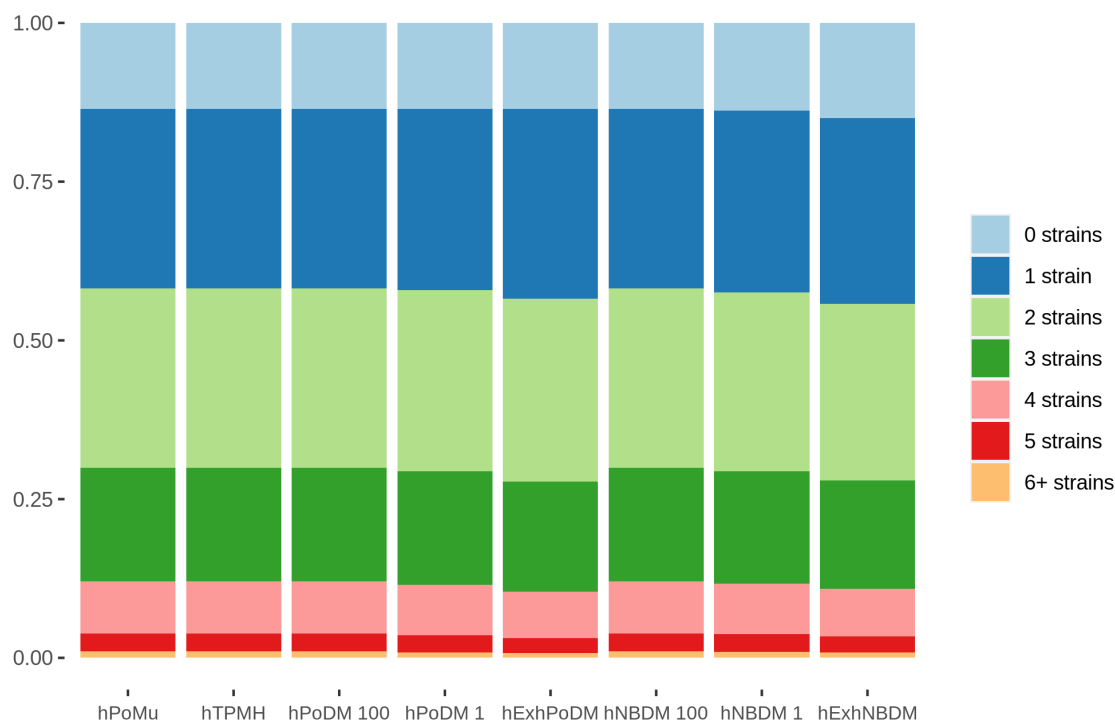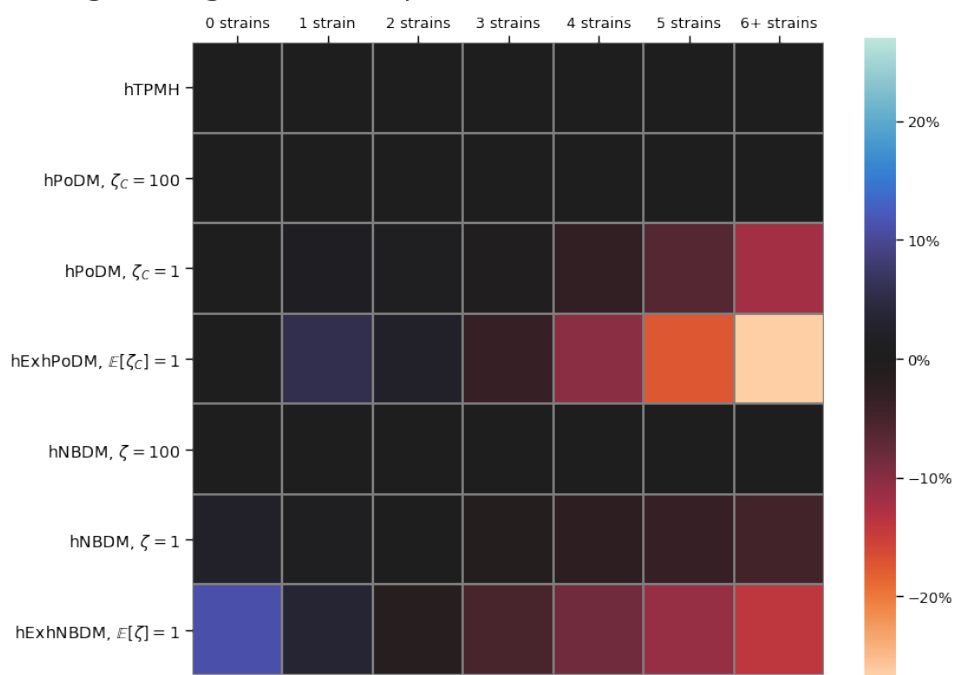
(b)

Figure 5.8: Droplets stratified by numbers of strains found inside droplets.
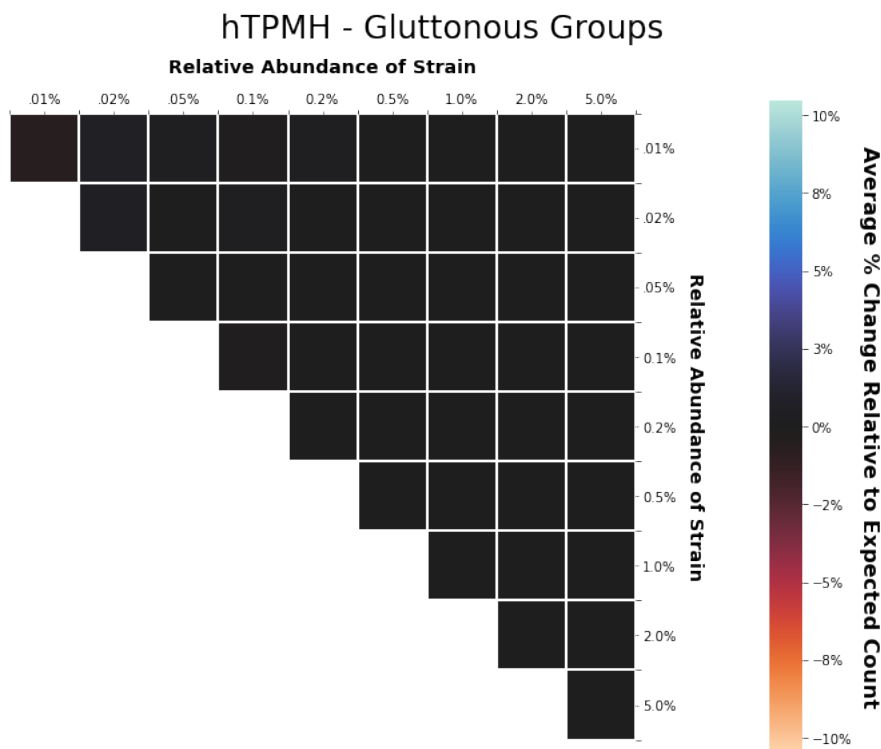
(a)

Figure 5.9: Gluttonous groups for low heterogeneity distributions.

5.4.2.1 Distributions Again Fall Into Same Groups

From figures 5.8a, 5.8b, 5.9a, 5.9b, 5.9c, 5.10a, 5.10b, and 5.10c we see that hTPMH, hPoDM with $\zeta_C = 100$, and hNBDM with $\zeta_C = 100$ are again (cf. section 4.4) either indistinguishable from or very similar to hPoMu. Any differences in average composition are only visible for medium and high heterogeneities, and easily visible usually only for high heterogeneity.

5.4.2.2 Increased Heterogeneity Decreases Number of Droplets with 3 or More Strains

From figures 5.8a and 5.8b we see that, both for the hPoDM and hNBDM families, as the compositional heterogeneity increases, the number of droplets with 3 or more strains decreases in both absolute and relative terms. Since droplets with 3 or more strains are a component of the gluttonous groups, their decrease with increased compositional heterogeneity is also reflected in figures 5.11a, 5.11c, 5.11b, and 5.11d. We can also see from figures 5.8b, 5.11a, 5.11c, 5.11b, and 5.11d that this decrease in the number of droplets with 3 or more strains is more pronounced in the absence of density heterogeneity.
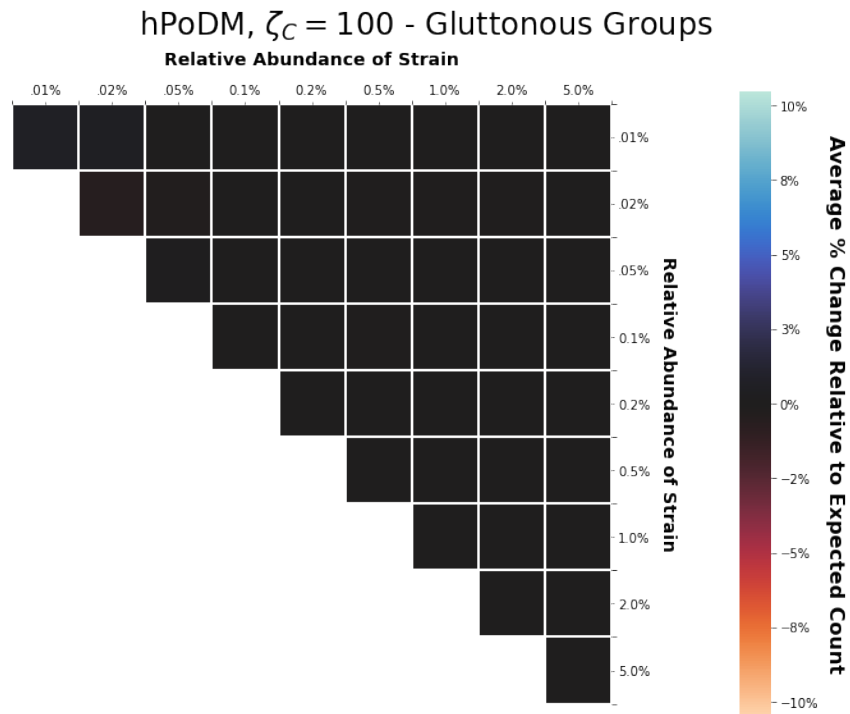
(b)
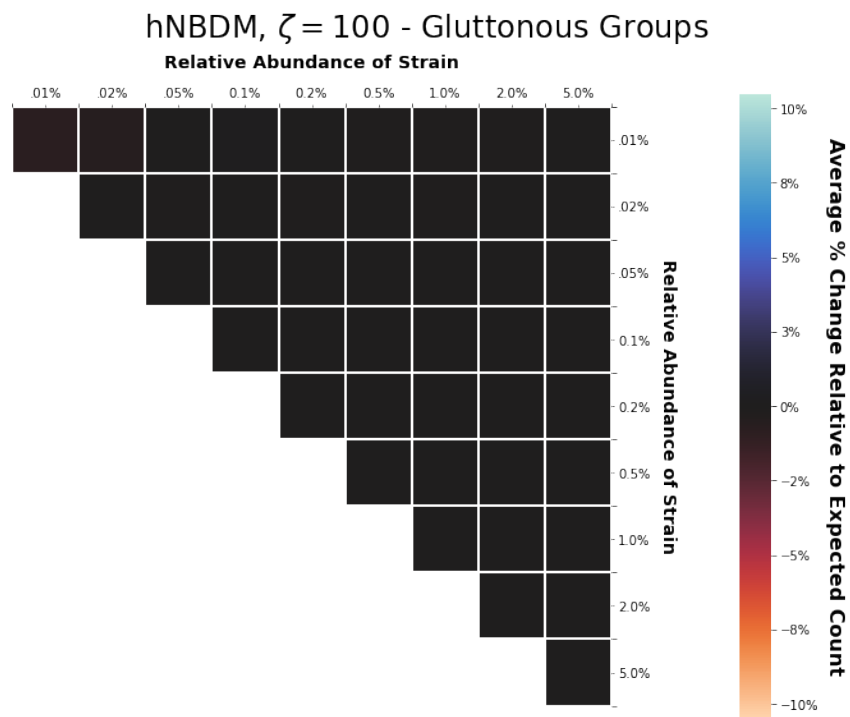
(c)

Figure 5.9: Gluttonous groups for low heterogeneity distributions.
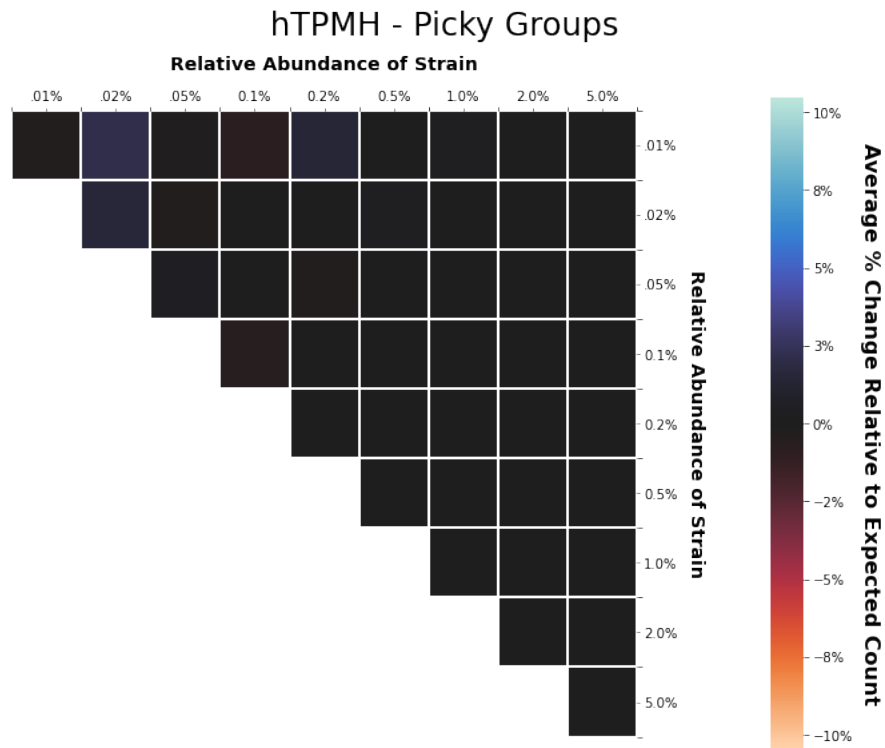
(a)

Figure 5.10: Picky groups for low heterogeneity distributions.

5.4.2.3 Effect of Increased Heterogeneity on Droplets with Fewer than 3 Strains

For droplets with fewer than 3 strains, the effects that occur as compositional heterogeneity increases depend on whether density heterogeneity is present.

Increased Compositional Heterogeneity Only In the absence of density heterogeneity, we see an increase in the number of droplets with one strain, with a smaller increase in the number of droplets with two strains. This is best evidenced in figure 5.8b. This slight increase in the number of droplets with two strains causes the slight increase in the number of picky treatments shown in figures 5.12a and 5.12b. As is clear also from figure 5.8a, there is little change in the number of empty droplets with zero strains regardless of the level of compositional heterogeneity.

Increased Compositional and Density Heterogeneity In the presence of density heterogeneity, as seen in figure 5.8b the number of droplets with a single strain does still increase slightly. However, in contrast to what happens in the absence of density heterogeneity, in the presence of density heterogeneity there is an observable decrease in the number of droplets
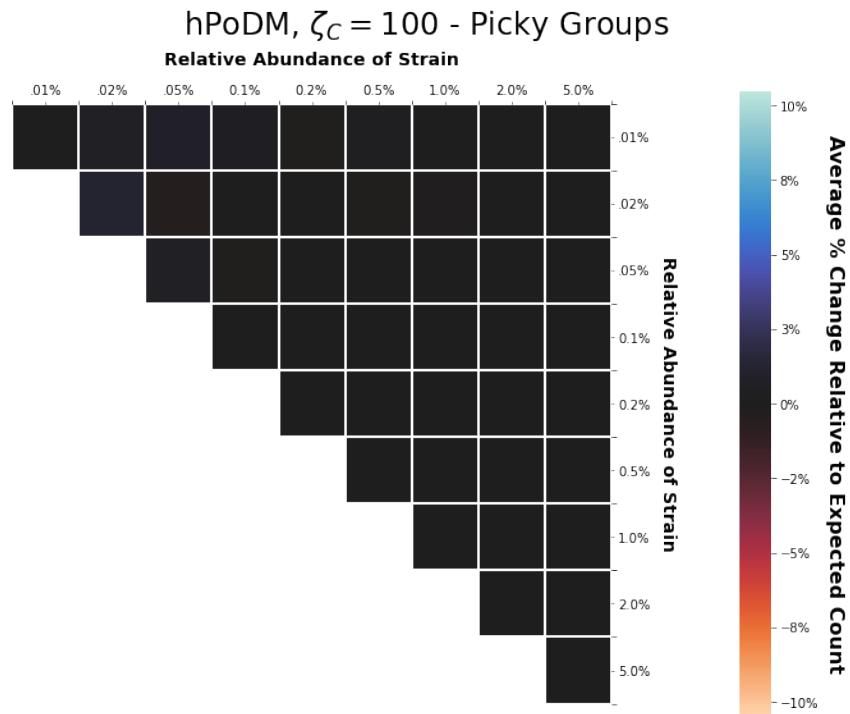
(b)
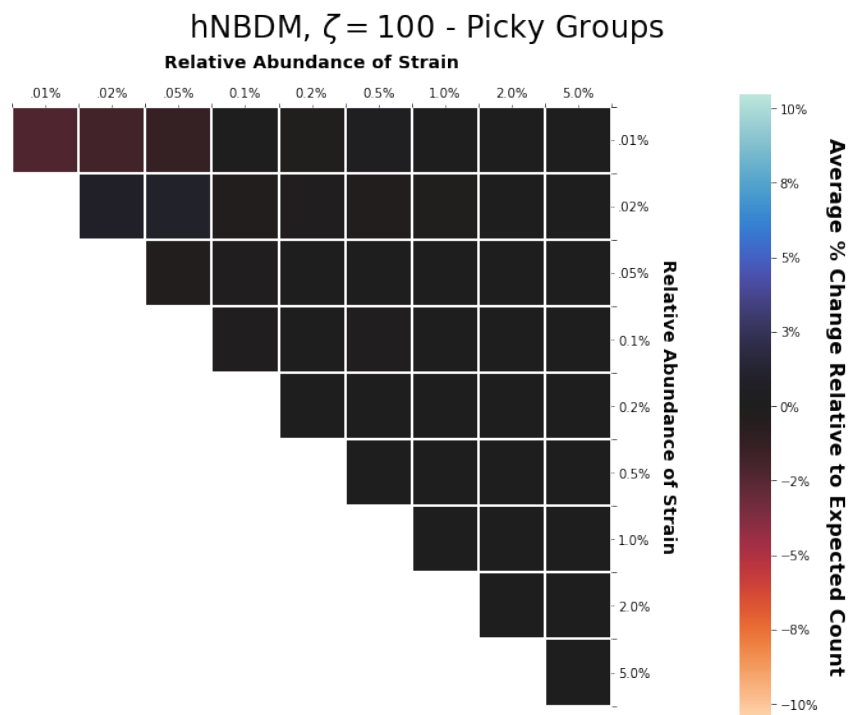
(c)

Figure 5.10: Picky groups for low heterogeneity distributions.
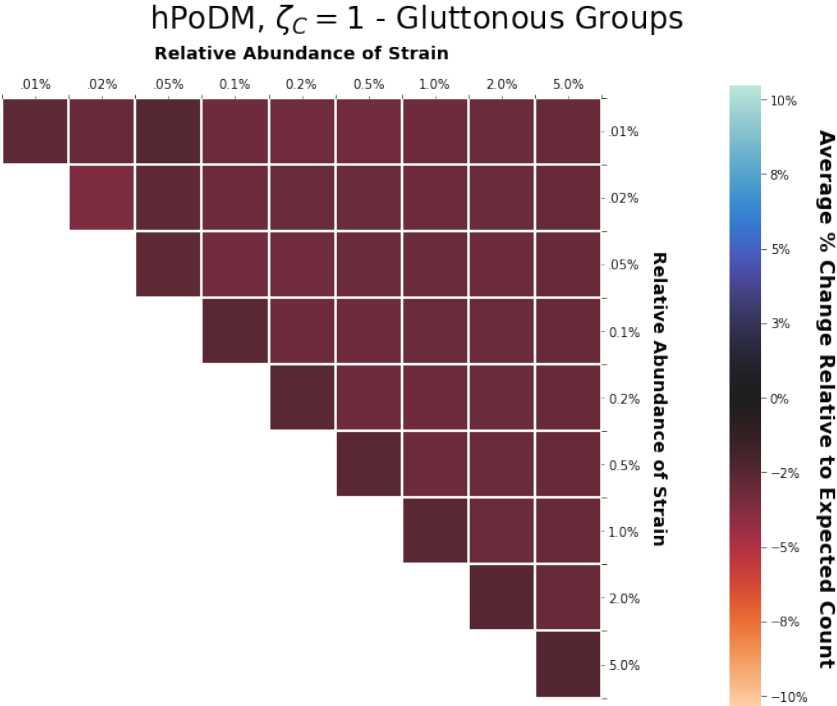
(a)
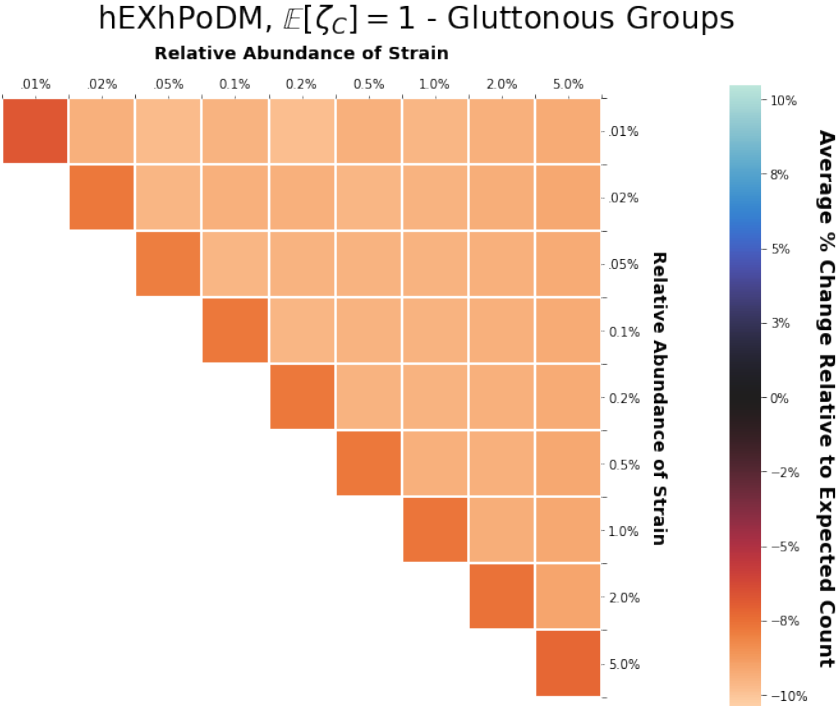
(b)

Figure 5.11: Gluttonous groups for high heterogeneity distributions.
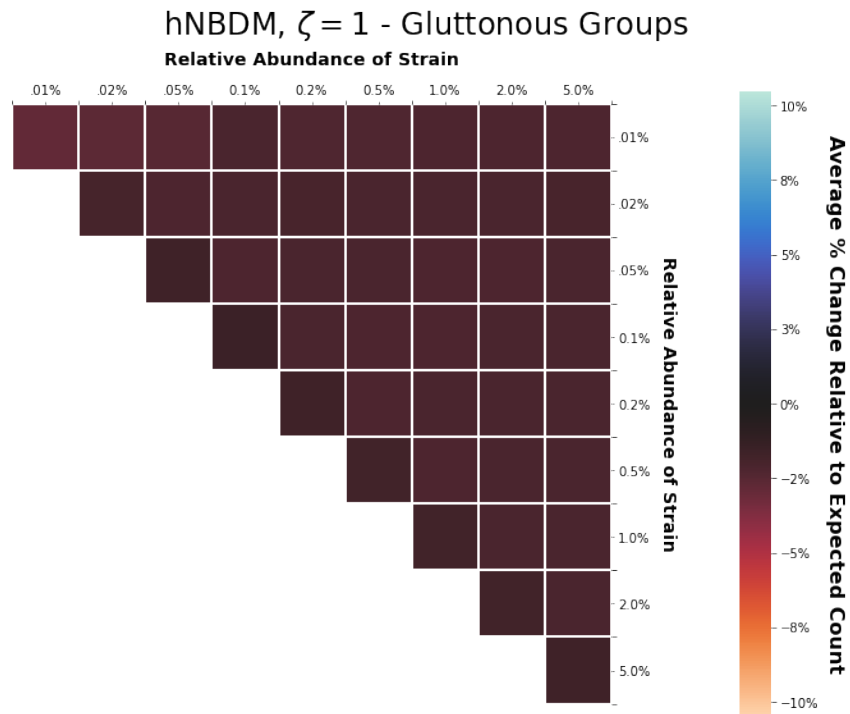
(c)
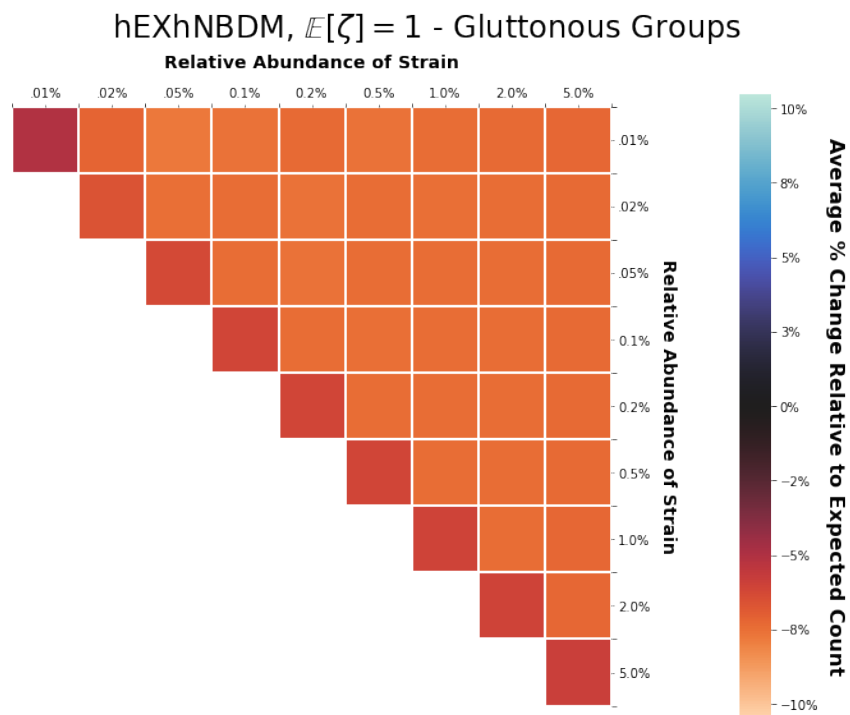
(d)

Figure 5.11: Gluttonous groups for high heterogeneity distributions.
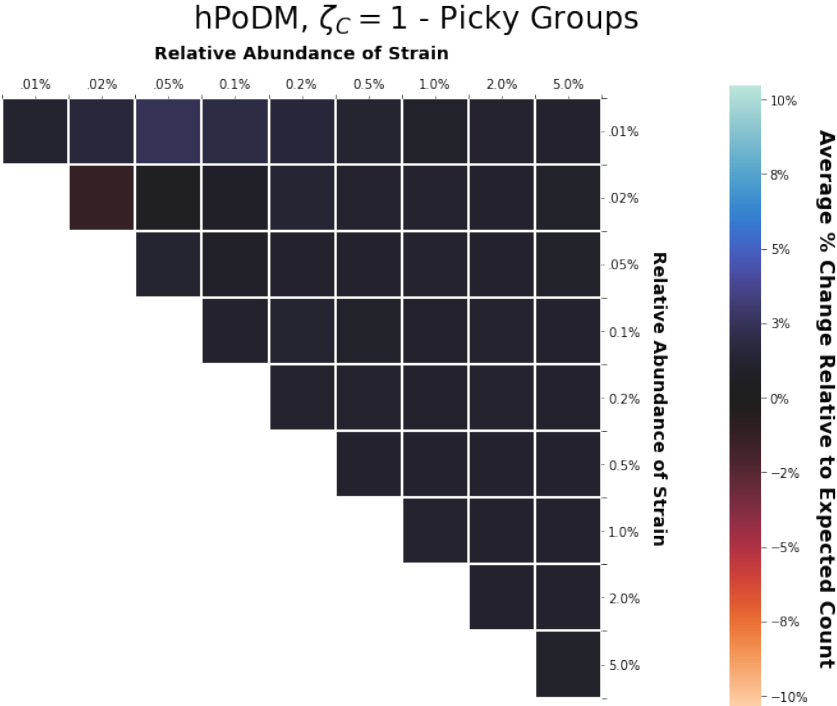
(a)
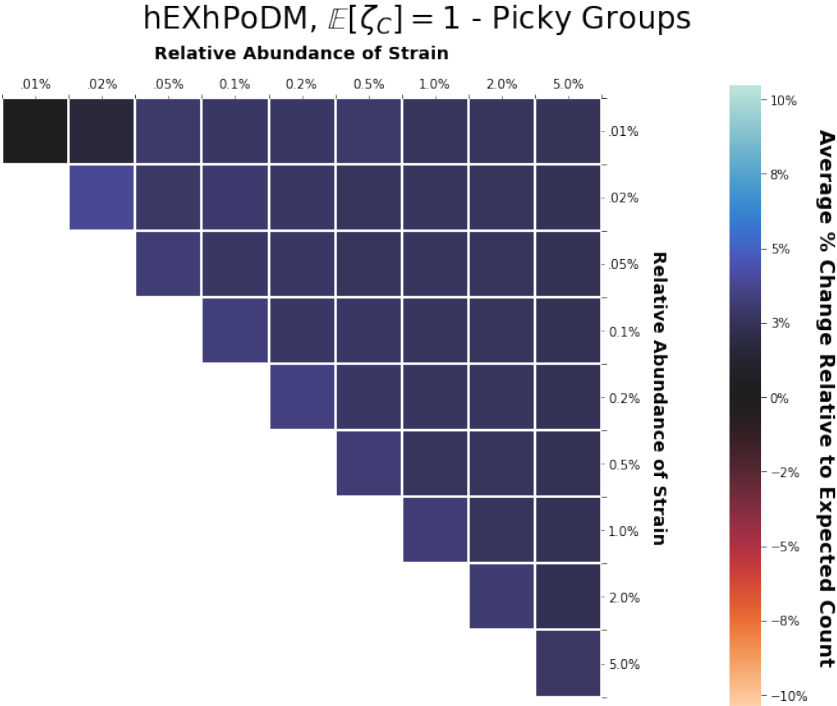
(b)

Figure 5.12: Picky groups for high heterogeneity distributions.
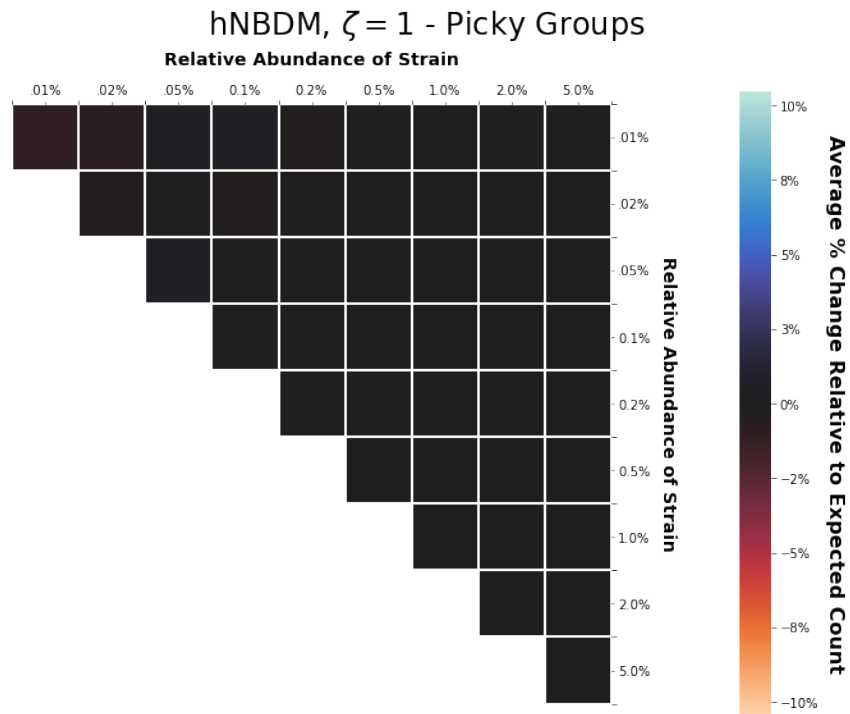
(c)
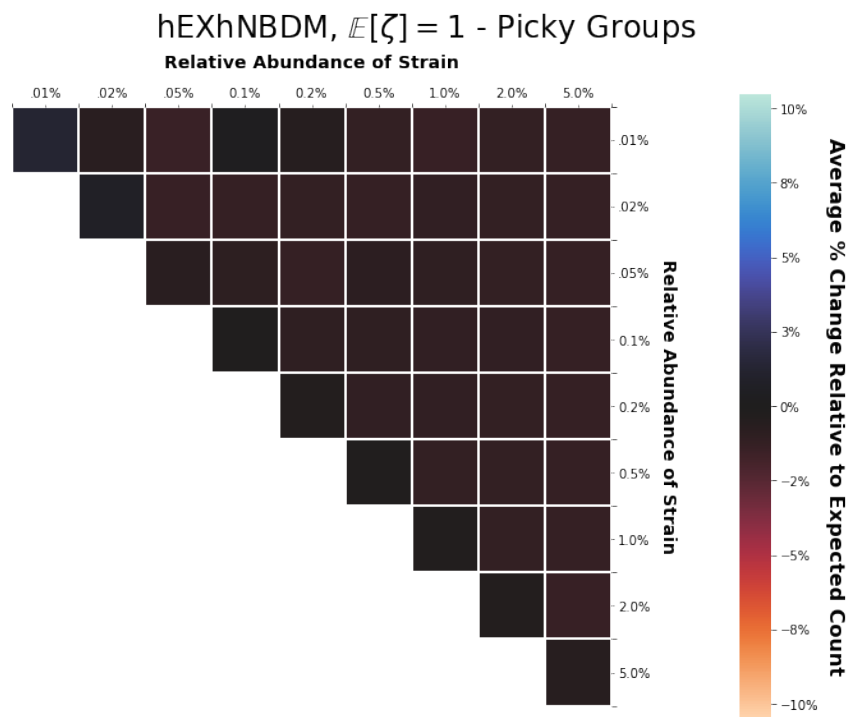
(d)

Figure 5.12: Picky groups for high heterogeneity distributions.

with two strains. While this trend is less obvious based only on figure 5.8b, it is clear based on the observed decrease in the number of picky treatments visible in figures 5.12c and 5.12d. Nevertheless it is also clear from figure 5.8b and from a comparison of figures 5.12d and 5.11d (and 5.12c and 5.11c) that the decrease in the number of droplets with two strains is not as pronounced (at least in relative terms) as the decrease in the number of droplets with three or more strains. Also in contrast to what happens in the absence of density heterogeneity, in the presence of density heterogeneity figures 5.8a and 5.8b show a marked increase in the number of empty droplets with zero strains as the compositional heterogeneity increases.

5.5 Discussion

Section 5.5.1 discusses the taxonomy of four groups of distributions identified in section 5.4. Section 5.5.2 discusses the distributions which are most likely to exhibit differences from hPoMu that are important in practice. Section 5.5.3 discusses how the interpretation of the results from section 5.4 depends fundamentally on our opinion of which droplets constitute "good data". Section 5.6 discusses how the results from section 5.4 indicate what the next step is in order to recover true interaction networks from MOREI data.

5.5.1 Overview of Groups of Distributions

The seven simulated distributions can be grouped into roughly four levels of assumption violation.

For the first two levels, the first consisting of hTPMH and the second consisting of hPoDM $\zeta_C = 100$ and hNBDM $\zeta = 100$, hPoMu is most likely an adequate working model. Cf. again section 4.5.2 from the previous chapter.

The third level consists of hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, which produce data that can be relatively easily distinguished from data produced by hPoMu. Cf. again section 4.5.3.

In all cases, the results using the contingency tables for the target parameters and resulting χ^2-divergences do nothing to change the qualitative conclusions that were made in the previous chapter. Similarity (or lack thereof) of these working models, as measured by log-likelihood ratios, corresponds to similarity (or lack thereof) of their contingency tables of target parameters.

The fourth level consists of hExhPoDM $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM $\mathbb{E}[\zeta] = 1$, which produce data so substantially different from that produced by hPoMu that it is almost impossible to *not* distinguish. It is very difficult to believe that hPoMu could possibly serve as an adequate working model for such data.

5.5.2 Very Different Distributions

The simulated hExhPoDM and hExhNBDM distributions technically do not belong to the ghNBDM family of distributions, since their concentration "parameters" do not have fixed

values, and instead are exponentially distributed. This distinction will be particularly relevant in sections 6.4.2.3 and 6.4.3.3.

Moreover, although the mean of these exponential distributions is 1, the behavior of the simulated hExhPoDM with $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM with $\mathbb{E}[\zeta] = 1$ distributions is substantially different from that of the hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$ distributions. This is most likely because concentration parameter values sampled from the Exponential(1) are $1 - \frac{1}{e} \approx 63.2\%$ of the time smaller than 1. (The median of the distribution is $\ln(2) \approx .693 < 1$.)

Thus the simulated hExhPoDM with $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM with $\mathbb{E}[\zeta] = 1$ can be assumed a priori to resemble more closely hPoDM and hNBDM distributions with values of ζ_C and ζ (substantially) smaller than 1. This corresponds to what we see in practice, because heterogeneities, as measured by over-dispersion, are inversely proportional to the concentration parameters. So for example $\zeta_C = \ln(2)$ would correspond to an hPoDM distribution with $\approx 44\%$ more heterogeneity than the hPoDM $\zeta_C = 1$ distribution.

Considering that the simulated hExhPoDM with $\mathbb{E}[\zeta_C] = 1$ and hExhNBDM with $\mathbb{E}[\zeta] = 1$ correspond $\approx 63.2\%$ to distributions with heterogeneities no less than those of hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, and $1 - \frac{1}{\sqrt{e}} \approx 39.3\%$ of the time to distributions with heterogeneities no less than *twice* those of hPoDM $\zeta_C = 1$ and hNBDM $\zeta = 1$, it is unsurprising that the aggregate/average heterogeneities observed for these distributions is much larger than those observed for either the hPoDM $\zeta_C = 1$ or hNBDM $\zeta = 1$ distributions.

Therefore fact that data generated by these two distributions is very unlikely to be adequately modeled by hPoMu further suggests that data generated by hPoDM or hNBDM distributions with ζ_C or ζ values not much smaller than 1 will also be inadequately modeled by hPoMu. It is a priori unclear which, if any, values of the concentration parameters smaller than 1 are realistic in practice.

5.5.3   The Effect of Heterogeneity on Data Throughput is Equivocal

The general effect that heterogeneities will have on the data throughput of MOREI is not straightforward. The answer depends on the particular combination of density heterogeneity and compositional heterogeneity present. Given both density and compositional heterogeneity, it is clear that the effect on data throughput is negative. Both the numbers of picky treatments and gluttonous treatments decrease.

In the presence of only compositional heterogeneity but no density heterogeneity, whether the effect on data throughput is positive depends primarily on our opinion of what is useful data. If our primary concern is avoiding "data starvation" and being able to say anything about the interactions of the rarest strains, then the strong decrease in the number of gluttonous treatment groups clearly means that the effect of heterogeneity is negative. On the other hand, if we consider only the data from the picky treatments useful, then in the absence of density heterogeneity the effect of compositional heterogeneity is slightly positive.

5.6 Conclusion

Findings and Contributions I confirmed that more severe failures of the assumptions from section 3.3.3 cause more severe discrepancies with the predictions derived from hPoMu. The nature of the effect depends on the chosen grouping (picky or gluttonous) of droplets defining the targeted estimands. This demonstrates that failures of the homogeneity assumptions of hPoMu potentially could be severe enough to require attention in practice.

Practical Implications The practical implications of this chapter remain limited by ignorance of which values of density and compositional concentration parameters are realistic in practice. If realistic values correspond to sufficiently small violations of the homogeneity assumptions, then this chapter actually provides evidence that such violations could be ignored in practice. Likewise, if realistic values correspond to sufficiently large violations of the homogeneity assumptions, then this chapter provides evidence of the opposite.

Next Steps and Open Questions What is also clear is that, if we wish to be able to predict generally what effects heterogeneities may have on the data throughput of MOREI in practice, we must be able to estimate both density heterogeneity and compositional heterogeneity from empirical data. This is the topic of chapter 6. Given such estimates from real-world data, we can get some sense for the amount of either type of heterogeneity one is likely to encounter in practice. That makes it possible to create more accurate simulations. More accurate simulations in turn better assess which inference methods are most capable of recovering true interaction networks from MOREI data.

Chapter 5.   Appendices

Appendix 5.A p-Values from χ^2 Approximation

Even for 15 million droplets, some categories corresponding to the rarest strains had very small expected counts (< 5). Thus the χ^2 approximation for the sampling distribution of the categorical divergence under the null distribution was unlikely to be accurate. In contrast, the Monte Carlo approximation to the null distribution is from 10^9 (one **b**illion) replicates. Therefore the p-values from the Monte Carlo approximation are more likely to be representative of the true p-values. For completeness, here I report the p-value distributions computed from the χ^2 approximation. While similar to the p-values from the Monte Carlo approximation, they appear to over-estimate the differences from the null.

Appendix 5.B   Additional Pearson Categorical Divergence Distributions

Here are two additional plots, one focusing on those distributions with compositional heterogeneity only, and the other on those distributions with both heterogeneities.
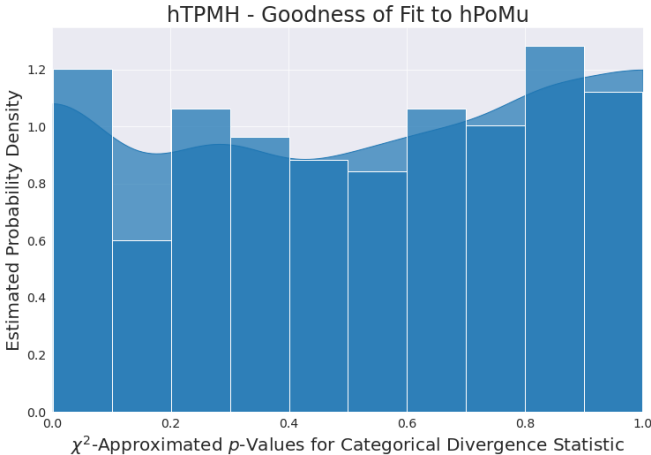
(a)
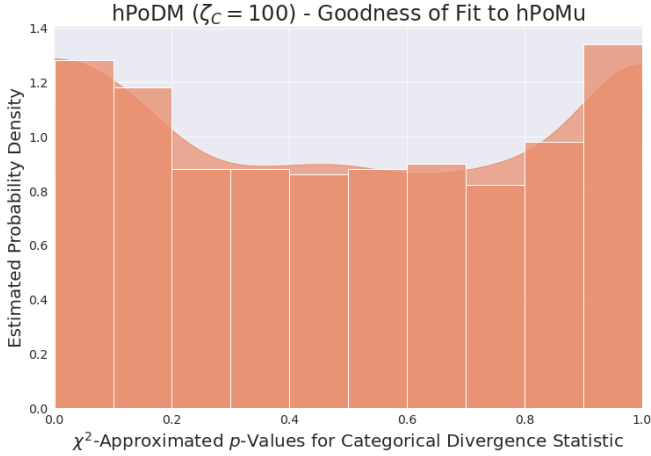
(b)
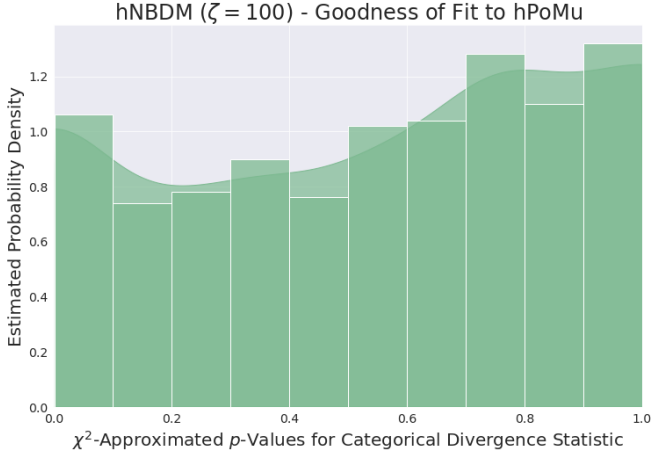
(c)

Figure 5.A.1: Approximate goodness of fit p-value distributions using χ^2 approximation.
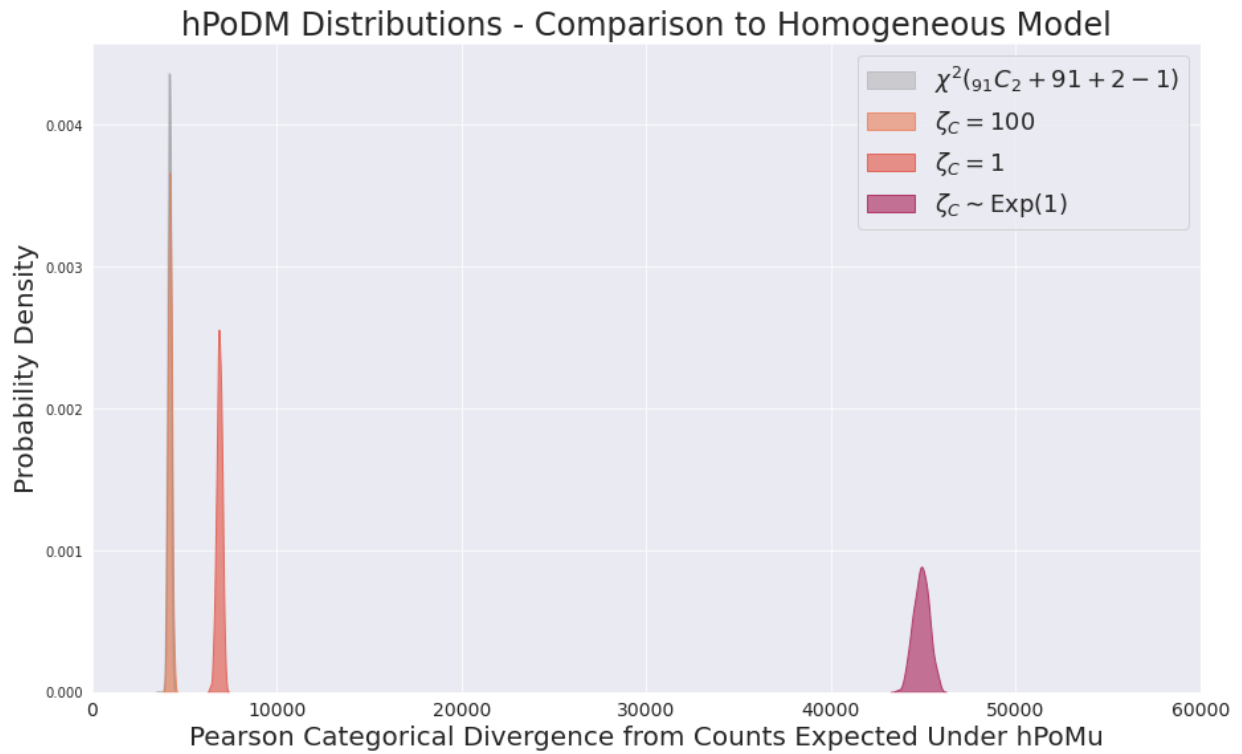
(a)
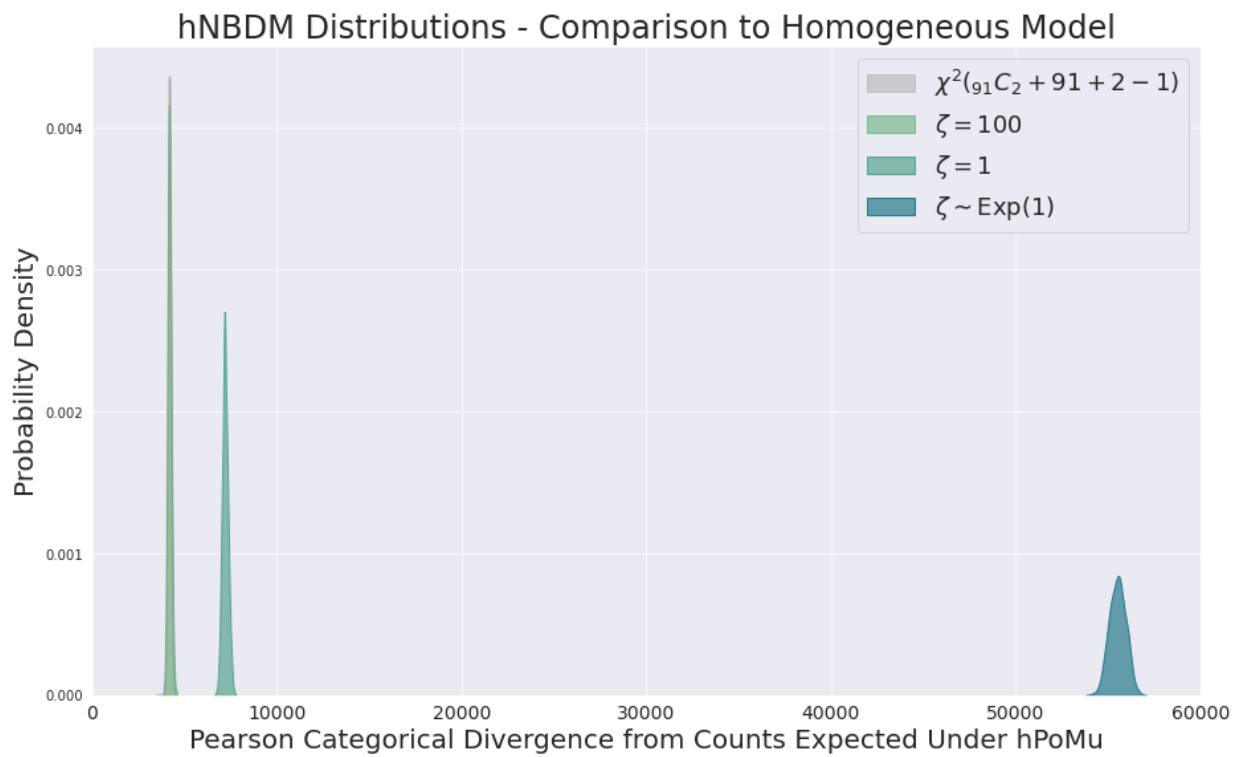
(b)

Appendix 5.C Multiple Representatives Problem

Another potential challenge arising from the randomness of the initial droplet formation process is the "multiple representatives problem", where within a given droplet there is initially a strain that is represented by multiple cells. Cf. figure 5.C.1. Even if we were more concerned about the effects randomness of the initial droplet formation process causes via the "multiple representatives problem" than the effects it has on data throughput, I argue we should still first focus on the effects it has on data throughput, for two reasons.

First, to the extent that the "multiple representatives problem" might create "noise" or "fluctuations" in the MOREI data, these effects will on average "cancel out" if the data throughput is great enough to ensure large sample sizes for both the treatment and control groups. Cf. section 7.8 (in particular section 7.8.2) for an example of when this is true. Second, even to the extent the "multiple representatives problem" might create systematic errors which do not "cancel out" given large enough sample sizes, these systematic errors should become easier to identify, and therefore also easier to account for, with greater data throughput. In either case, figuring out where we can and need to (or where we can't) account for effects caused by the "multiple representatives problem" begins with identifying the bottlenecks in data throughput that the particular form of randomness of the initial droplet formation process causes.
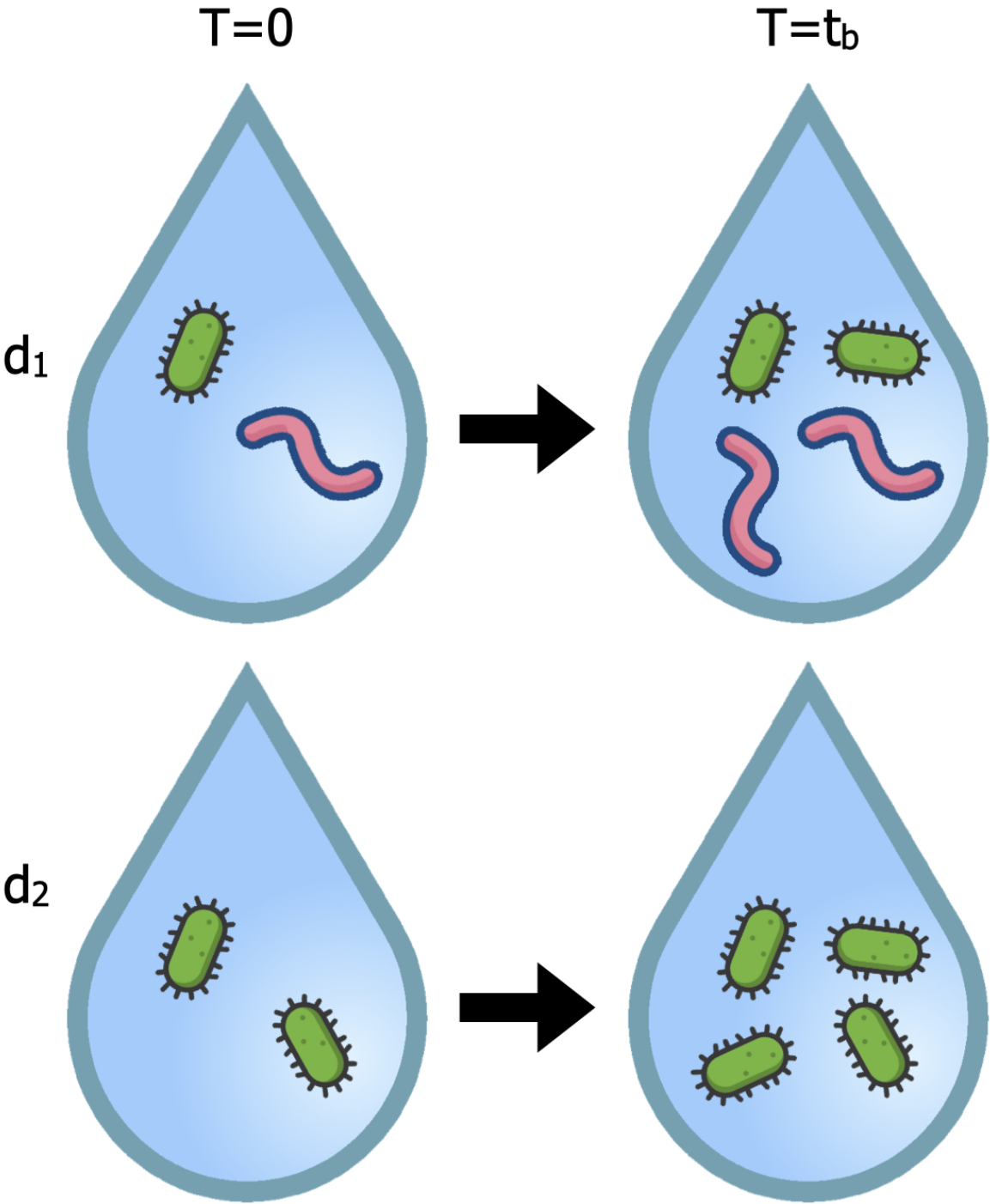
Figure 5.C.1: Multiple Representatives Problem. The droplets are only observed at time t_b. If we incorrectly assume both droplets began with the same number of green cells, then we incorrectly conclude that the pink cells halve the growth rate of the green cells. Yet the pink cells have no effect on the growth rate of the green cells. Cf. section 7.8.2.1.

Chapter 6

Estimators for Density and Compositional Heterogeneities

Herein I present both plugin and maximum likelihood estimators for failures of the assumptions from section 3.3.3. (The failures correspond to the so-called "concentration" parameters defined in sections 3.6, 3.7, and 3.8.) See section 6.2. I show that failures of the assumptions from section 3.3.3 are understandable non-parametrically, despite the parametric definitions implicit from sections 3.6, 3.7, and 3.8. See sections 6.2.1.2, 6.2.3.5, 6.2.3.6, and 6.5.2 for arguments to this effect. In particular, the estimators manage to still behave reasonably even under (mild) model misspecification. See sections 6.4.2.3 and 6.4.3.3.

Section 6.1 provides context. Section 6.2 goes over technical details related to defining the estimators. Section 6.3 describes some implementation details of the analyses. Section 6.4 explains the results of the analyses. Section 6.5 interprets the results and explains how they are relevant to modelling the initial formation of droplets. Finally section 6.6 concludes.

Chapter 6 Contents

6.1 Background and Significance

Broader field   The problem for this chapter belongs to the broader field of point estimation. See for example any of [LC98], [Kee10], [Ser80], or [Cra46] for general introductions to this field. Cf. also the discussion earlier from the introduction to Part II. This chapter considers the plug-in (also known as "method of moments") and maximum likelihood approaches to point estimation in particular. See [LC98, Example 4.7, p.456] for the definition and basic theory of plug-in estimators, which recommends [Cra46, Section 33.1] and [Ser80, Section 4.3.1] for further reading. Likewise, see [Kee10, Section 8.4] or [LC98, Section 6.3, p. 444] for the definition and basic theory of maximum likelihood estimators, the latter of which also recommends [Str85, Sections 64.4 and 84.2] and [Efr82] for further reading.

Specific problem   To account for effects of the randomness of the initial droplet formation process we need to make realistic predictions about what these effects will be. Previous chapters showed that different statistical working models for the initial droplet formation process can make substantially different predictions. Thus, to make realistic predictions, we need to evaluate which of these statistical working models is most realistic. However, the more general statistical working models have additional parameters compared to hPoMu, and unlike the parameters in hPoMu these additional parameters are not obviously connected to the configuration of the experiment. Hence, to evaluate which of these statistical working models is most realistic, we need to estimate these additional parameters from real data.

Therefore the specific problem for this chapter is to demonstrate that these additional parameters are identifiable from the observed data generated by unincubated droplets.

Using the hierarchical count-categorical framework (cf. section 3.2), we can separate the estimation problems for the density concentration parameters and the compositional concentration parameters. Doing so, estimating the density concentration parameter for the ghNBDM family (cf. again section 3.8) is the same as estimating the density concentration parameter for the negative binomial distribution, and estimating the compositional concentration parameter for the ghNBDM family is the same as estimating the compositional concentration parameter the Dirichlet-Multinomial.

Because of the use of spikein genes (cf. again section 1.4.3), we can assume that the observed data generated by unincubated droplets is (roughly) integer count data. Plentiful previous work already exists about how to estimate the density concentration parameter (or "equivalent" parameters) of the negative binomial distribution from (univariate) count data. See for example any of [Ans50], [SW62], [WFY84], [Bow84], [RP85], [WFY86], [CP89], [Pie90], [AEE92], [Nak97], [LS07], or [YHV13]. Less previous work seems to exist regarding estimation of the compositional concentration parameter of the Dirichlet-Multinomial distribution from multivariate count data. See e.g. [Skl14], [YS14], [Bou20], or [HCSB20].

Particular approach   I derive plugin and maximum likelihood estimators for these additional parameters and demonstrate their effectiveness in practice. Plugin and maximum

likelihood estimation are probably the most widely employed approaches for point estimation. Therefore employing them is a useful starting point for this problem.

The maximum likelihood estimators are more efficient when the models are correctly specified. (This is unsurprising in light of general results about the asymptotic optimality of maximum likelihood estimation; cf. [LC98, Chapter 6].) Nevertheless the difference in efficiency is small for the density concentration parameter, confirming previous reports [Bow84][Nak97] of the somewhat pathological nature of that particular estimation problem.

Previous work [CP89] [Nak97] has used maximum pseudo-likelihood or maximum quasi-likelihood approaches for the density concentration of the negative binomial distribution. However, such approaches are not as standard as plugin ("method of moments") or maximum likelihood estimation. Moreover, there does not appear to be any consensus in the literature that those approaches actually perform better. Similar reasoning also applies for choosing not to use the methods of neither [WFY84] (density concentration of negative binomial) nor [YS14] (compositional concentration of Dirichlet-Multinomial).

This chapter also attempts to isolate non-parametric definitions which correspond to the concentration parameters in the special cases of these parametric working models. Being able to characterize these as non-parametric estimands (more precisely pathwise differentiable functionals definable on non-parametric statistical models) will facilitate future work applying targeted minimum-loss based estimation (TMLE) [vdLR11] to this problem.

6.2 Inference of Concentration Parameters: ML and Plugin Approaches

Using the hierarchical count-categorical framework defined in section 3.2, I treat the estimation of the density concentration ζ_D and compositional concentration ζ_C parameters of the ghNBDM family (cf. section 3.8) as two separate estimation problems. The first seeks to estimate the density concentration ζ_D from the empirical count distribution. The second seeks to estimate the compositional concentration ζ_C from the empirical categorical distributions[1].

For both estimation problems, I employ two strategies: (1) plugin ("method of moments") estimation, and (2) maximum likelihood ("ML") estimation. Both approaches give statistically consistent and asymptotically normal estimators. The maximum likelihood strategy has the advantage of being more efficient and asymptotically optimal (in the sense of satisfying the Cramer-Rao lower bound). The plugin strategy has the advantage of being computationally simpler (and thus also easier for non-specialists to implement). I argue in sections 6.2.1.2, 6.2.3.5, and 6.5.2 that the plugin strategy also has the advantage of being "semi-parametric"[2], at least in the sense that its estimand is still straightforward to interpret even

---

[1]Distribution*s* (plural) because again technically speaking there is a different categorical distribution for each possible value of the count distribution.

[2] Section 4 of [Nak97] argues (at least for density concentration) that the plugin estimator is also "semi-parametric" when compared to the MLE due to having less bias and standard error. This is because the plugin estimator retains consistency (and supposedly also efficiency) for any distribution with the same

under model misspecification, although this is debatable.

Sections 6.2.1 and 6.2.2 discuss the estimation problem corresponding to the count distributions and the density concentration parameter ζ_D. This is formulated as a parametric estimation problem involving the negative binomial distribution. Section 6.2.1 provides details of the plugin strategy, while section 6.2.2 provides details of the ML strategy.

Sections 6.2.3 and 6.2.4 discuss the estimation problem corresponding to the categorical distributions and the compositional concentration parameter ζ_C. This is formulated as a univariate parametric estimation problem involving the Dirichlet-Multinomial distribution. The remaining nuisance parameters (the strain[3] frequencies) can be assumed to be either already known or estimated separately (cf. section 6.A). Section 6.2.3 provides details of the plugin strategy, while section 6.2.4 provides details of the ML strategy.

Notation

The notation $\mathbb{I}\{A\}$ denotes the indicator function for the event A.

Given any estimand which is a conditional expectation, i.e. of the form $\mathbb{E}[X \mid A]$ for some random variable X and some event A in the underlying sigma-algebra, the corresponding empirical (arithmetic) mean is denoted

$$\hat{\boldsymbol{\mu}}[X \mid A] := \frac{\sum_{d \in [D]} X_d \mathbb{I}\{A\}}{\sum_{d \in [D]} \mathbb{I}\{A\}} \, . \tag{6.2.1}$$

In particular the unconditional empirical mean is

$$\hat{\boldsymbol{\mu}}[X] := \frac{1}{D} \sum_{d \in [D]} X_d \, . \tag{6.2.2}$$

Analogously a conditional variance estimand $\mathrm{Var}[X \mid A]$ has a corresponding empirical conditional variance:

$$\hat{\mathrm{var}}[X \mid A] := \frac{\sum_{d \in [D]} (X_d - \hat{\boldsymbol{\mu}}[X \mid A])^2 \mathbb{I}\{A\}}{\sum_{d \in [D]} \mathbb{I}\{A\}} \, , \tag{6.2.3}$$

---

overdispersion relative to the Poisson while the ML estimator does not. Cf. sections 3.4.4.1 and 3.7.1.2 of this work with formula (2.3) and section 4 of [Nak97]. This work does not investigate the estimators' behavior under model misspecification enough to evaluate this robustness claim. The true values of the effective density concentration (cf. section 6.2.1.2) and the effective compositional concentration (cf. section 6.2.3.5) are never computed for the misspecifed models. Moreover the misspecified models (hExhPoDM and hExhNBDM) are only slight modifications of the ghNBDM family. An investigation of the estimators' behavior under model misspecification thorough enough to make a legitimate and informed judgment about this issue is a worthwhile subject for future work.

[3]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.

and an unconditional variance estimand $\mathrm{Var}[X]$ has the empirical counterpart:

$$\mathrm{v\hat{a}r}[X] := \frac{1}{D} \sum_{d \in [D]} (X_d - \hat{\boldsymbol{\mu}}[X])^2 \,. \tag{6.2.4}$$

Generalizing the above, for a conditional *covariance* estimand $\mathrm{Cov}[X, Y \mid A]$ its empirical counterpart is denoted and defined as

$$\mathrm{c\hat{o}v}[X, Y \mid A] := \frac{\sum_{d \in [D]} (X_d - \hat{\boldsymbol{\mu}}[X \mid A])(Y_d - \hat{\boldsymbol{\mu}}[Y \mid A])\mathbb{I}\{A\}}{\sum_{d \in [D]} \mathbb{I}\{A\}} \,, \tag{6.2.5}$$

and likewise for an unconditional covariance estimand $\mathrm{Cov}[X, Y]$ its empirical counterpart is denoted and defined as

$$\mathrm{c\hat{o}v}[X, Y] := \frac{1}{D} \sum_{d \in [D]} (X_d - \hat{\boldsymbol{\mu}}[X])(Y_d - \hat{\boldsymbol{\mu}}[Y]) \,. \tag{6.2.6}$$

6.2.1 Plugin Estimator for Density Concentration

Section 6.2.1.1 derives the formula for the plugin estimator. Section 6.2.1.2 discusses how the plugin strategy implies an estimand that is defined for count distributions more general than the negative binomial distribution alone.

6.2.1.1 Derivation of Plugin Estimator for Density Concentration

Recall that when $N(0)$ is negative binomial distributed[4], the variance equals

$$\mathrm{Var}[N(0)] = \lambda \left(1 + \frac{\lambda}{\zeta_D S}\right) \,, \tag{6.2.7}$$

while the variance that would have been expected under the Poisson distribution equals

$$\mathbb{E}[N(0)] = \lambda \,. \tag{6.2.8}$$

Thus when $N(0)$ is negative binomial distributed its over-dispersion relative to the Poisson distribution equals

$$\frac{\mathrm{Var}[N(0)] - \mathbb{E}[N(0)]}{\mathbb{E}[N(0)]} = \frac{\lambda}{\zeta_D S} = \frac{\mathbb{E}[N(0)]}{\zeta_D S} \,. \tag{6.2.9}$$

Solving for ζ_D leads to

$$\zeta_D = \frac{1}{S} \cdot \frac{(\mathbb{E}[N(0)])^2}{\mathrm{Var}[N(0)] - \mathbb{E}[N(0)]} \,. \tag{6.2.10}$$

The above relationship (6.2.10) motivates the following plugin estimator for ζ_D:

$$\hat{\zeta_D} := \begin{cases} \infty & \mathrm{v\hat{a}r}[N(0)] \leq \hat{\boldsymbol{\mu}}[N(0)] \\ \frac{1}{S} \frac{(\hat{\boldsymbol{\mu}}[N(0)])^2}{\mathrm{v\hat{a}r}[N(0)] - \hat{\boldsymbol{\mu}}[N(0)]} & \text{else} \end{cases} \,. \tag{6.2.11}$$

Cf. the derivation of the analogous equation (3.10) in [Ans50].

---

[4]Using the parametrization for the negative binomial described in section 3.7.1.1.

6.2.1.2 Effective Density Concentration

If one defines for any possible distribution of $N(0)$ with finite variance (e.g. including the Poisson) the "effective density concentration" as

$$\zeta_{Deff} := \begin{cases} \infty & \text{Var}[N(0)] \leq \mathbb{E}[N(0)] \\ \frac{1}{S} \frac{(\mathbb{E}[N(0)])^2}{\text{Var}[N(0)] - \mathbb{E}[N(0)]} & \text{else} \end{cases} , \tag{6.2.12}$$

then the above plugin estimator (6.2.11) is always consistent for the "effective density concentration" ζ_{Deff}. Consistency follows from the weak law of large numbers and the continuous mapping theorem for convergence in probability.

The effective density concentration ζ_{Deff} being infinite merely flags that the distribution of $N(0)$ is not over-dispersed relative to the Poisson. An obvious limitation of this is that provides no quantification of the extent to which the distribution might be *under*-dispersed relative to the Poisson. Such under-dispersed distributions are not necessarily "exotic", consider for example the distribution of $Y = \frac{1}{2}X$ when $X \sim$ Pois. Then $\text{Var}[Y] = \left(\frac{1}{2}\right)^2 \text{Var}[X] = \frac{1}{4}\mathbb{E}[X] < \frac{1}{2}\mathbb{E}[X] = \mathbb{E}[Y]$. This of course works also for pX for any $0 < p < 1$. Considering the possible ramifications of under-dispersion is left to future work.

When the distribution of $N(0)$ *is* over-dispersed relative to the Poisson, the effective density concentration measures the over-dispersion on a normalized and inverted scale. Because the effective density concentration ζ_{Deff} equals the density concentration parameter ζ_D whenever $N(0)$ is negative binomial distributed, the estimand ζ_{Deff} may be considered a generalization of the parameter ζ_D to other distributions. This resembles the argument made in [Nak97, section 4] that the negative binomial plugin estimator[5] is only "semi-parametric".

6.2.2 Maximum Likelihood Estimator for Density Concentration

This section explains how to use the maximum likelihood strategy to estimate the density concentration ζ_D. Section 6.2.2.1 computes the relevant likelihood and log likelihood for the count distributions, while section 6.2.2.2 computes the score. Section 6.2.2.3 uses those results to derive the maximum likelihood estimators for the negative binomial distribution. Sections 6.2.2.4 and 6.2.2.5 survey observations from previous literature about this estimator.

6.2.2.1 Negative Binomial (Log) Likelihood

Starting from equation (3.7.1),

- if we assume for mathematical convenience that the count distributions of all droplets are mutually independent,

---

[5][Nak97] considers a reparameterized version of the negative binomial distribution such that the corresponding plugin estimator is the reciprocal of that considered herein. This would correspond to an estimand that is the reciprocal of the effective density concentration.

- using the identity[6] $\frac{\Gamma(x+M)}{\Gamma(x)} = \prod_{m=0}^{M-1}(x+m)$ for all positive integers M,

then the following is the full likelihood for a batch of D droplets:

$$\mathcal{L}(\lambda, \zeta_D) = \prod_{d \in [D]} \left[\left(\prod_{\nu=0}^{n_d-1} (\zeta_D S + \nu) \right) \cdot \left(1 + \frac{\lambda}{\zeta_D S} \right)^{-\zeta_D S} \cdot \left(\frac{1}{\zeta_D S + \lambda} \right)^{n_d} \cdot \frac{\lambda^{n_d}}{n_d!} \right]. \qquad (6.2.13)$$

(Recall that $[D] := \{1, \ldots, D\}$.) Thus this is the log likelihood:

$$\ell(\lambda, \zeta_D) :=$$

$$\log(\mathcal{L}(\lambda, \zeta_D)) = \sum_{d \in [D]} \left[\sum_{\nu=0}^{n_d-1} \log(\zeta_D S + \nu) \right] - D\zeta_D S \log\left(1 + \frac{\lambda}{\zeta_D S} \right) +$$

$$D\overline{n} \log\left(\frac{1}{\zeta_D S + \lambda} \right) + D\overline{n} \log(\lambda) - \sum_{d \in [D]} \log(n_d!), \qquad (6.2.14)$$

where I have defined $\overline{n} := \frac{1}{D} \sum_{d \in [D]} n_d$. (Each variable n_d corresponds to a possible value of the random variable $N_d(0)$, the number of cells in droplet d.)

6.2.2.2 Negative Binomial Score

It follows that the score with respect to λ is

$$\frac{\partial \ell(\lambda, \zeta_D)}{\partial \lambda} = -\frac{D\zeta_D S}{\zeta_D S + \lambda} + \frac{\overline{n}}{\lambda} \cdot \left(\frac{D\zeta_D S}{\zeta_D S + \lambda} \right). \qquad (6.2.15)$$

Similarly, from (very) tedious calculus we get the following expression for the score with respect to ζ_D:

$$\frac{\partial \ell(\lambda, \zeta_D)}{\partial \zeta_D} = \sum_{d \in [D]} \left[\sum_{\nu=0}^{n_d-1} \frac{S}{\zeta_D S + \nu} \right] + \frac{DS\lambda}{\zeta_D S + \lambda} -$$

$$DS \log\left(1 + \frac{\lambda}{\zeta_D S} \right) - \frac{DS\overline{n}}{\zeta_D S + \lambda}. \qquad (6.2.16)$$

6.2.2.3 Negative Binomial Maximum Likelihood Estimator

Setting $\frac{\partial \ell(\lambda, \zeta_D)}{\partial \lambda} = 0$, equation (6.2.15) gives us that the maximum likelihood estimator $\hat{\lambda}$ for λ is $\hat{\lambda} = \overline{n}$. This means that the plugin and maximum likelihood estimators for λ coincide.

---

[6]This follows from how Γ interpolates the factorials, i.e. $\Gamma(x+1) = x\Gamma(x)$.

Setting $\frac{\partial \ell(\lambda, \zeta_D)}{\partial \zeta_D} = 0$, and substituting into (6.2.16) the relationship $\lambda = \hat{\lambda} = \overline{n}$, we get that, if the maximum likelihood estimator $\hat{\zeta}_D$ of ζ_D exists, then it must be a solution of the following equation:

$$D \log \left(1 + \frac{\overline{n}}{\hat{\zeta}_D S}\right) = \sum_{d \in [D]} \left[\sum_{\nu=0}^{n_d - 1} \frac{1}{\hat{\zeta}_D S + \nu}\right] . \tag{6.2.17}$$

Rearranging terms for the purposes of computational efficiency, the equation (6.2.17) above can be rewritten

$$D \log \left(1 + \frac{\overline{n}}{\hat{\zeta}_D S}\right) = \sum_{m=1}^{M} \frac{D_{\geq m}}{\hat{\zeta}_D S + (m - 1)} , \tag{6.2.18}$$

where $M := \max_{d \in [D]} N_d(0)$ and $D_{\geq m} := |\{d : N_d(0) \geq m\}|$.

6.2.2.4 Existence, Uniqueness, and Finiteness Conditions

It is believed[Ans50][WFY84][WFY86] that a solution to the above equation (6.2.18) exists if and only if $\hat{\text{var}}[N(0)] > \hat{\boldsymbol{\mu}}[N(0)]$ (i.e. if and only if the empirical distribution is overdispersed with respect to the Poisson), and that if a solution exists it is always unique. A purported[7] proof of these claims is given in [AEE92].

Note how the above purported conditions for the existence and uniqueness of the negative binomial MLE are the same as the conditions for the finiteness of the plugin estimator. When the MLE does not exist, it is because the likelihood is increasing but has no maximum. Therefore in those situations it makes sense to define the MLE to be infinite (cf. [Ans50]). Thus the plugin estimator and MLE agree on when the estimate should be finite or infinite. Just like for the plugin estimator, the MLE being infinite has the interpretation that the Poisson distribution fits the data better than any negative binomial distribution.

6.2.2.5 Estimation Difficulties

The two parameter (i.e. both λ and ζ_D unknown) estimation problem for the negative binomial may be intrinsically difficult due to there being no complete sufficient statistic, corresponding to how the two parameter negative binomial family is not an exponential family [WFY86].

At least for (relatively) small samples [Bow84][CP89][Pie90][LS07], previous work has raised the concern that the MLE has unusually large bias [SW62] [WFY84] [Bow84] [Pie90] [LS07]. However at least one source [Pie90] conjectures that the MLE becomes a more viable option for large sample sizes.

The maximum likelihood estimation of ζ_D is believed to be particularly difficult for "small" values of λ and large values of ζ_D [SW62][WFY84][Bow84][LS07]. This may be due to the contours of the likelihood function being relatively "flat" or slowly increasing

---

[7]The proof is probably correct, but I have inspected it only cursorily and not in detail. Therefore I feel I cannot legitimately attest to its validity firsthand.

with respect to ζ_D[WFY86]. However, other work [CP89] has found that the small λ and large ζ_D regime is also the most difficult for other kinds of estimators of ζ_D. [Bow84] found that $\mathrm{v\hat{a}r}[N(0)] \leq \hat{\boldsymbol{\mu}}[N(0)]$ occurred most often in the small λ, large ζ_D regime, a condition for which both the plugin and ML estimators return infinite estimates (cf. section 6.2.2.4).

Because of this, some authors have suggested that the MLE does not perform much better than the corresponding plugin estimator (despite having lower variance) [Ans50] [SW62] [WFY84] and thus that the plugin estimator should be preferred due to its greater computational simplicity [Ans50] [WFY84] [YHV13].

Other authors have suggested instead reparameterizing the negative binomial in terms of $\frac{1}{\zeta_D S}$ [RP85][CP89][Pie90][Nak97][LS07]. This makes sense because the difficulties occur mostly for large values of ζ_D for which the resulting distributions are all very similar to the corresponding Poisson[Bow84][LS07] and because this removes the need for infinite estimates [CP89][LS07]. [RP85] argue that reparameterizing the negative binomial using $\frac{1}{\zeta_D S}$ allows a continuous transition between the negative binomial, which is over-dispersed relative to the Poisson, and the ("positive") binomial, which is under-dispersed relative to the Poisson. Nevertheless even the reparameterized MLE was not always necessarily found to behave much better than the reparameterized plugin estimator [Pie90].

Exact formulae for the asymptotic variance of both the plugin and ML estimators were given in [Ans50], and exact formulae for the asymptotic bias of both the plugin and ML estimators were given in [SW62]. ([Nak97] gives exact formulae for the asymptotic variances of the $\frac{1}{\zeta_D S}$ reparameterized versions of the plugin and ML estimators.) Because the asymptotic formulae of the plugin and ML estimators have "the same general structure"[SW62] for both the bias and variance[8], the larger sample sizes typical for this problem may lead to the performance of the MLE and plugin estimator being comparable.

6.2.3 Plugin Estimator for Compositional Concentration

Section 6.2.3.1 outlines the big picture ideas behind the plugin estimator's somewhat complicated derivation. Section 6.2.3.2 provides the details behind most of the derivation. Section 6.2.3.3 discusses a possible ambiguity and choices for resolving it. Section 6.2.3.4 gives the formal definition of the final plugin estimator used for all later analyses. Section 6.2.3.5 discusses how the plugin strategy implies an estimand that is defined for categorical distributions more general than the Dirichlet-Multinomial distribution alone.

6.2.3.1 Overview of Derivation of Plugin Estimator for Compositional Concentration

The general idea is as follows:

- For each number of cells $n \geq 2$ and each tuple of strains (s_1, s_2) get a plugin estimator for $\frac{1}{1+\zeta_C}$.

---

[8]Cf. Fig.3 of [SW62] and equations 3.6 and 3.11 of [Ans50].

- Take a weighted average to get a single plugin estimator for $\frac{1}{1+\zeta_C}$.

- Solve to get a plugin estimator for ζ_C.

The motivation for averaging before solving for ζ_C in the expression $\frac{1}{1+\zeta_C S}$ is because I anticipate the former operation to be "better behaved" than the reciprocals involved in the latter. Solving after averaging makes it possible to take reciprocals only once. This hopefully leads to a plugin estimator for ζ_C "better behaved" than plugin estimators derived by solving (multiple times) first and then averaging. Nevertheless the latter in principle is possible too.

Note that when $n = 1$, the Dirichlet-Multinomial distribution reduces completely to the corresponding Multinomial ("Multinoulli") distribution and any compositional concentration parameter is completely unidentifiable. (This corresponds to how $n - 1 = 0$ in the overdispersion formula, but follows ultimately from the probability mass function of the Dirichlet-Multinomial distribution.)

6.2.3.2 Solving for Compositional Concentration Parameter

Whenever $\mathbb{E}[\vec{\mathbf{N}}(0) \mid N(0)]$ is Dirichlet-Multinomial distributed, for all fixed $n \geq 2$ one has, it follows from (3.B.2) that for any strain s:

$$\frac{1}{1 + \zeta_C S} = \frac{\mathrm{Var}[N^{(s)}(0) \mid N(0) = n] - nf^{(s)}(1 - f^{(s)})}{n(n-1)f^{(s)}(1 - f^{(s)})} . \tag{6.2.19}$$

Similarly, it follows from (3.B.4) that for any pair of strains s_1, s_2:

$$\frac{1}{1 + \zeta_C S} = \frac{-\mathrm{Cov}[N^{(s_1)}(0), N^{(s_2)}(0) \mid N(0) = n] - nf^{(s_1)}f^{(s_2)}}{n(n-1)f^{(s_1)}f^{(s_2)}} . \tag{6.2.20}$$

Therefore, for any convex combination c_σ, c_{σ_1, σ_2} of the right hand sides of (6.2.19) and (6.2.20) one has

$$\frac{1}{1 + \zeta_C S} = \frac{1}{n(n-1)} \left[\sum_{\sigma \in [S]} c_\sigma \left[\frac{\mathrm{Var}[N^{(\sigma)}(0)] - nf^{(\sigma)}(1 - f^{(\sigma)})}{f^{(\sigma)}(1 - f^{(\sigma)})} \right] \right.$$
$$\left. + \sum_{\sigma_1 \in [S]} \sum_{\sigma_2 \neq \sigma_1} c_{\sigma_1, \sigma_2} \left[\frac{-\mathrm{Cov}[N^{(\sigma_1)}(0), N^{(\sigma_2)}(0)] - nf^{(\sigma_1)}f^{(\sigma_2)}}{f^{(\sigma_1)}f^{(\sigma_2)}} \right] \right] . \tag{6.2.21}$$

Therefore one has further that

$$\frac{1}{1 + \zeta_C S} = \sum_{n \geq 2} \left(\left[\frac{1}{n(n-1)} \left[\sum_{\sigma \in [S]} c_\sigma \left[\frac{\mathrm{Var}[N^{(\sigma)}(0) \mid N(0) = n] - nf^{(\sigma)}(1 - f^{(\sigma)})}{f^{(\sigma)}(1 - f^{(\sigma)})} \right] \right. \right. \right.$$
$$\left. \left. \left. + \sum_{\sigma_1 \in [S]} \sum_{\sigma_2 \neq \sigma_1} c_{\sigma_1, \sigma_2} \left[\frac{-\mathrm{Cov}[N^{(\sigma_1)}(0), N^{(\sigma_2)}(0) \mid N(0) = n] - nf^{(\sigma_1)}f^{(\sigma_2)}}{f^{(\sigma_1)}f^{(\sigma_2)}} \right] \right] \right] \right. \tag{6.2.22}$$
$$\left. \cdot \mathbb{P}(N(0) = n \mid N(0) \geq 2) \right) ,$$

allowing one to solve for the compositional concentration parameter

$$\zeta_C = \frac{1}{S}\left(\left(\sum_{n\geq 2}\left(\left[\frac{1}{n(n-1)}\left[\sum_{\sigma\in[S]}c_\sigma\left[\frac{\mathrm{Var}[N^{(\sigma)}(0)\mid N(0)=n]-nf^{(\sigma)}(1-f^{(\sigma)})}{f^{(\sigma)}(1-f^{(\sigma)})}\right]\right.\right.\right.\right.\right.$$
$$\left.\left.\left.+\sum_{\sigma_1\in[S]}\sum_{\sigma_2\neq\sigma_1}c_{\sigma_1,\sigma_2}\left[\frac{-\mathrm{Cov}[N^{(\sigma_1)}(0),N^{(\sigma_2)}(0)\mid N(0)=n]-nf^{(\sigma_1)}f^{(\sigma_2)}}{f^{(\sigma_1)}f^{(\sigma_2)}}\right]\right]\right]$$
$$\left.\left.\left.\cdot\,\mathbb{P}(N(0)=n\mid N(0)\geq 2)\right)\right)^{-1}-1\right).$$

(6.2.23)

6.2.3.3   Choice of Convex Coefficients

Choosing $c_\sigma, c_{\sigma_1,\sigma_2} = \frac{1}{S^2}$ for all strains (and pairs thereof) corresponds to the L_2 projection of the covariance matrices onto the vector space of matrices whose entries all equal a common value. Thus the corresponding plugin estimator might initially appear attractive as a "least-squares" estimator.

However, the corresponding plugin estimator can have an extremely slow rate of convergence, e.g. when there are many "rare" strains. Preliminary investigations (data not shown) found that the corresponding plugin estimator effectively failed to converge towards the true value at all even after sampling 15 million droplets.

A more pragmatic strategy is to choose the weights $c_\sigma, c_{\sigma_1,\sigma_2}$ such that estimates with the most data behind them also receive the most weight. This means assigning higher weights to (pairs of) more common strains. Using the choices $c_\sigma = (f^{(\sigma)})^2$ and $c_{\sigma_1,\sigma_2} = f^{(\sigma_1)}f^{(\sigma_2)}$, the resulting formula for ζ_C is

$$\zeta_C = \frac{1}{S}\left(\left(\sum_{n\geq 2}\left(\left[\frac{1}{n(n-1)}\left[\sum_{\sigma\in[S]}(f^{(\sigma)})^2\left[\frac{\mathrm{Var}[N^{(\sigma)}(0)\mid N(0)=n]-nf^{(\sigma)}(1-f^{(\sigma)})}{f^{(\sigma)}(1-f^{(\sigma)})}\right]\right.\right.\right.\right.\right.$$
$$\left.\left.\left.+\sum_{\sigma_1\in[S]}\sum_{\sigma_2\neq\sigma_1}f^{(\sigma_1)}f^{(\sigma_2)}\left[\frac{-\mathrm{Cov}[N^{(\sigma_1)}(0),N^{(\sigma_2)}(0)\mid N(0)=n]-nf^{(\sigma_1)}f^{(\sigma_2)}}{f^{(\sigma_1)}f^{(\sigma_2)}}\right]\right]\right]$$
$$\left.\left.\left.\cdot\,\mathbb{P}(N(0)=n\mid N(0)\geq 2)\right)\right)^{-1}-1\right).$$

(6.2.24)

The corresponding plugin estimator converges much faster to the true value ζ_C.

6.2.3.4   Plugin Estimator Definition

The resulting plugin estimator for ζ_C is

$$\hat{\zeta}_{C\,final} = \begin{cases} \infty & \frac{1}{1+\hat{\zeta}_C S} \leq 0 \\ 0 & \frac{1}{1+\hat{\zeta}_C S} \geq 1 \\ \hat{\zeta}_C & \text{else} \end{cases}$$

(6.2.25)

where $\hat{\zeta}_C$ is defined as

$$
\hat{\zeta}_C := \frac{1}{S} \left(\left(\sum_{n \geq 2} \left(\left[\frac{1}{n(n-1)} \left[\sum_{\sigma \in [S]} (\hat{f}^{(\sigma)})^2 \left[\frac{\hat{\mathrm{var}}[N^{(\sigma)}(0) \mid N(0) = n] - n\hat{f}^{(\sigma)}(1 - \hat{f}^{(\sigma)})}{\hat{f}^{(\sigma)}(1 - \hat{f}^{(\sigma)})} \right] \right. \right. \right. \right.
$$
$$
\left. \left. \left. + \sum_{\sigma_1 \in [S]} \sum_{\sigma_2 \neq \sigma_1} \hat{f}^{(\sigma_1)} \hat{f}^{(\sigma_2)} \left[\frac{-\hat{\mathrm{cov}}[N^{(\sigma_1)}(0), N^{(\sigma_2)}(0) \mid N(0) = n] - n\hat{f}^{(\sigma_1)}\hat{f}^{(\sigma_2)}}{\hat{f}^{(\sigma_1)}\hat{f}^{(\sigma_2)}} \right] \right] \right] \right.
$$
$$
\left. \cdot \, \mathbb{P}(N(0) = n \mid N(0) \geq 2) \right) \right)^{-1} - 1 \right) . \tag{6.2.26}
$$

In the above definition, $\hat{f}^{(\sigma)}$ denote statistically consistent estimators of the frequencies $f^{(\sigma)}$. One possibile definition for such estimators (which was used for the implementation seen in section 6.4) is given in section 6.A. I did not want to assume that the scientist had, or wanted to rely on, a priori accurate estimates of the true frequency values. In principle, if such knowledge were available and trusted, then one would use those values for $\hat{f}^{(\sigma)}$ in equation 6.2.25.

Note that when $\frac{1}{1+\hat{\zeta}_C} > 1$, the result of solving for $\hat{\zeta}_C$ is less than 0. The interpretation in this case is extremely unclear given that the limit as $\zeta_C \to 0$ corresponds to "infinite heterogeneity". This is the reason for thresholding the values of $\hat{\zeta}_{C\,final}$ below by 0.

6.2.3.5 Effective Compositional Concentration

Using the chosen values of c_σ and c_{σ_1, σ_2} gives one possible choice for a definition of "effective compositional concentration". This estimand exists for any distribution of $\vec{\mathbf{N}}(0)$ with finite covariances, not just when $\mathbb{E}[\vec{\mathbf{N}}(0) \mid N(0)]$ is Dirichlet-Multinomial distributed:

$$
(\zeta_{C\,eff})_{final} = \begin{cases} \infty & \frac{1}{1+\zeta_{C\,eff}S} \leq 0 \\ 0 & \frac{1}{1+\zeta_{C\,eff}S} \geq 1 \\ \zeta_{C\,eff} & \text{else} \end{cases} \tag{6.2.27}
$$

with $\zeta_{C\,eff}$ defined (when $\frac{1}{1+\zeta_{C\,eff}S} \neq 0$) as

$$
\zeta_{C\,eff} = \frac{1}{S} \left(\left(\sum_{n \geq 2} \left(\left[\frac{1}{n(n-1)} \left[\sum_{\sigma \in [S]} (f^{(\sigma)})^2 \left[\frac{\mathrm{Var}[N^{(\sigma)}(0) \mid N(0) = n] - nf^{(\sigma)}(1 - f^{(\sigma)})}{f^{(\sigma)}(1 - f^{(\sigma)})} \right] \right. \right. \right. \right.
$$
$$
\left. \left. \left. + \sum_{\sigma_1 \in [S]} \sum_{\sigma_2 \neq \sigma_1} f^{(\sigma_1)} f^{(\sigma_2)} \left[\frac{-\mathrm{Cov}[N^{(\sigma_1)}(0), N^{(\sigma_2)}(0) \mid N(0) = n] - nf^{(\sigma_1)}f^{(\sigma_2)}}{f^{(\sigma_1)}f^{(\sigma_2)}} \right] \right] \right] \right.
$$
$$
\left. \cdot \, \mathbb{P}(N(0) = n \mid N(0) \geq 2) \right) \right)^{-1} - 1 \right) , \tag{6.2.28}
$$

The law of large numbers combined with the continuous mapping theorem for convergence in probability guarantees the consistency of the estimator (6.2.25) for the effective compositional concentration $(\zeta_{Ceff})_{final}$ whenever, under the given distribution for $\vec{\mathbf{N}}(0)$, the chosen estimators $\hat{f}^{(\sigma)}$ are also consistent for the true frequencies. (Asymptotic consistency is actually guaranteed, due to the same reasons, for the analogous plugin estimator and estimand given any choice of convex coefficients $c_\sigma, c_{\sigma_1,\sigma_2}$. The result does *not* depend on the chosen definitions $c_\sigma := f^{(\sigma)}$, $c_{\sigma_1,\sigma_2} := f^{(\sigma_1)} f^{(\sigma_2)}$.)

Note also that $(\zeta_{Ceff})_{final}$ is thresholded below by 0 for the same reason that $\hat{\zeta}_{Cfinal}$ is thresholded below by 0.

Because the effective compositional concentration, as well as ζ_{Ceff}, both equal the compositional concentration parameter ζ_C whenever $\mathbb{E}[\vec{\mathbf{N}}(0) \mid N(0)]$ is Dirichlet-Multinomial distributed, the estimand $(\zeta_{Ceff})_{final}$ may be considered a generalization of the parameter ζ_C to other distributions.

6.2.3.6 Effective Compositional Concentration without Frequencies

There could be distributions which are not parameterized in terms of frequencies $f^{(s)}$ to which we might still want to apply the notion of effective compositional concentration. One strategy we could employ is based on the expectations of the multinomial and Dirichlet-Multinomial distributions. For the marginal distribution corresponding to any strain s, when $\vec{\mathbf{N}}(0)$ conditional on $N(0) = n$ is multinomial or Dirichlet-Multinomial distributed, then one always has

$$\mathbb{E}\left[N^{(s)}(0) \mid N(0) = n\right] = nf^{(s)}. \tag{6.2.29}$$

Rearranging the relationship from (6.2.29) gives us a proxy which we can use in place of the frequency parameters $f^{(s)}$ (at least when $n \geq 1$):

$$f^{(s)} = \frac{\mathbb{E}\left[N^{(s)}(0) \mid N(0) = n\right]}{n}, \tag{6.2.30}$$

related to the so-called "effective frequency", see section 6.A for a derivation. Note this has the drawback of being dependent on the value of n.

Using these proxies for the frequencies (6.2.30), under the above distributional assumptions we can rewrite equation (6.2.19) as

$$\frac{1}{1+\zeta_C S} = \frac{1}{n-1} \cdot \left[\frac{n \operatorname{Var}[N^{(s)}(0) \mid N(0) = n] - \mathbb{E}\left[N^{(s)}(0) \mid N(0) = n\right]\left(n - \mathbb{E}\left[N^{(s)}(0) \mid N(0) = n\right]\right)}{\mathbb{E}[N^{(s)}(0) \mid N(0) = n]\left(n - \mathbb{E}[N^{(s)}(0) \mid N(0) = n]\right)}\right], \tag{6.2.31}$$

and rewrite equation (6.2.20) as

$$\frac{1}{1+\zeta_C S} = \frac{1}{n-1} \cdot \left[\frac{-n \operatorname{Cov}[N^{(s_1)}(0), N^{(s_2)}(0) \mid N(0) = n] - \mathbb{E}\left[N^{(s_1)}(0) \mid N(0) = n\right]\mathbb{E}\left[N^{(s_2)}(0) \mid N(0) = n\right]}{\mathbb{E}[N^{(s_1)}(0) \mid N(0) = n]\mathbb{E}[N^{(s_2)}(0) \mid N(0) = n]}\right]. \tag{6.2.32}$$

Based on this we can easily generalize the formulae from sections 6.2.3.2-6.2.3.5 to define a notion of effective compositional concentration that does not require the distribution to have frequency parameters $f^{(s)}$. Again, cf. section 6.A.

I do not claim that this is the only possibly proxy one could use in place of the frequency parameters for more general distributions. However, it does have the benefit of being widely applicable. Any definition of effective compositional concentration will (seemingly) require the distribution to have finite second (conditional) cumulants, which in turn guarantees the finiteness of the first (conditional) cumulants in formulae (6.2.31) and (6.2.32) above.

6.2.4 Maximum Likelihood Estimator for Compositional Concentration

This section explains how to use maximum likelihood to estimate the compositional concentration ζ_C. Section 6.2.4.1 computes the likelihood and log likelihood for the categorical distributions, while section 6.2.4.2 computes the score. Section 6.2.4.3 uses those results to derive the maximum likelihood estimators for the Dirichlet-Multinomial distribution.

6.2.4.1 Dirichlet-Multinomial (Log) Likelihood

Starting from equation (3.6.1),

- if we assume for mathematical convenience that the distributions of all droplets are mutually independent,

- using the identity[9] $\frac{\Gamma(x+M)}{\Gamma(x)} = \prod_{m=0}^{M-1}(x+m)$ for all positive integers M,

- if we assume that the droplets are Dirichlet-Multinomial distributed conditional on their count distributions,

then this is the full likelihood for an entire batch of D droplets:

$$
\mathcal{L}(\vec{f}, \zeta_C) = \prod_{d \in [D]} \left[\mathbb{P}(N_d(0) = n_d) \cdot \left(\prod_{\nu=0}^{n_d-1} (\zeta_C S + \nu)^{-1} \right) \cdot \right.
$$
$$
\left. \prod_{s \in [S]} \left(\prod_{\nu=0}^{n_d^{(s)}-1} (\zeta_C S f^{(s)} + \nu) \right) \cdot \binom{n_d}{n_d^{(1)} \cdots n_d^{(S)}} \right],
\tag{6.2.33}
$$

where (as a reminder) by definition $n_d = \sum_{s \in [S]} n_d^{(s)}$.

---

[9]Cf. footnote 6.

Therefore this is the full log likelihood:

$$\ell(\vec{f}, \zeta_C) :=$$

$$\log\left(\mathcal{L}(\vec{f}, \zeta_C)\right) = \sum_{d \in [D]} \left[\log\left(\mathbb{P}(N_d(0) = n_d)\right) - \sum_{\nu=0}^{n_d - 1} \log(\zeta_C S + \nu) + \right.$$

$$\left. \sum_{s \in [S]} \left(\sum_{\nu=0}^{n_d^{(s)} - 1} \log(\zeta_C S f^{(s)} + \nu) \right) + \log\left(\binom{n_d}{n_d^{(1)} \cdots n_d^{(S)}} \right) \right] . \tag{6.2.34}$$

6.2.4.2 Dirichlet-Multinomial Score

Assuming that the count distribution has no dependence[10] on ζ_C, it follows that the score with respect to ζ_C is

$$\frac{\partial \ell(\vec{f}, \zeta_C)}{\partial \zeta_C} = \sum_{d \in [D]} \left[-\sum_{\nu=0}^{n_d - 1} \frac{S}{\zeta_C S + \nu} + \sum_{s \in [S]} \left(\sum_{\nu=0}^{n_d^{(s)} - 1} \frac{S f^{(s)}}{\zeta_C S f^{(s)} + \nu} \right) \right] . \tag{6.2.35}$$

Using the commutativity of addition (i.e. rearranging terms) for purposes of computational efficiency, the above may be rewritten

$$\frac{\partial \ell(\vec{f}, \zeta_C)}{\partial \zeta_C} = -\sum_{m=1}^{M} \frac{D_{\geq m} S}{\zeta_C S + (m - 1)} + \sum_{s \in [S]} \left(\sum_{m=1}^{M^{(s)}} \frac{D_{\geq m}^{(s)} S f^{(s)}}{\zeta_C S f^{(s)} + (m - 1)} \right) , \tag{6.2.36}$$

where I have defined

$$M := \max_{d \in [D]} N_d(0) ,$$

$$\forall s \in [S] \quad M^{(s)} := \max_{d \in [D]} N_d^{(s)}(0) ,$$

$$D_{\geq m} := |\{d : N_d(0) \geq m\}| ,$$

$$\forall s \in [S] \quad D_{\geq m}^{(s)} := \left| \left\{ d : N_d^{(s)}(0) \geq m \right\} \right| . \tag{6.2.37}$$

This is basically the same approach to grouping terms as suggested in [Skl14].

6.2.4.3 Dirichlet-Multinomial Maximum Likelihood Estimator

Assuming that the frequencies $\vec{f} := (f_1, \ldots, f^{(S)})$ are already known or estimated (and thus don't need to be treated as nuisance parameters), it follows from the above formula (6.2.36)

---

[10]Or ignoring such a dependence if it exists, e.g. as in the case of hNBDM, and choosing to consider the maximum *conditional* likelihood estimator instead of the MLE sensu stricto.

for the score that a maximum likelihood estimator $\hat{\zeta}_C$ for the compositional concentration ζ_C must be a solution of the following equation:

$$\sum_{m=1}^{M} \frac{D_{\geq m}}{\hat{\zeta_C}S + (m-1)} = \sum_{s \in S} \left(\sum_{m=1}^{M^{(s)}} \frac{D_{\geq m}^{(s)} f^{(s)}}{\hat{\zeta_C} S f^{(s)} + (m-1)} \right) . \qquad (6.2.38)$$

I only claim that satisfying the above equation is necessary, not sufficient. In particular, I make no claims regarding the conditions under which a root for the above equation exists, nor regarding the conditions under which a root (if it exists) will be unique. These conditions appear to be unknown.

6.3 Methods

Section 6.3.1 explains how the empirical distributions of the plugin heterogeneity estimators were computed as a function of batch size. Section 6.3.2 provides analogous details for the ML estimators. Section 6.3.3 explains how these results were depicted. Complete implementation details can be found in the code at the relevant GitLab repository. See https://gitlab.com/krinsman/droplets.

6.3.1 Computing Plugin Estimator Distributions

For each of the seven simulated distributions, for each of the 500 simulations, for each of the three batch sizes (small = $10,000$ droplets, medium = $500,000$ droplets, large = $15,000,000$ droplets), using NumPy [HMvdW+20] version 1.20.2 I partitioned the simulation results into equal parts of the given batch size. Since each simulation contains $15,000,000$ droplets, this corresponds to $1,500$ small batches per simulation, 30 medium batches per simulation, and 1 large batch per simulation. Then for each of the batches, I evaluated the plugin density heterogeneity estimator (6.2.11) and the plugin compositional heterogeneity estimator (6.2.25) and stored the results.

6.3.2 Computing MLE Distributions

Simulated droplets were partitioned into batches exactly as was done for the plugin estimators. Cf. section 6.3.1 above for details.

I tried to use Brent's method [Bre73, Ch. 3-4] whenever possible to estimate roots of the score functions, because Powell's method [Pow70] was often unstable. Preliminary investigations (data not shown) suggested that Powell's method often either failed to converge, or failed to converge to a reasonable value (for example returning a negative answer, or one that is far too large), under conditions where Brent's method converged to a reasonable answer similar to the truth. While it was always possible to supply a reasonable interval of values for Brent's method to restrict its search, the method requires the signs of the function to

differ at both ends of the interval. It was difficult to systematically choose (i.e. without manual intervention) interval endpoints which would guarantee opposite signs at both ends of the interval for all possible simulated datasets. Therefore Powell's method often had to be used as an alternative, which may have introduced many inaccurate estimates. While gradient-based methods probably would have been more accurate and converged faster, this level of accuracy seemed sufficient for a proof of principle. Estimates that were less than zero were obviously erroneous and therefore discarded – this affected fewer than 3% of the estimates in general and usually less (data not shown).

For the Negative Binomial (Density Concentration):

- Check that the sample variance is larger than the sample mean, otherwise return ∞ and terminate.

- Set initial guess value for the root-finding algorithm to be the value of the plugin estimator[11], equation (6.2.11).

- If the score function, computed using the terms of equation (6.2.18) rearranged to one side to match (6.2.16), had different signs at 10^{-4} and 10^4:

  - Use Brent's method [Bre73, Ch. 3-4] implemented in SciPy [VGO+20] with brackets 10^{-4} and 10^4 to look for a root of the score function.

- If the score function, computed using the terms of equation (6.2.18) rearranged to one side to match (6.2.16), had the same sign at 10^{-4} and 10^4:

  - Use Powell's hybrid method [Pow70] as implemented in the HYBRD routine of MINPACK [MGH80][Cow84, Ch. 5] and made available through SciPy [VGO+20] to search for a root of the score function.
  - If the method fails to converge and the final guess is greater than 10^4, return ∞.
  - Otherwise (even if the method fails to converge) return the final guess (or discard if the final guess is less than zero).

For the Dirichlet-Multinomial (Compositional Concentration):

- Set the initial guess value for the root-finding algorithm to be the value of the plugin estimator, equation (6.2.25).

- If the plugin estimator was larger than 10^6 (including ∞), set the initial guess value to be 0.5×10^6.

- If the score function, computed using equation (6.2.36), had different signs at 10^{-6} and 10^6:

---

[11]This never exceeded 10,000 when the sample variance was larger than the sample mean.

  - Use Brent's method [Bre73, Ch. 3-4] implemented in SciPy [VGO$^+$20] with brackets 10^{-6} and 10^6 to look for a root of the score function.

- If the score function, computed using equation (6.2.36), had the same sign at 10^{-6} and 10^6:

  - Use Powell's hybrid method [Pow70] as implemented in the HYBRD routine of MINPACK [MGH80][Cow84, Ch. 5] and made available through SciPy [VGO$^+$20] to search for a root of the score function.
  - If the method fails to converge and the final guess is greater than 10^6, return ∞.
  - Otherwise (even if the method fails to converge) return the final guess (or discard if the final guess is less than zero).

6.3.3   Plotting Estimator Distributions

I plotted results using Matplotlib [Hun07] version 3.4.1 and Seaborn [Was21] version 0.11.1. Kernel density estimates were made using the default Seaborn [Was21] settings[12]. I computed empirical cumulative density functions using StatsModels [SP10] version 0.12.2, which uses the binary search algorithm from NumPy [HMvdW$^+$20]. Finally, I computed the empirical survival functions from the empirical cumulative density functions.

6.4   Results

Section 6.4.1 explains what indicates that the plugin estimators are consistent in practice when the concentration parameters are infinite. Section 6.4.2 explains the analogous evidence when the concentration parameters are finite.

6.4.1   Estimators are Consistent for Infinite Concentrations

When the true value of the estimand is infinite, we see the correct behavior for a consistent estimator. Section 6.4.1.1 explains how the survival functions support this conclusion. Section 6.4.1.2 explains how the distributions of incorrect finite estimates also supports this conclusion.

6.4.1.1   Survival Functions Show Anticipated Behavior

When the true value of the concentration estimand is infinite, we see for both the density (as in figures 6.1a, 6.2a, 6.3a, 6.4a for the plugin estimators and figures 6.1b, 6.2b, 6.3b, 6.4b for the ML estimators) and compositional concentration estimators (as in figure 6.5a for the plugin estimator and figure 6.5b for the ML estimator) that the empirical survival

---

[12]Via SciPy [VGO$^+$20], which chooses bandwidths according to Scott's Rule [Sco92].

function generally moves to the upper right as the batch size increases. This is the correct behavior for a consistent estimator, since in the case that all estimates were infinite the survival function would be the horizontal line $y = 1$.

6.4.1.2 Incorrect Finite Estimates Shift Further to the Right as Batch Size Increases

Moreover, as the size of the batches increases (cf. figures 6.6a, 6.7a, 6.8a, 6.9a, and 6.10a for the plugin estimators and figures 6.6b, 6.7b, 6.8b, 6.9b, and 6.10b for the ML estimators) the distribution of the remaining estimates shifts further and further to the right. In other words, even when the estimators incorrectly estimate a finite value, the estimated finite value still tends to increase (and thus "better approximates infinity") as the data size increases.

6.4.2 Estimators are Consistent for Finite Concentrations

When the true value of the concentration estimand is finite, we also see the correct behavior for a consistent estimator. Section 6.4.2.1 explains how the pattern of infinite estimates supports this conclusion. Section 6.4.2.2 explains how the estimators satisfy the two minimal requirements for the behavior of a "decent" estimator when the model is correctly specified. Section 6.4.2.2 explains why the estimates are also good even in the case when the model is not correctly specified.

6.4.2.1 Incorrect Infinite Estimates Only Occur for Distributions Very Similar to hPoMu

For the distributions which are very similar to the homogeneous case with infinite concentration, the concentration estimates are sometimes infinite (cf. figures 6.11a, 6.12a, and 6.13a for the plugin estimators and figures 6.11b, 6.12b, and 6.13b for the ML estimators). Nevertheless, these figures also show that, as the batch size increases, the proportion of concentration estimates which are infinite decreases. This appears to occur more quickly for the compositional concentration estimators (see figures 6.11a and 6.13a for the plugin estimators and figures 6.11b and 6.13b for the ML estimators) than for the density concentration estimators (see figure 6.12a for the plugin estimator and figure 6.12b for the ML estimator).

For the distributions which are less similar to hPoMu, when the true value of the concentration estimand is finite, the estimated values are either never or almost never infinite. This is why survival functions are not reported neither for the density concentration estimates of hNBDM with $\zeta = 1$ and hExhNBDM with $\mathbb{E}[\zeta] = 1$ nor for the compositional concentration estimates of hPoDM with $\zeta_C = 1$, hNBDM with $\zeta = 1$, hExhPoDM with $\mathbb{E}[\zeta_C] = 1$, and hExhNBDM with $\mathbb{E}[\zeta] = 1$.
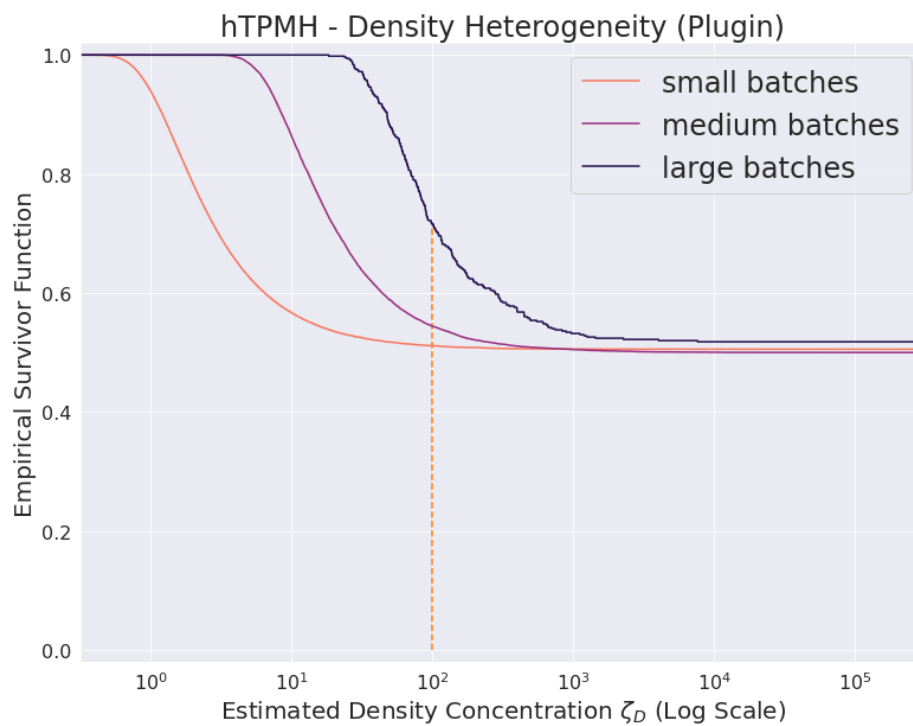
(a)
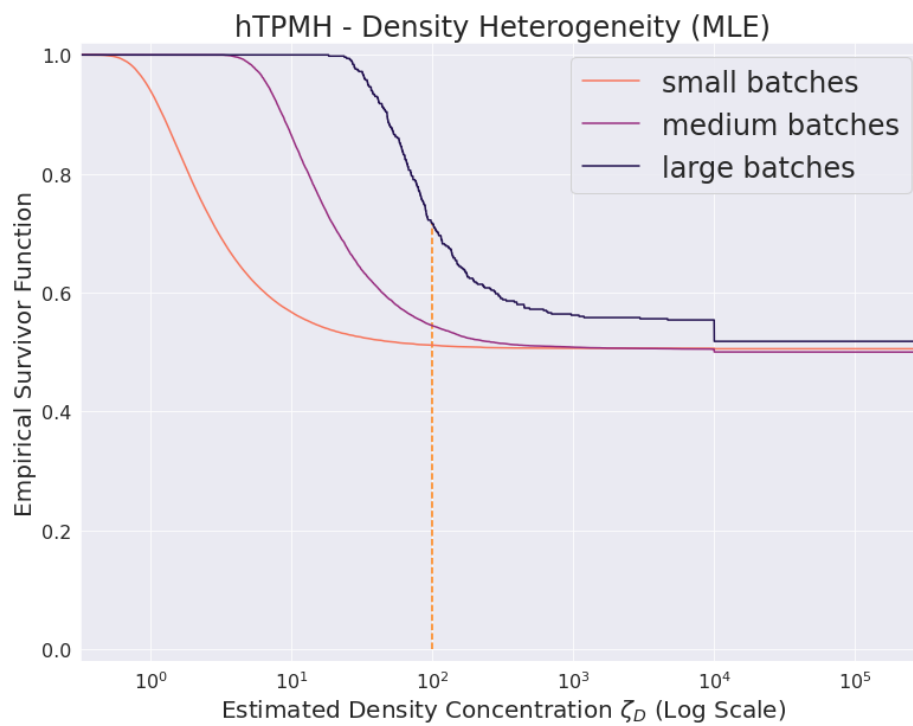
(b)

Figure 6.1: Survival functions of hTPMH density concentration estimates.
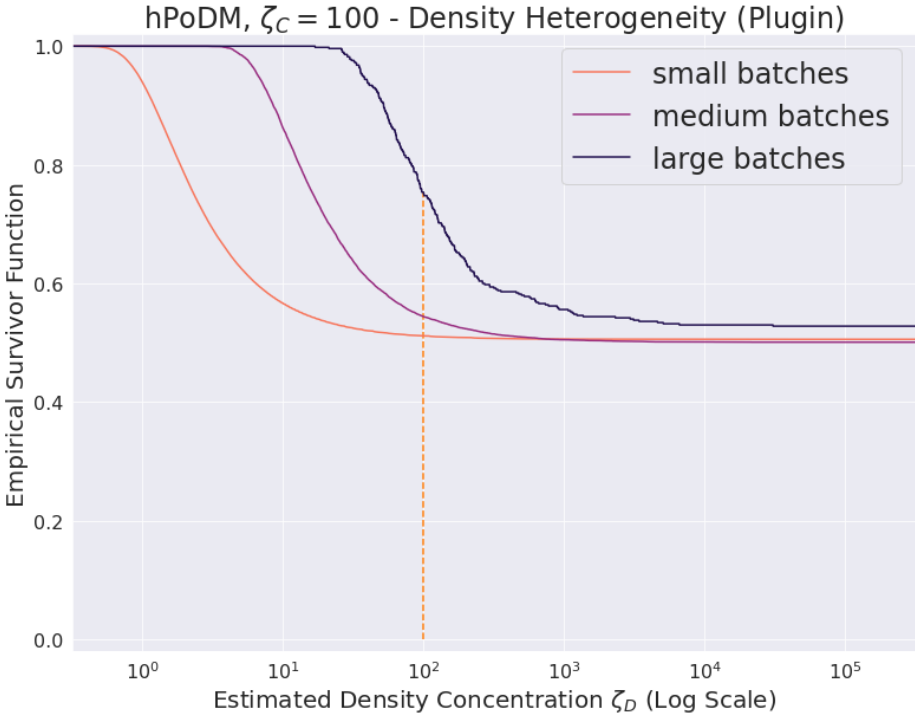
(a)
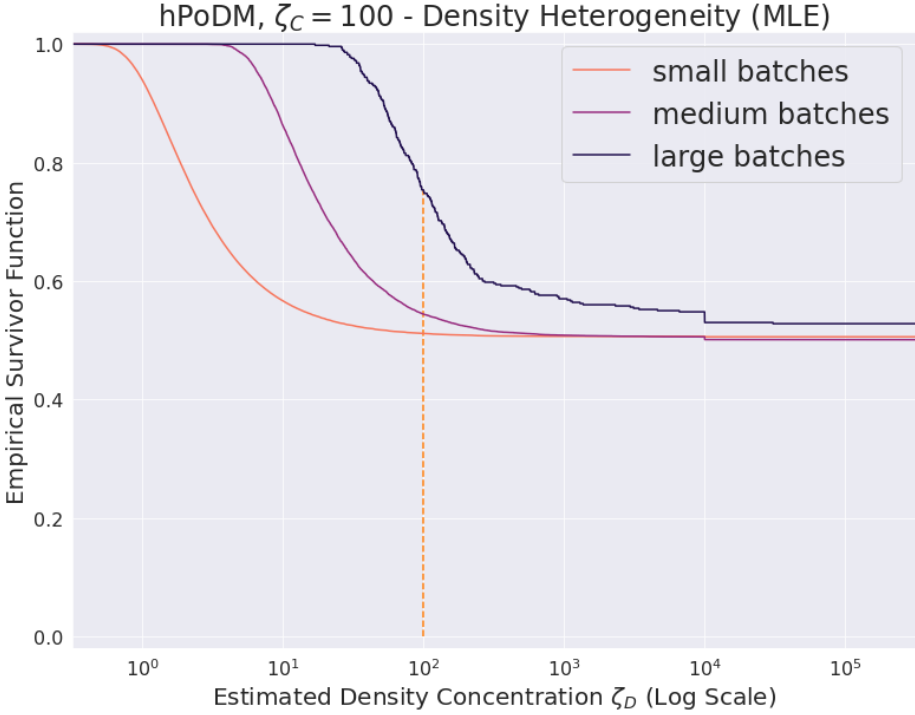
(b)

Figure 6.2: Survival functions of hPoDM ($\zeta_C = 100$) density concentration estimates.
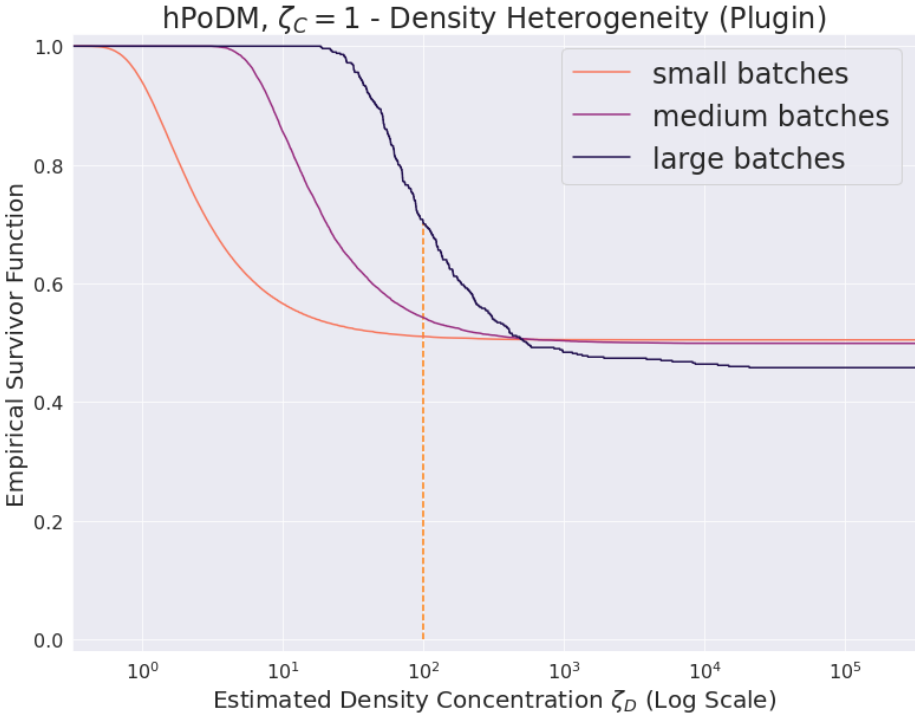
(a)
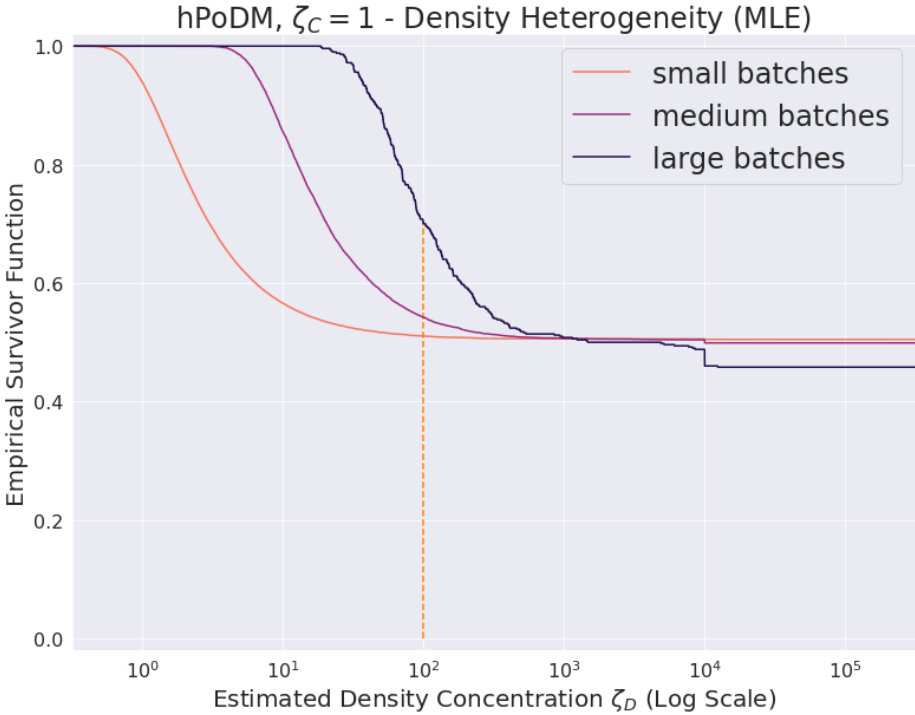
(b)

Figure 6.3: Survival functions of hPoDM ($\zeta_C = 1$) density concentration estimates.
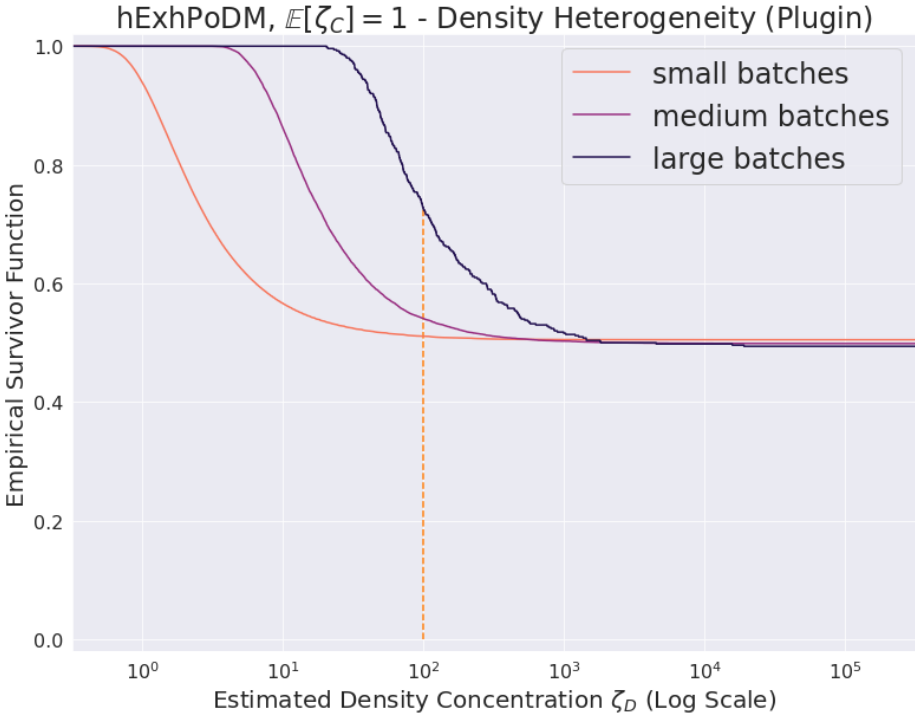
(a)
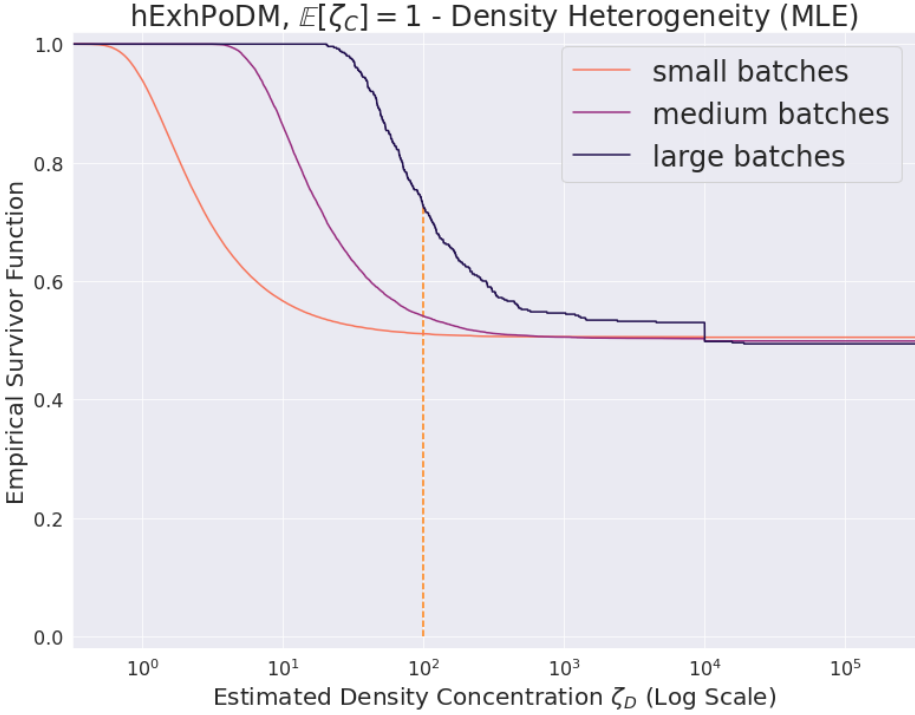
(b)

Figure 6.4: Survival functions of hExhPoDM density concentration estimates.
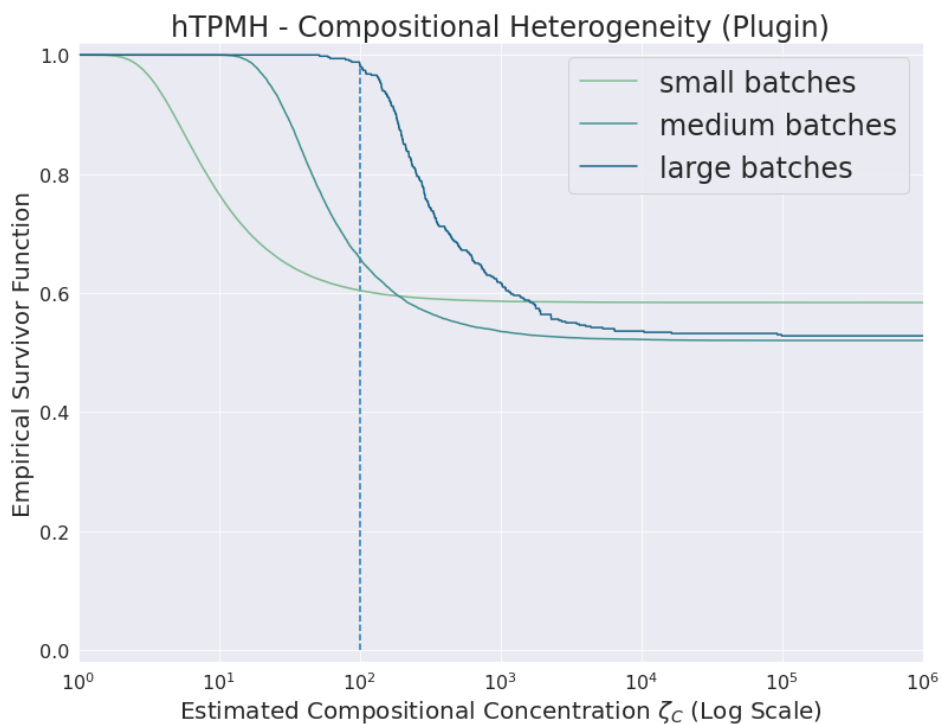
(a)
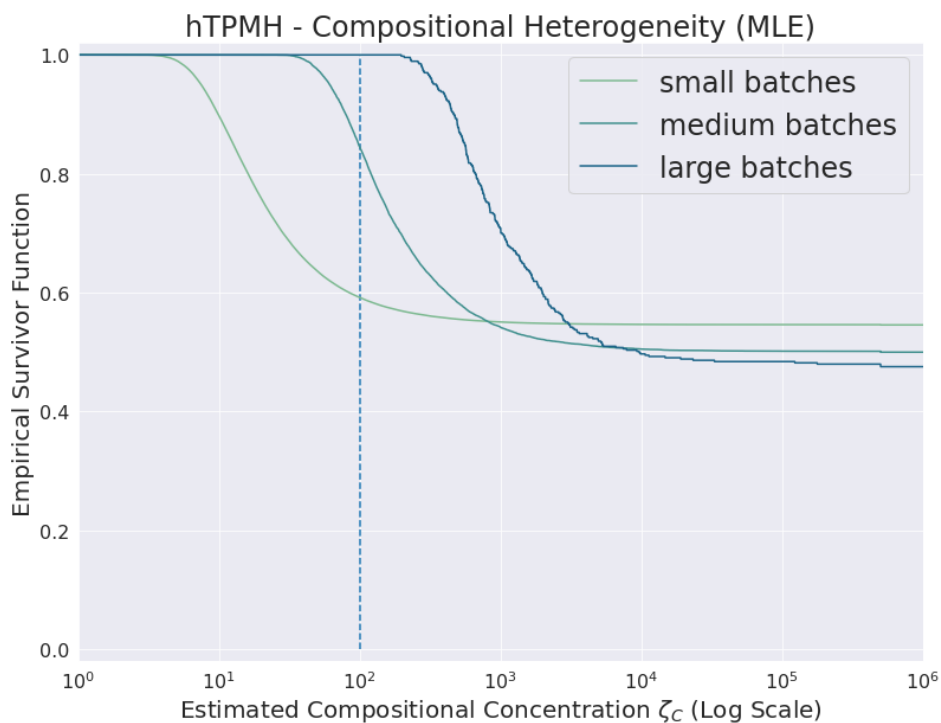
(b)

Figure 6.5: Survival functions of hTPMH compositional concentration estimates.
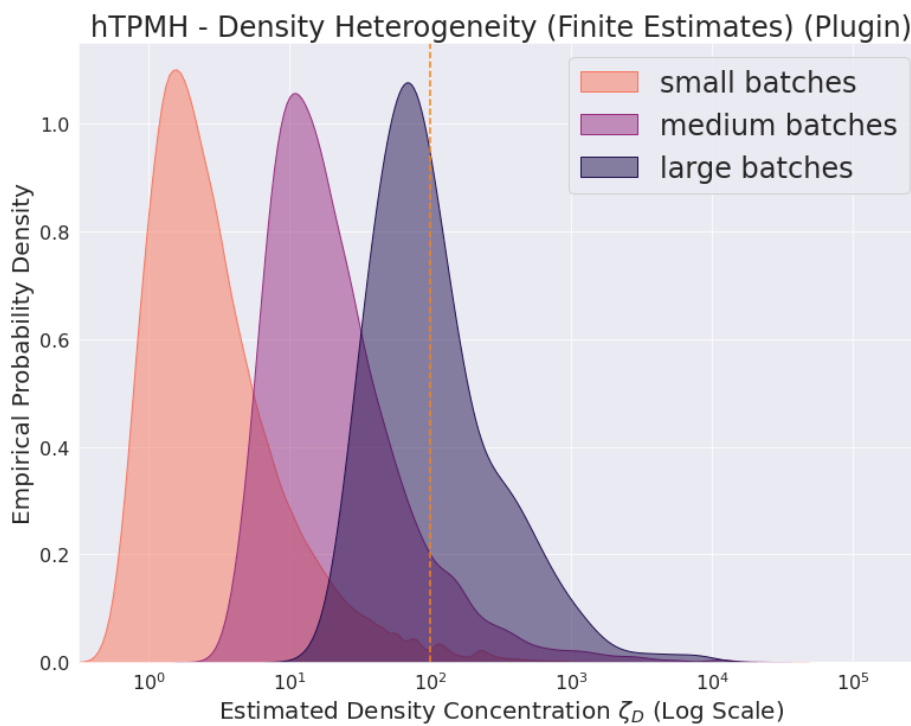
(a)
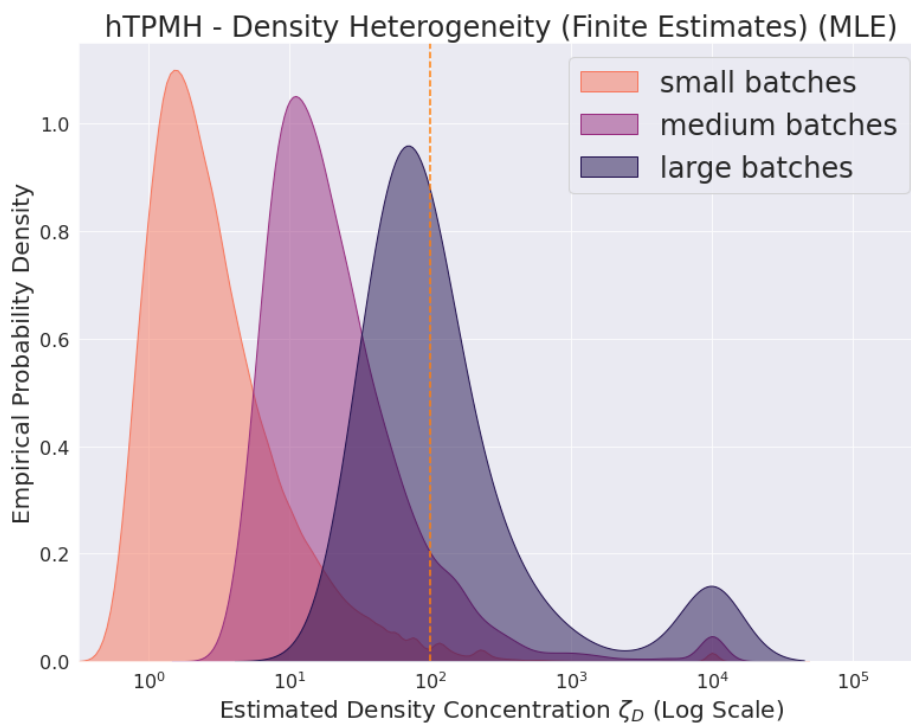
(b)

Figure 6.6: KDE plots of hTPMH density concentration estimates.
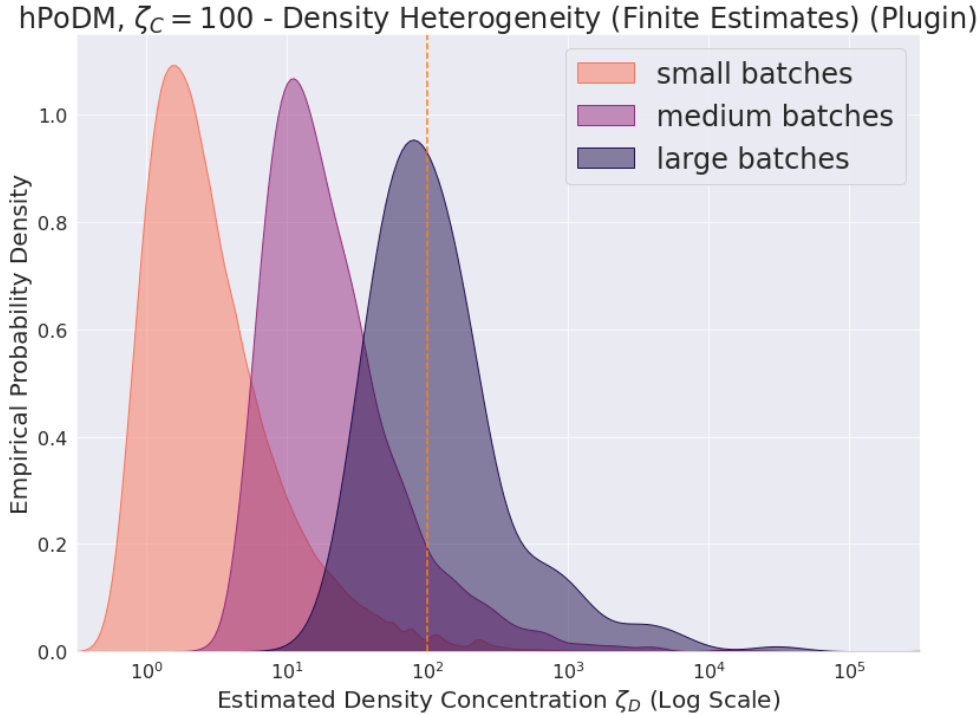
(a)
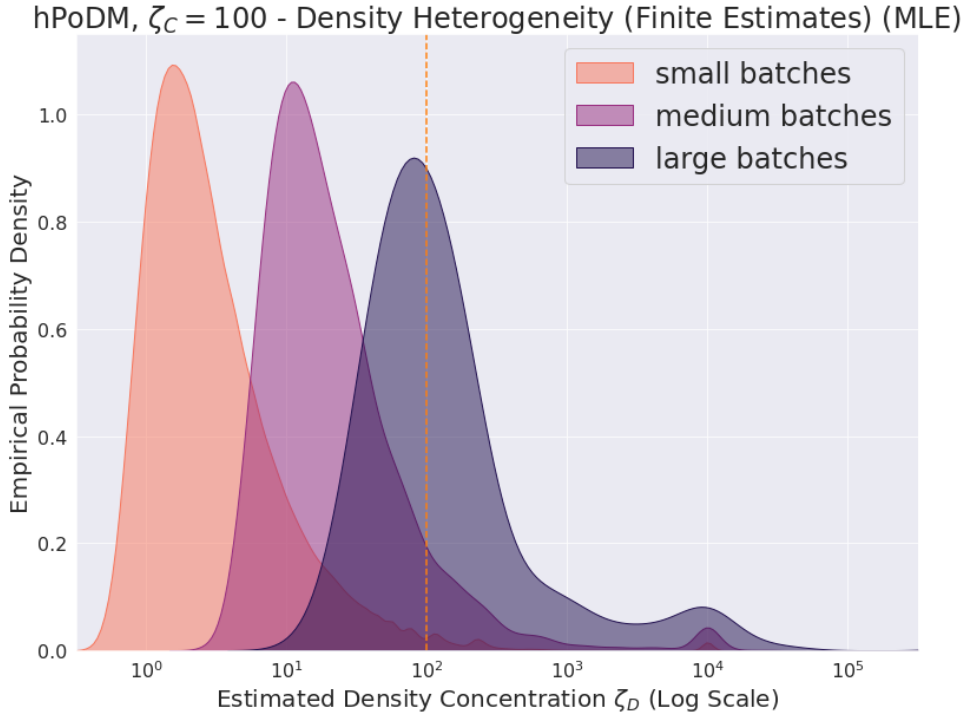
(b)

Figure 6.7: KDE plots of hPoDM ($\zeta_C = 100$) density concentration estimates.
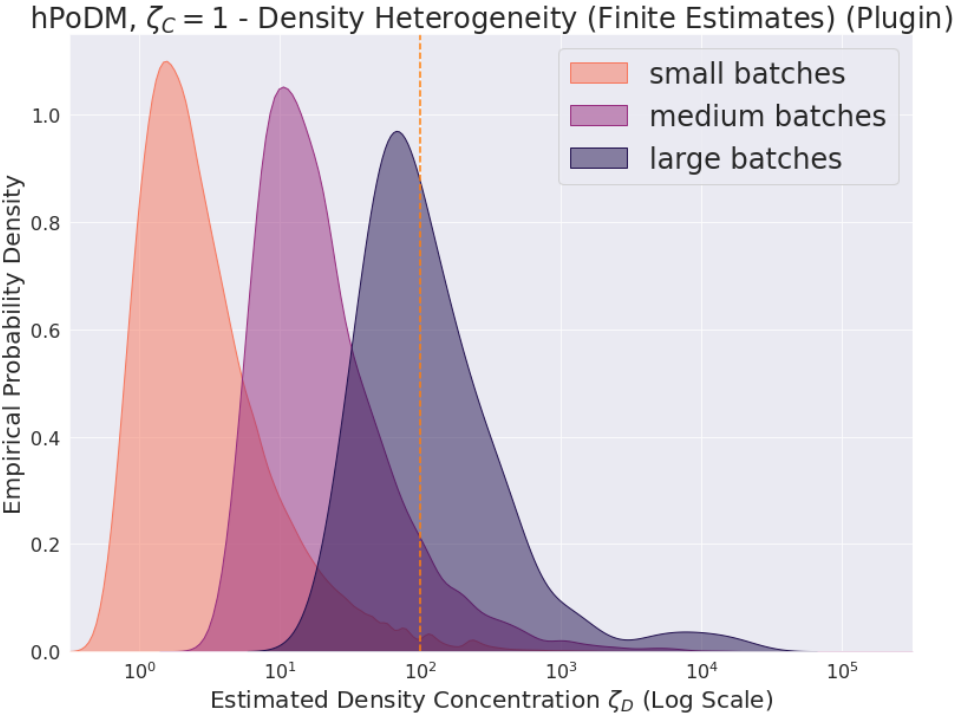
(a)
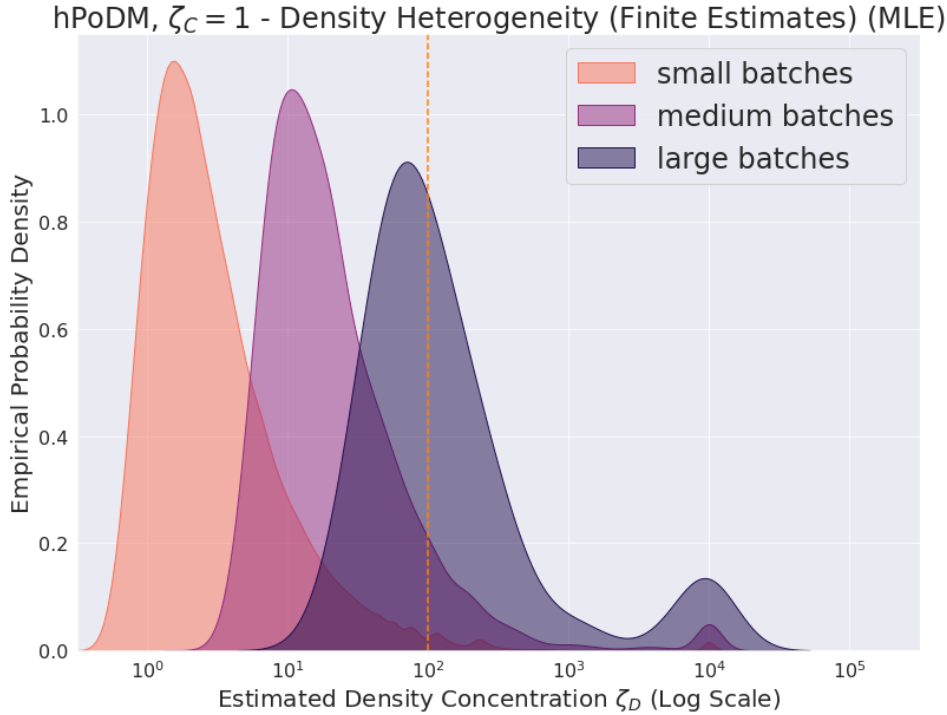
(b)

Figure 6.8: KDE plots of hPoDM ($\zeta_C = 1$) density concentration estimates.
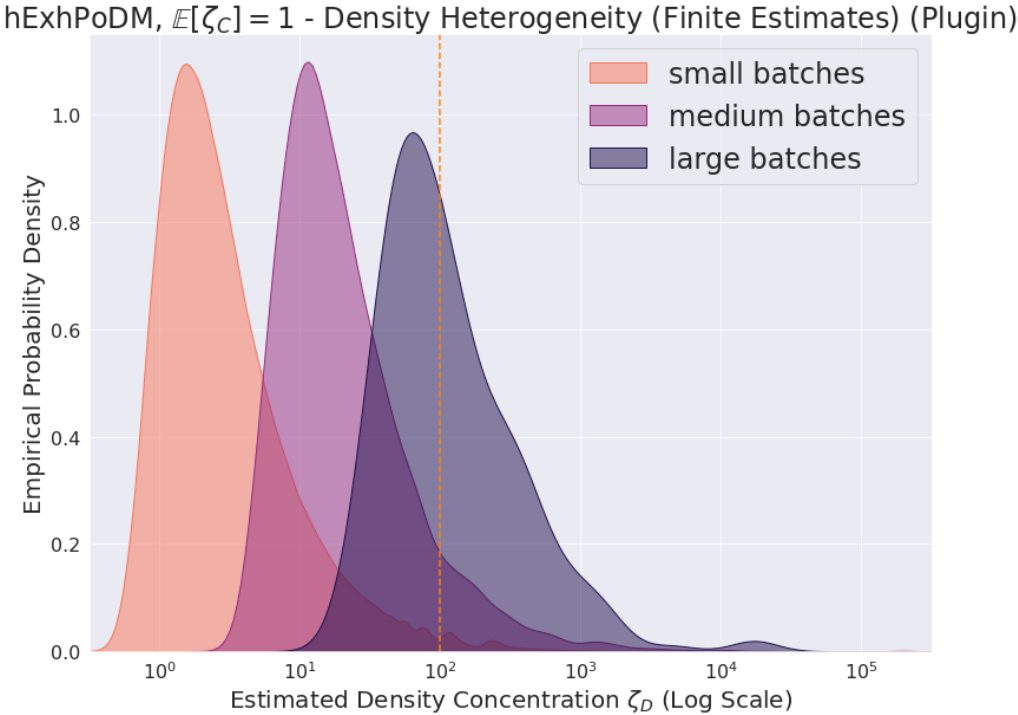
(a)
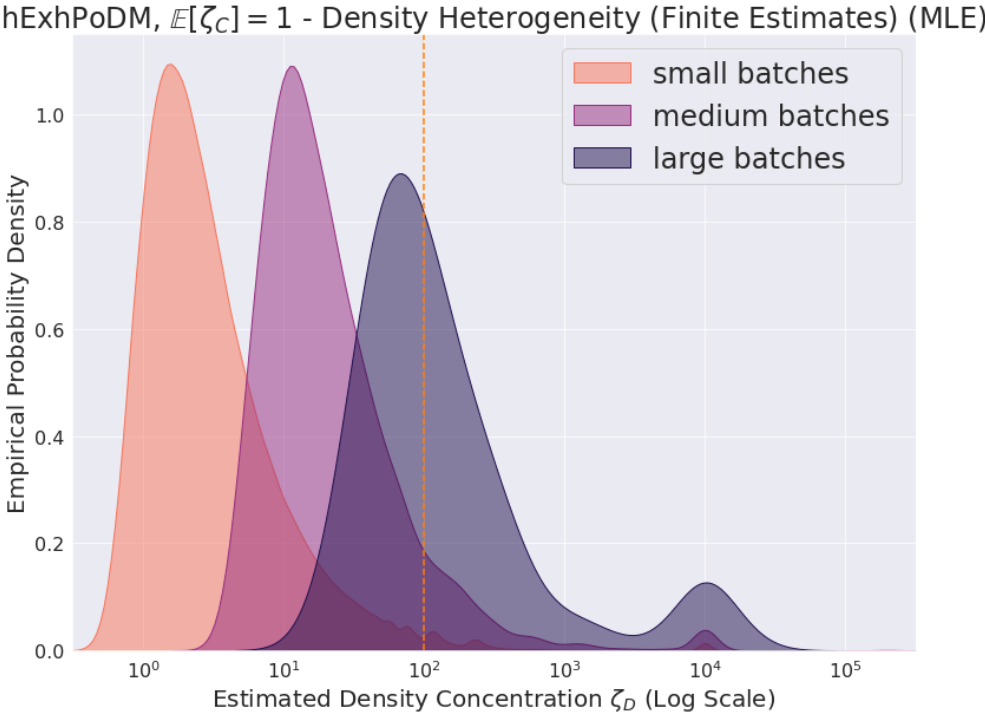
(b)

Figure 6.9: KDE plots of hExhPoDM density concentration estimates.
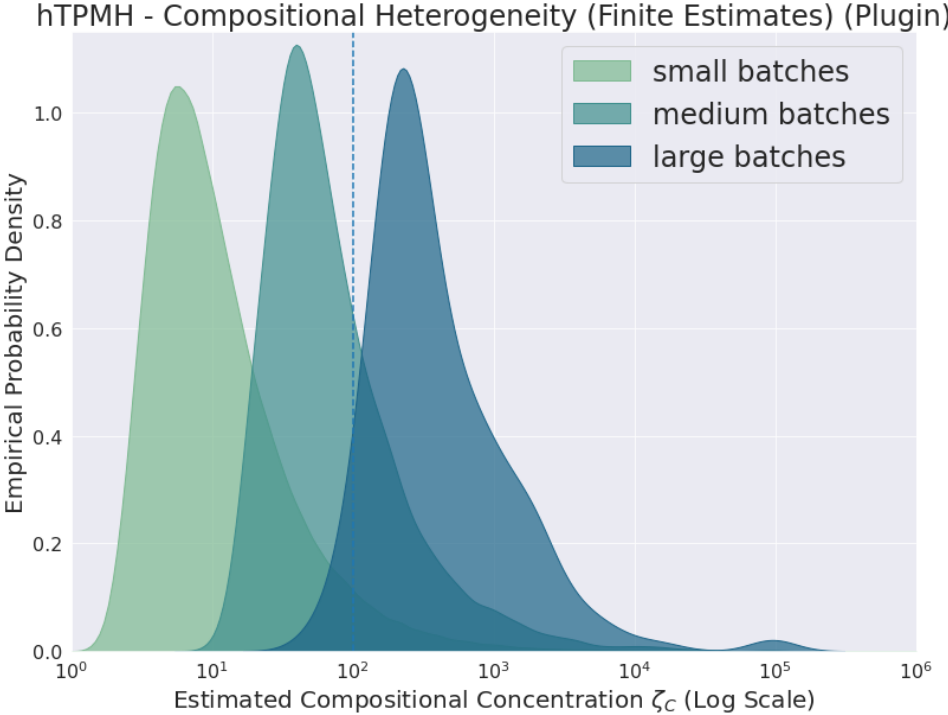
(a)
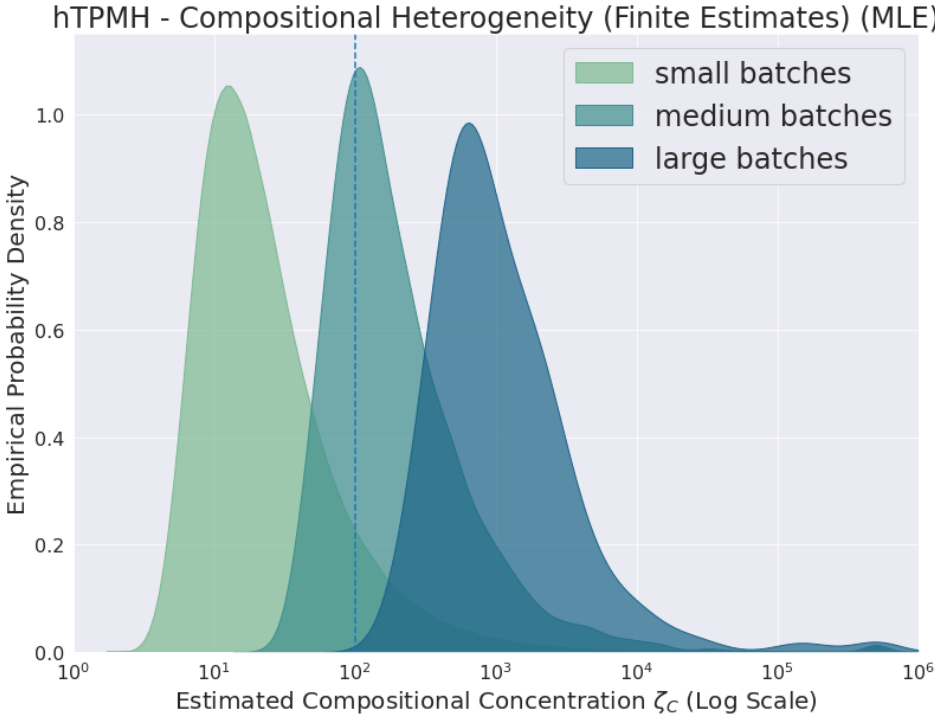
(b)

Figure 6.10: KDE plots of hTPMH compositional concentration estimates.
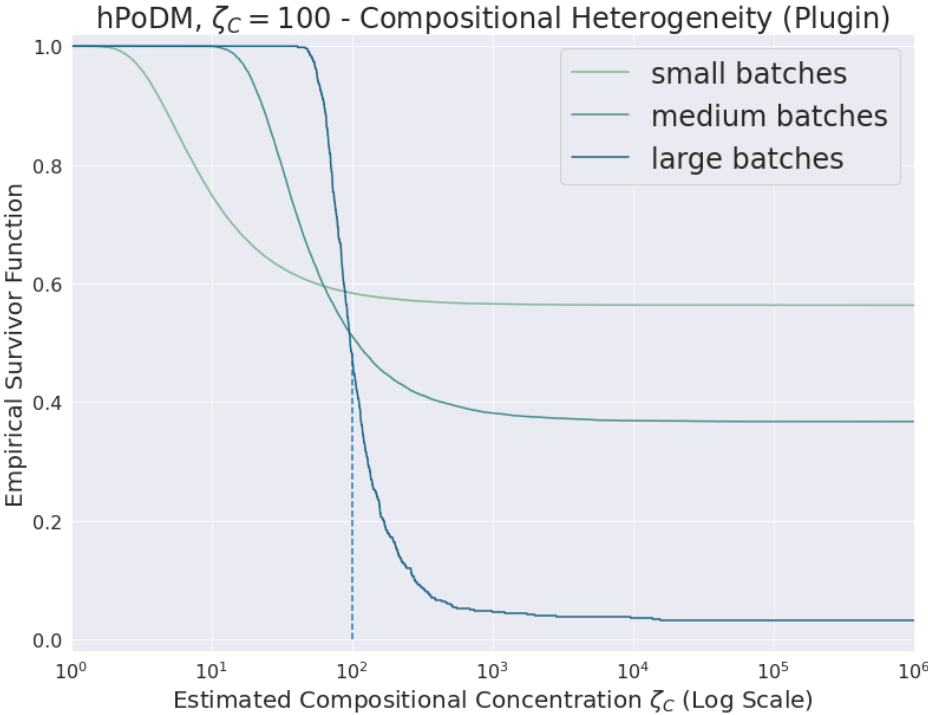
(a)
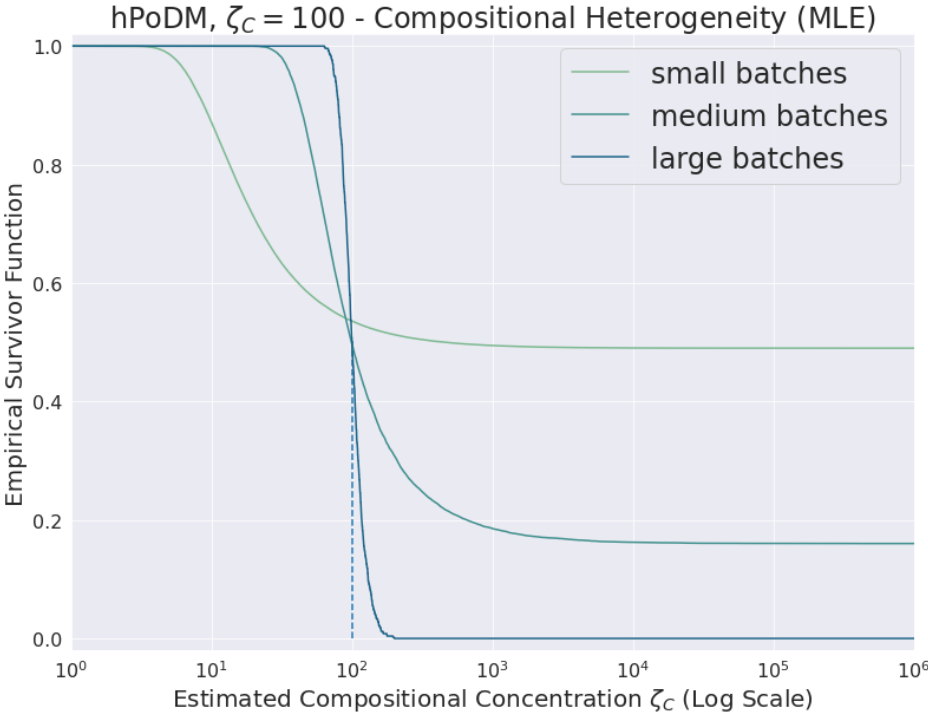
(b)

Figure 6.11: Survival functions of hPoDM ($\zeta_C = 100$) compositional concentration estimates.
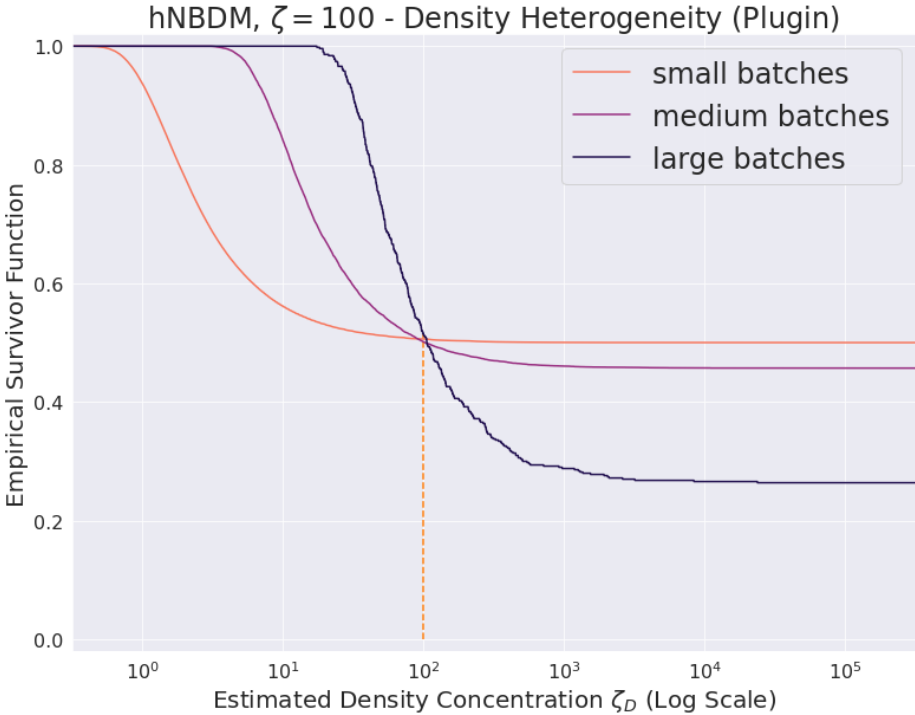
(a)
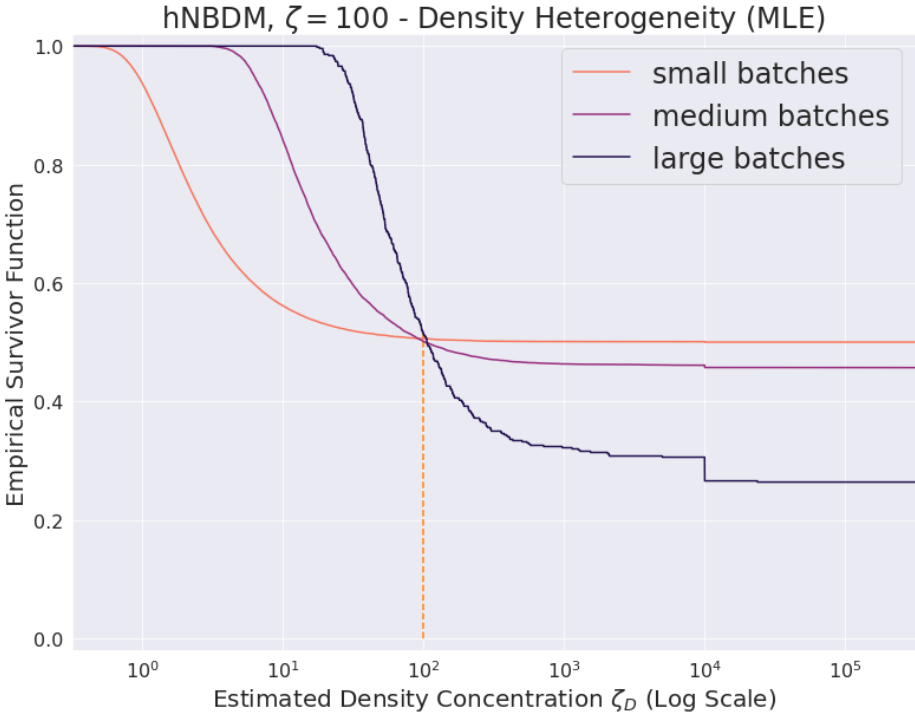
(b)

Figure 6.12: Survival functions of hNBDM ($\zeta = 100$) density concentration estimates.
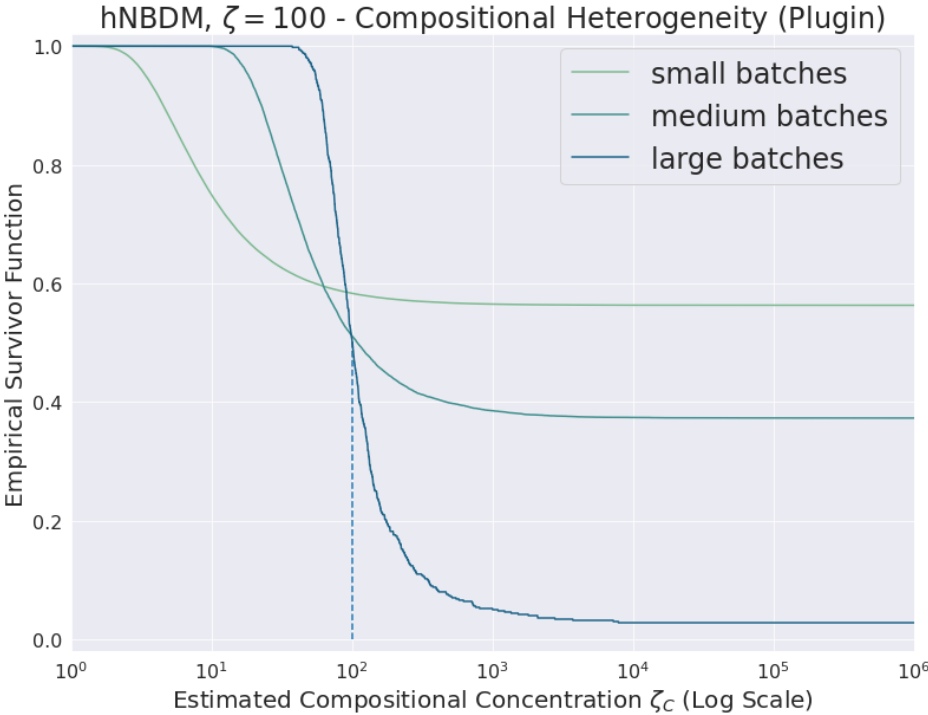
(a)
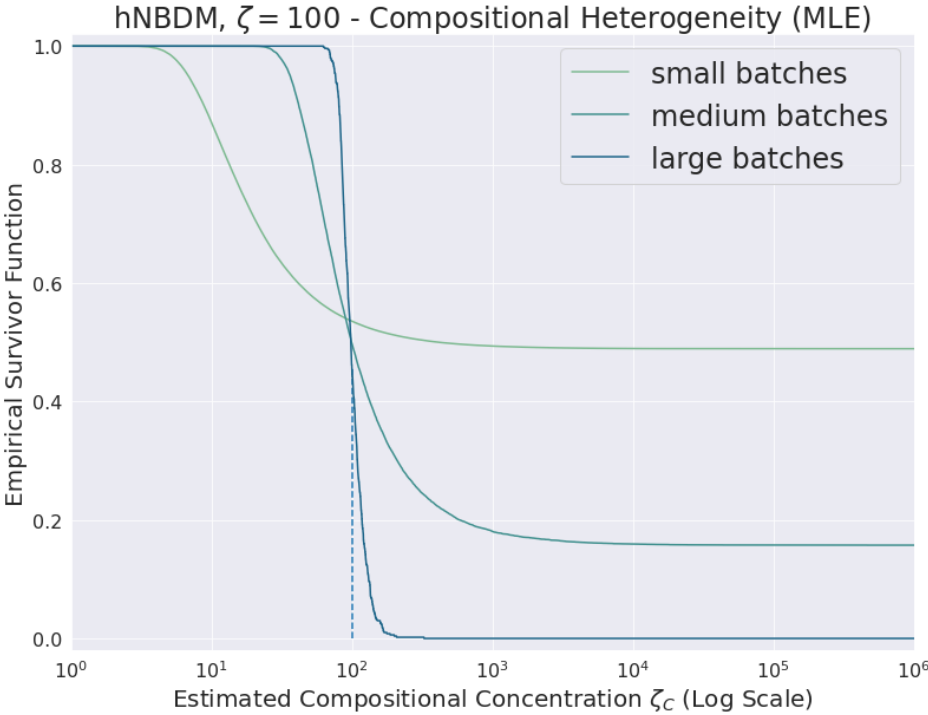
(b)

Figure 6.13: Survival functions of hNBDM ($\zeta = 100$) compositional concentration estimates.

6.4.2.2 Accuracy and Precision Both Increase with Batch Size

When the true value of the concentration estimand is finite, and the model is correctly specified, we see (cf. figures 6.14a, 6.15a, 6.16a, 6.17a, 6.18a, and 6.19a for the plugin estimators and figures 6.14b, 6.15b, 6.16b, 6.17b, 6.18b, and 6.19b for the ML estimators) that the distribution of the (finite) estimates is centered around the true value.

When the limit approached by the estimators is finite, even when the model is *not* correctly specified, the variance of the distribution of estimated concentrations decreases as the batch size increases. Cf. figures 6.14a, 6.15a, 6.16a, 6.17a, 6.18a, 6.19a, 6.20a, 6.21a, and 6.22a for the plugin estimators and figures 6.14b, 6.15b, 6.16b, 6.17b, 6.18b, 6.19b, 6.20b, 6.21b, and 6.22b for the ML estimators.

6.4.2.3 Results are Qualitatively Correct under Model Misspecification

For hExhPoDM with $\mathbb{E}[\zeta_C] = 1$, and hExhNBDM with $\mathbb{E}[\zeta] = 1$, the model is misspecified[13], since these distributions do not have single values of concentration parameters. Nevertheless, from the results in previous chapters it is also clear that these are the most heterogeneous distributions. These estimators therefore give qualitatively correct answers for these distributions, despite the model being misspecified, because (cf. figures 6.20a, 6.21a, and 6.22a for the plugin estimators and figures 6.20b, 6.21b, and 6.22b for the ML estimators) the lowest concentration estimates occur for these distributions.

6.4.3 Comparing ML Estimator and Plugin Estimator Results

An important reason for using ML estimators over plugin estimators is increased efficiency. Cf. the discussion from the beginning of section 6.2. Therefore I evaluate whether the precision of the ML estimator appears to increase much more rapidly than that of its plugin counterpart for the density estimators in section 6.4.3.1 and for the concentration estimators in section 6.4.3.2. One reason for using the plugin estimators over ML estimators is because the plugin estimators arguably make fewer model assumptions. Again, cf. the discussion at the beginning of section 6.2. Therefore in section 6.4.3.3 I compare the behaviors of the plugin and ML estimators for the misspecified models.

6.4.3.1 Precision of ML Estimator vs. Plugin Estimator for Density Concentration

For density concentration, both ML and plugin appear to demonstrate effectively the same efficiency in practice. See figures 6.1a, 6.1b, 6.2a, 6.2b, 6.3a, 6.3b, 6.4a, and 6.4b for survival curves when the true value of the estimand is infinite, figures 6.12a and 6.12b for survival

---

[13] Note that technically for hExhPoDM, only the categorical distribution, corresponding to the compositional concentration, is misspecified relative to the chosen estimators. For hExhNBDM both the count distribution, corresponding to the density concentration, and the categorical distribution are misspecified relative to the chosen estimators.
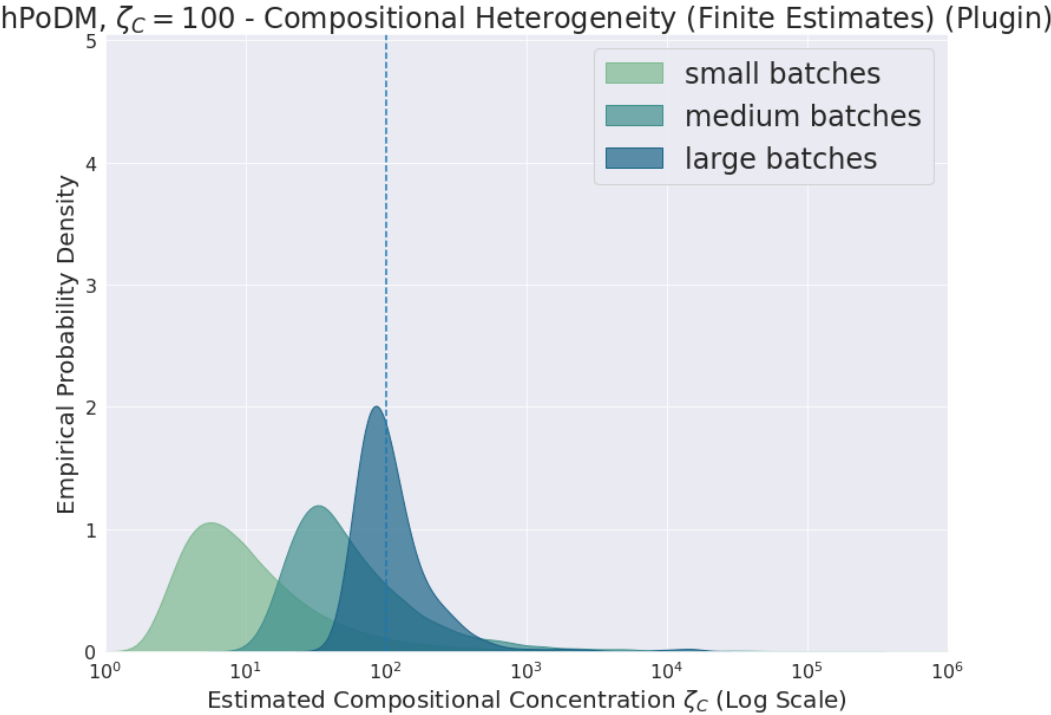
(a)
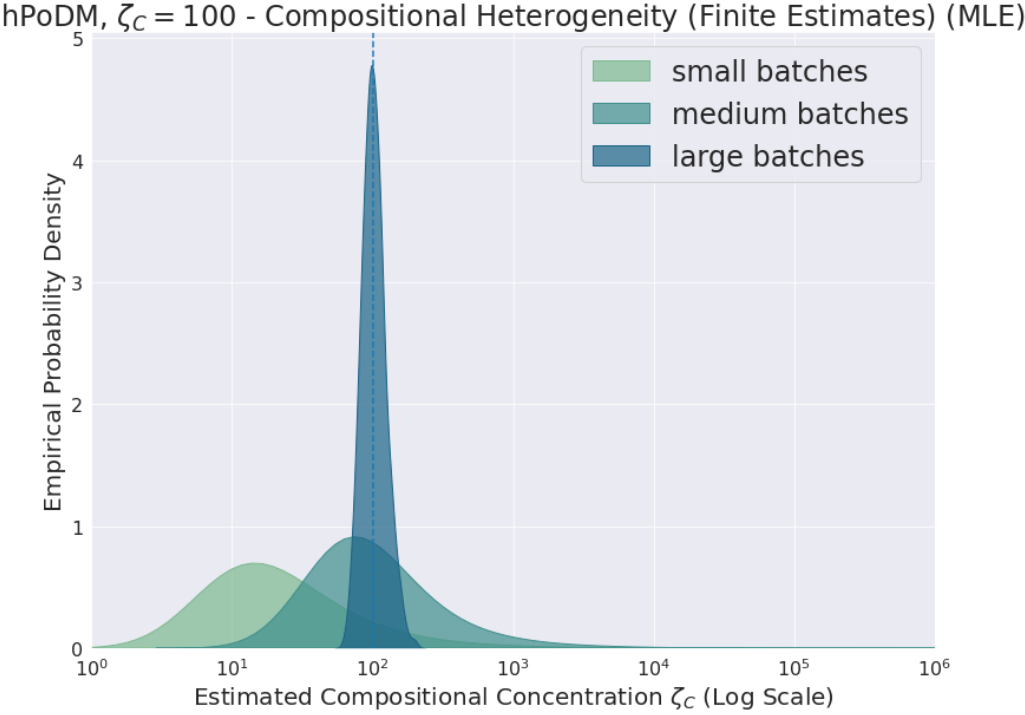
(b)

Figure 6.14: KDE plots of hPoDM ($\zeta_C = 100$) compositional concentration estimates.
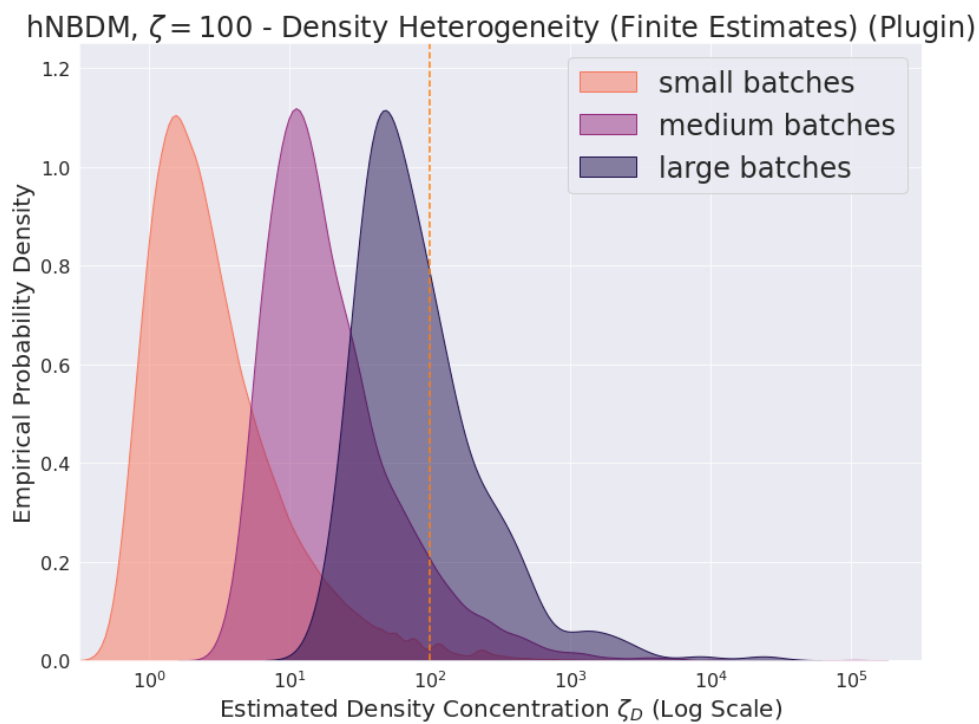
(a)
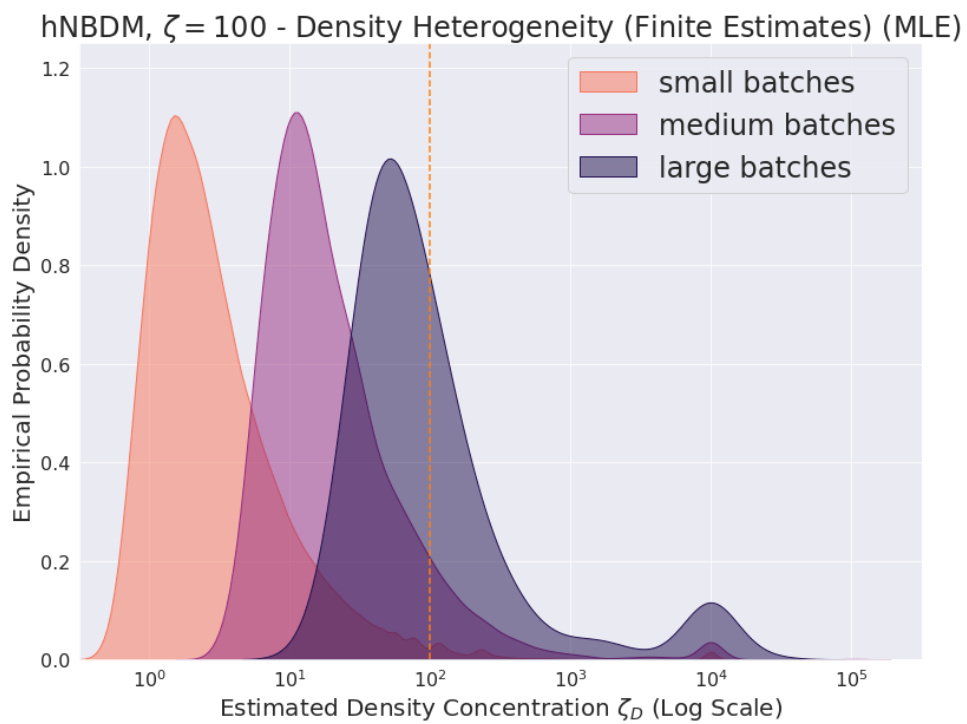
(b)

Figure 6.15: KDE plots of hNBDM ($\zeta = 100$) density concentration estimates.
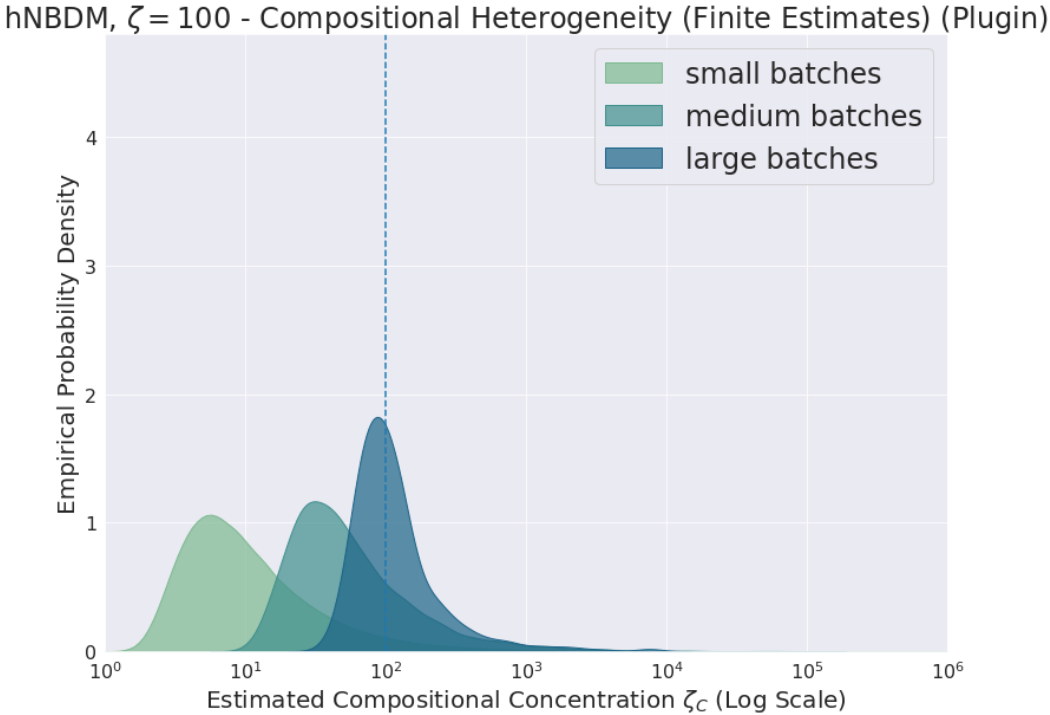
(a)
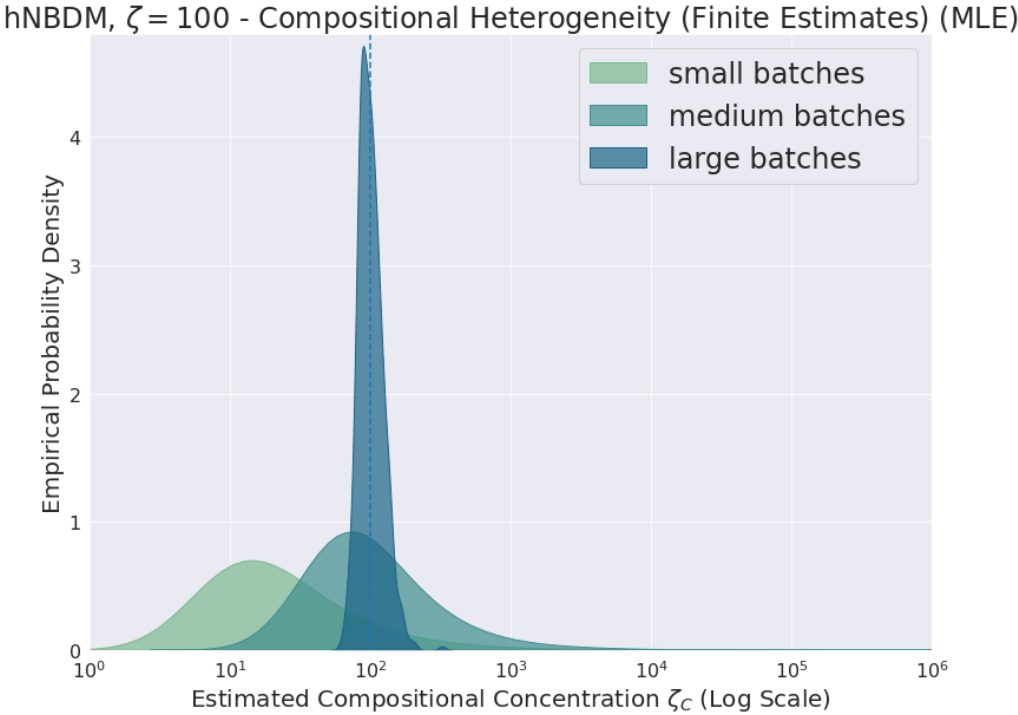
(b)

Figure 6.16: KDE plots of hNBDM ($\zeta = 100$) compositional concentration estimates.
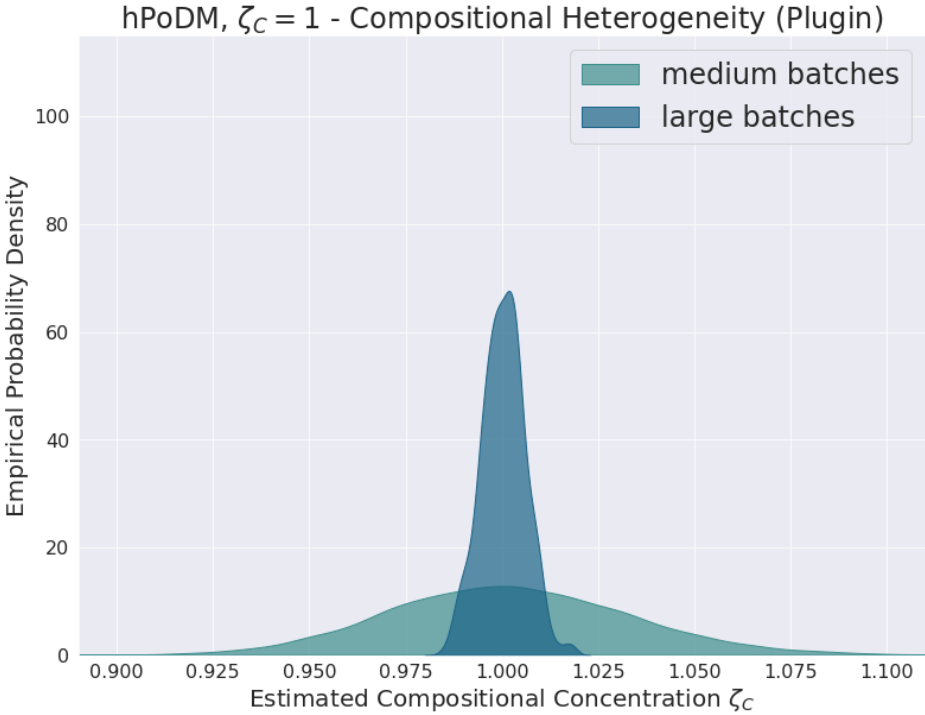
(a)
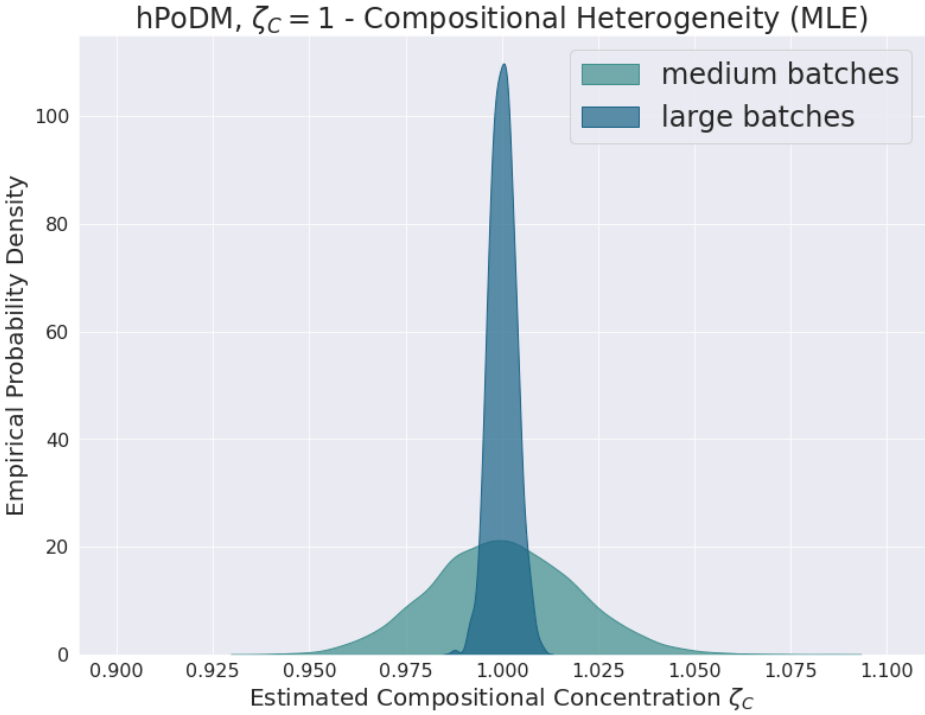
(b)

Figure 6.17: KDE plots of hPoDM ($\zeta_C = 1$) compositional concentration estimates.
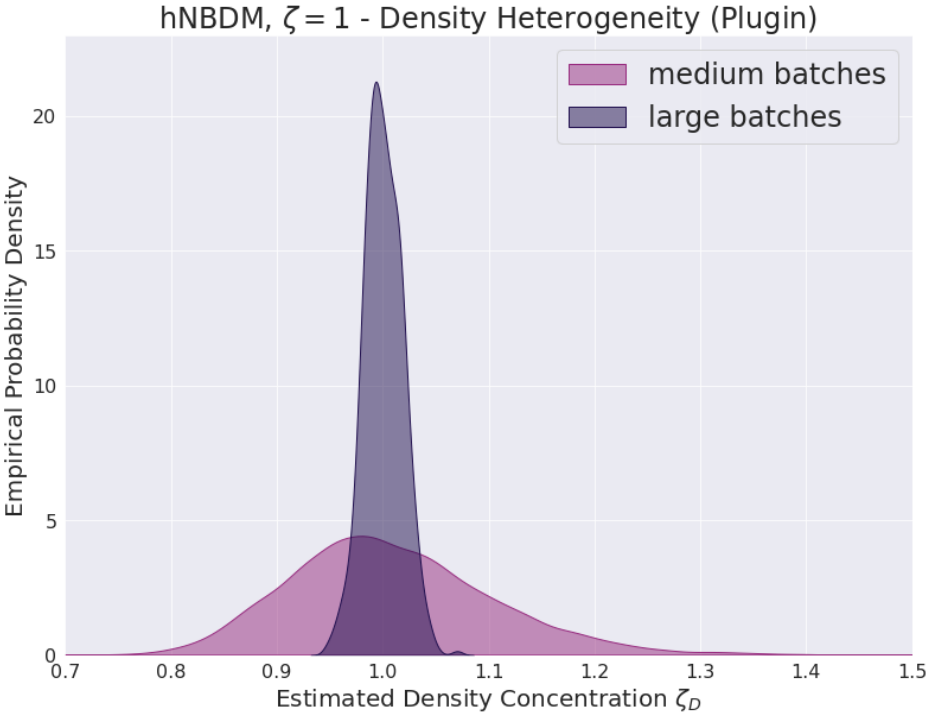
(a)
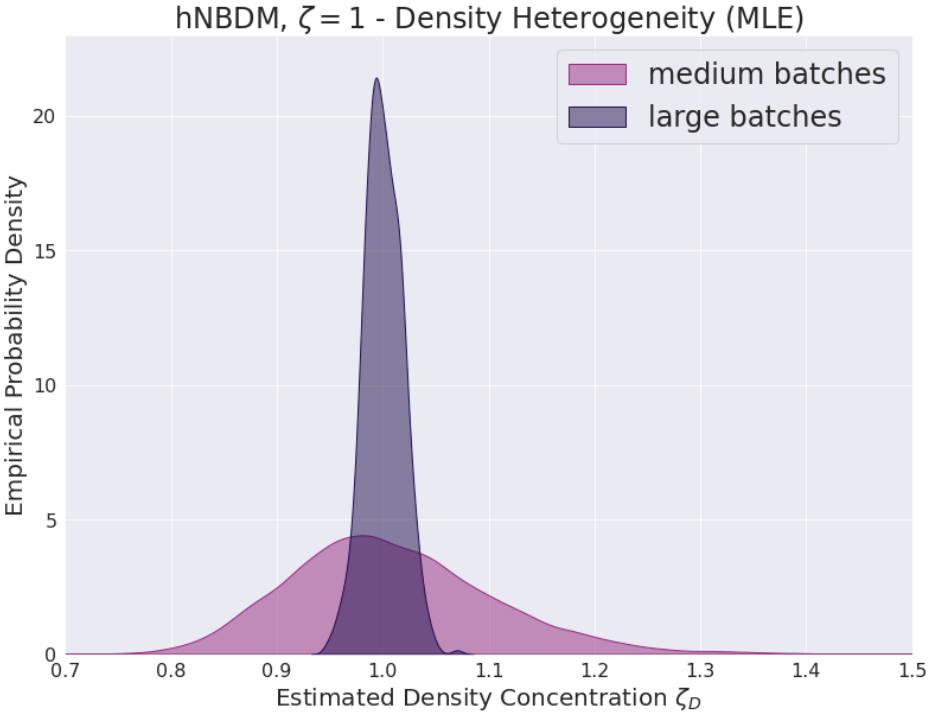
(b)

Figure 6.18: KDE plots of hNBDM ($\zeta = 1$) density concentration estimates.
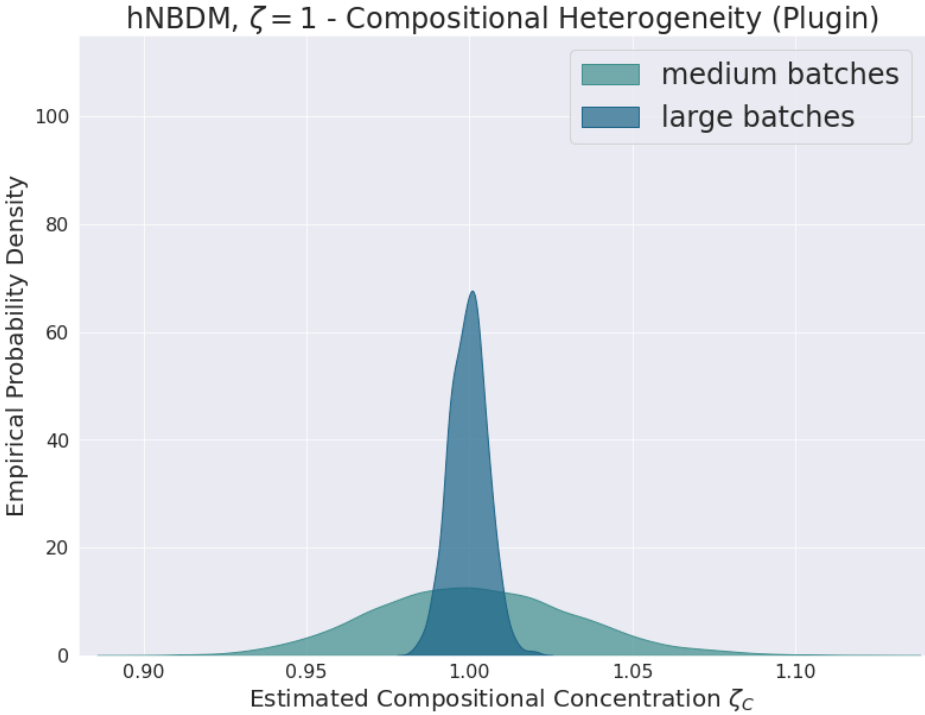
(a)
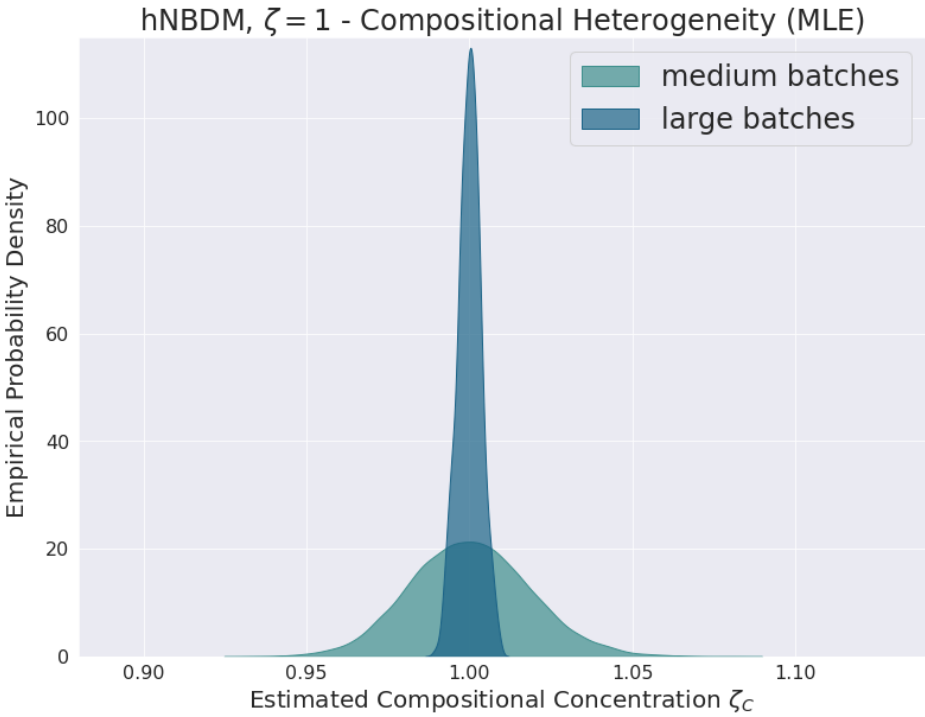
(b)

Figure 6.19: KDE plots of hNBDM ($\zeta = 1$) compositional concentration estimates.
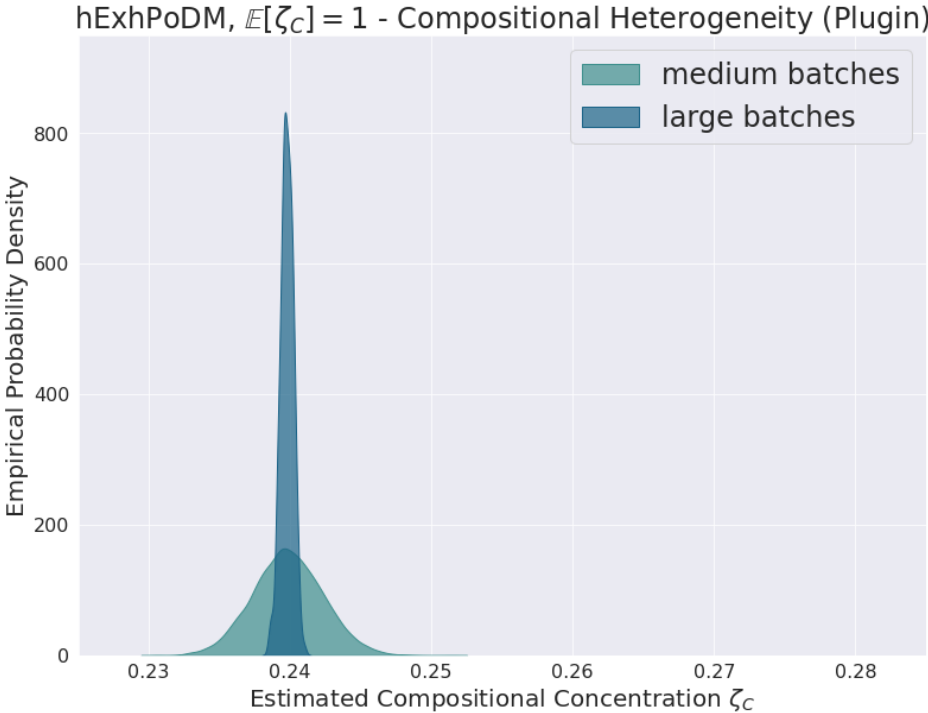
(a)
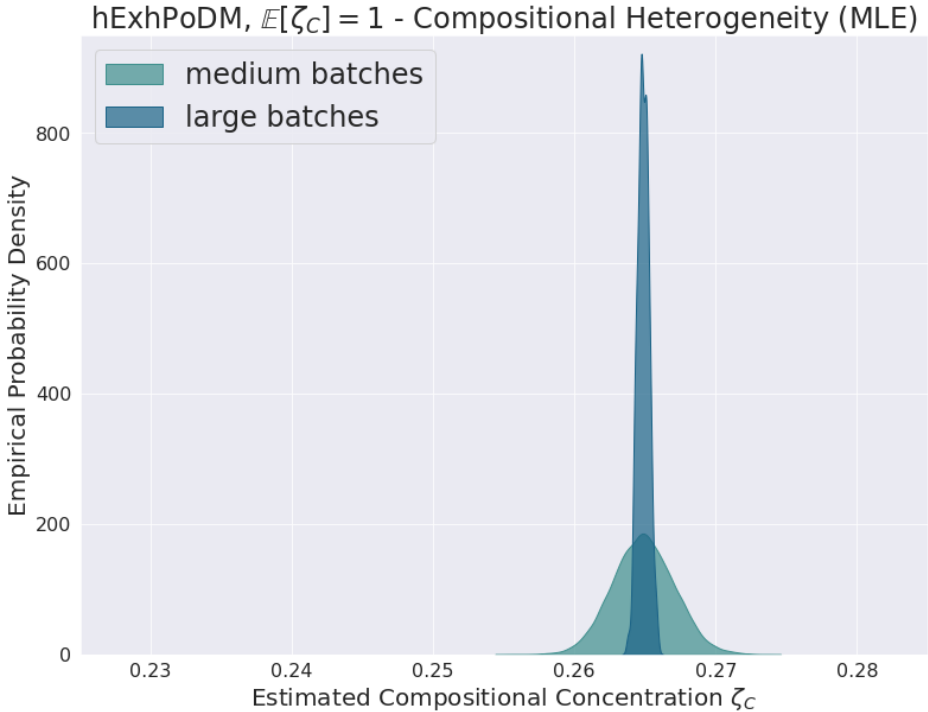
(b)

Figure 6.20: KDE plots of hExhPoDM compositional concentration estimates.

hEXhNBDM, $\mathbb{E}[\zeta] = 1$ - Density Heterogeneity (Plugin)

(a)

hEXhNBDM, $\mathbb{E}[\zeta] = 1$ - Density Heterogeneity (MLE)
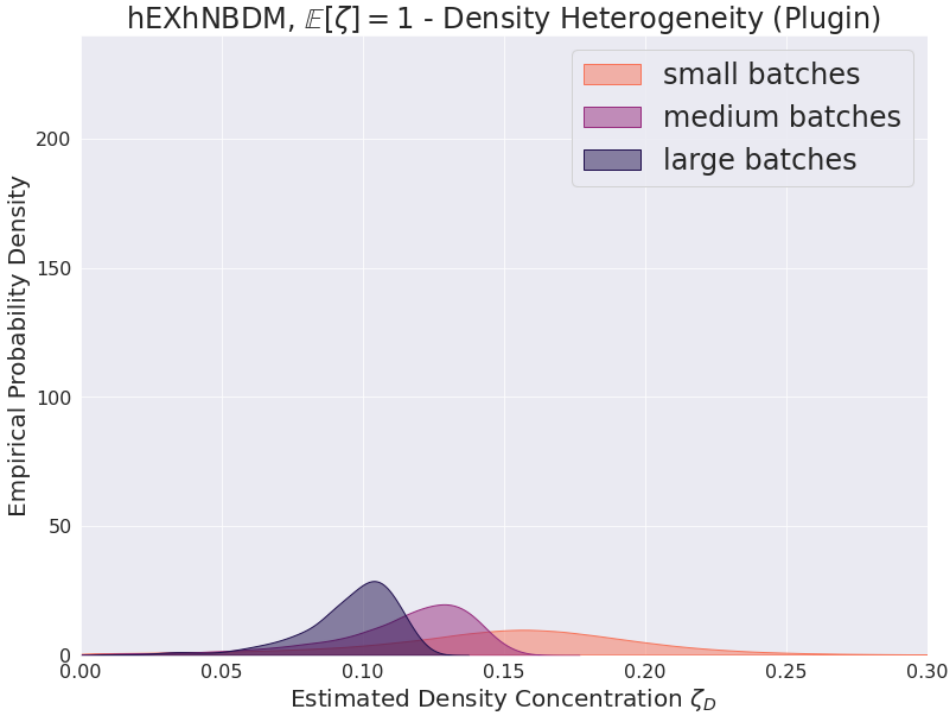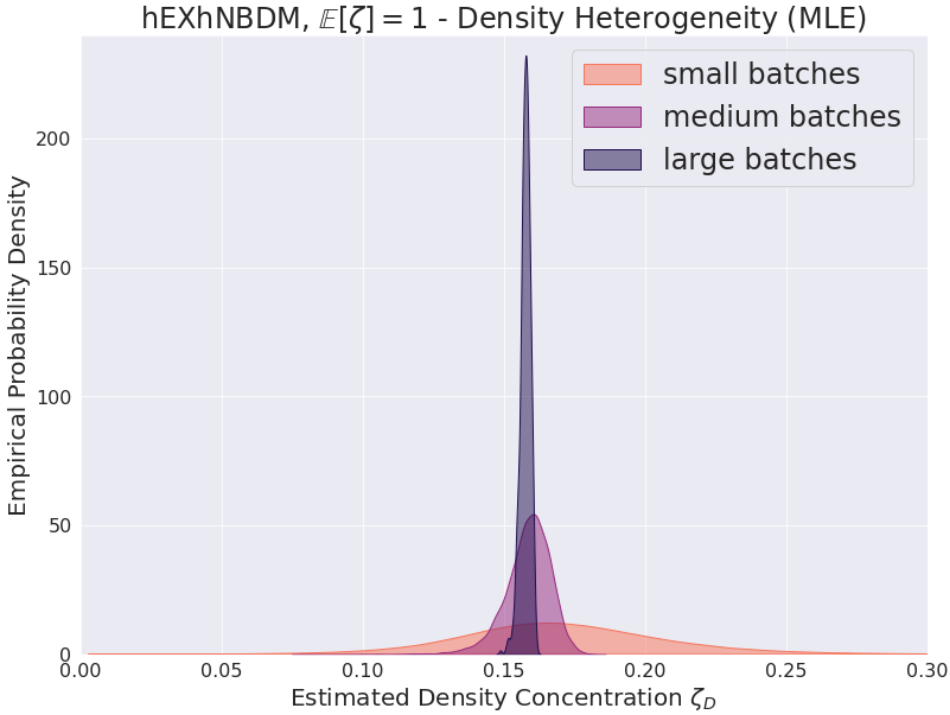
(b)

Figure 6.21: KDE plots of hExhNBDM density concentration estimates.
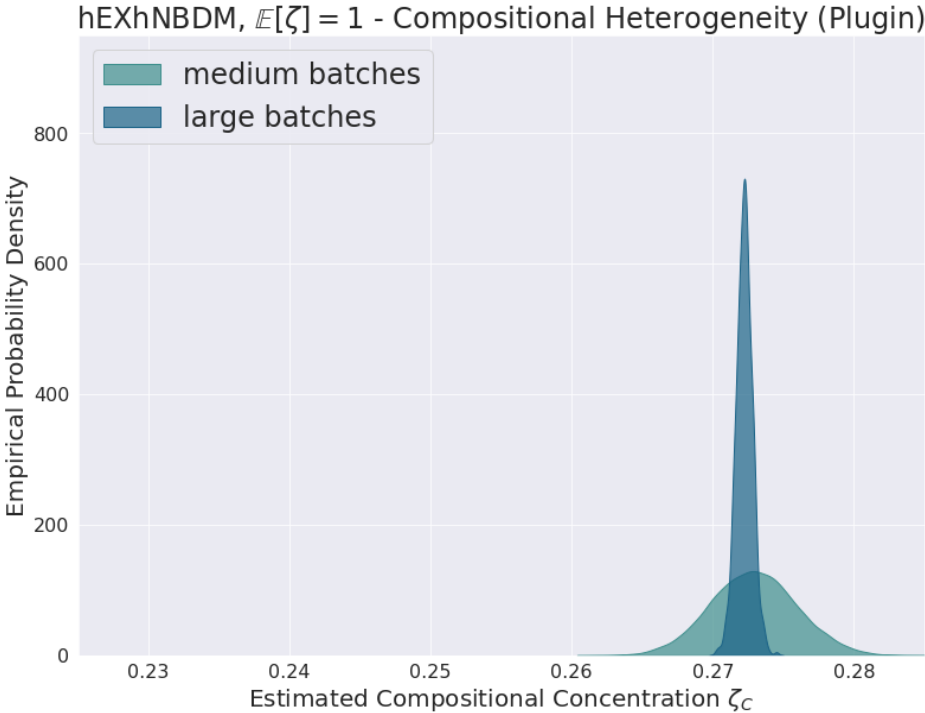
(a)
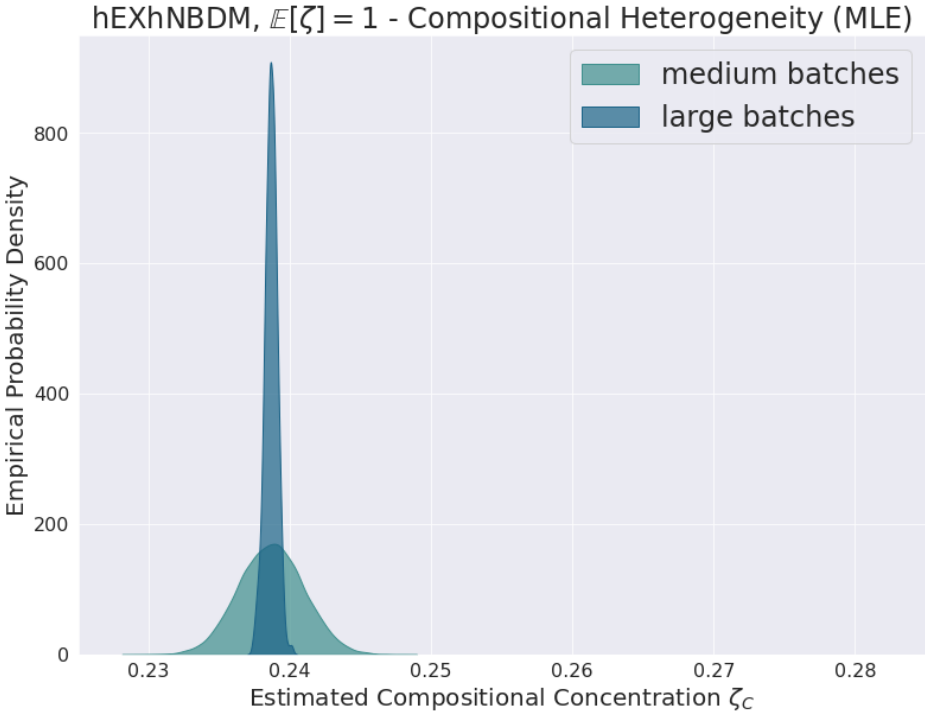
(b)

Figure 6.22: KDE plots of hExhNBDM density concentration estimates.

curves when the true value of the estimand is finite, figures 6.6a, 6.6b, 6.7a, 6.7b, 6.8a, 6.8b, 6.9a, and 6.9b for KDE plots when the true value of the estimand is infinite, and figures 6.15a, 6.15b, 6.18a, and 6.18b for KDE plots when the true value of the estimand is finite.

I double-checked the exact values and the estimates produced by the two estimators really are different, even though these differences are not apparent[14] from the aforementioned graphs. This makes sense given that the median difference in the estimated values was never greater than 3×10^{-4} (data not shown) and was often smaller (especially for the larger batches). Evidently the difference in the asymptotic efficiency (as computed in [Ans50]) of the negative binomial ML and plugin estimators is negligible for the chosen sample sizes.

Bizarrely, hExhNBDM appeared to be the only distribution for which the ML density estimator was noticeably more efficient than the plugin density estimator, despite (or perhaps because?) of the model misspecification. Compare figures 6.21a and 6.21b. I am unsure why this only occurred for hExhNBDM.

The bump in the density concentration estimates which consistently occurs around 10^4 for the ML estimates appears to be an artifact of the difficulties encountered when relying on Powell's method [Pow70] instead of Brent's method [Bre73] to find the root of the score, cf. the discussion in section 6.3.2. See figures 6.1b, 6.6b, 6.2b, 6.7b, 6.12b, 6.15b, 6.3b, 6.8b, 6.4b, and 6.9b. These bumps likely would not occur if the MLE was better implemented.

6.4.3.2 Precision of ML Estimator vs. Plugin Estimator for Compositional concentration

Unlike the situation for the density estimators, the compositional ML estimator is noticeably more efficient than the compositional plugin estimator for all of the distributions, regardless of whether the model was correctly specified.

See figures 6.5a and 6.5b for survival curves when the true value of the estimand is infinite. See figures 6.11a, 6.11b, 6.13a, and 6.13b for survival curves when the true value of the estimand is finite. Despite how difficult it usually is for most of the estimators to distinguish distributions with infinite concentration from those with concentration ≈ 100, for the large batches the ML estimator remarkably produced almost no infinite estimates. The increased efficiency of the ML estimator is evident from the more rapid decrease compared to the plugin estimator of the survival curves when the true value of the estimand is finite.

See figures 6.10a and 6.10b for KDE plots when the true value of the estimand is infinite. The greater efficiency of the ML estimator is reflected by the distribution of finite estimates shifting further to the right more quickly. When the true value of the estimand is finite, as well as when the model is misspecified, the greater efficiency of the ML estimator compared to the plugin estimator is reflected by the narrower distribution of estimates. See figures 6.14a, 6.14b, 6.16a, 6.16b, 6.17a, 6.17b, 6.19a, 6.19b, 6.20a, 6.20b, 6.22a, and 6.22b.

---

[14]Except for the bumps around 10^4, see below.

6.4.3.3 Behavior of ML vs Plugin Estimators Under Misspecification

Under the misspecified[15] models hExhPoDM and hExhNBDM, the effect on the limit of switching from plugin to ML was opposite for hExhPoDM compared to hExhNBDM. For the compositional concentration, switching from the plugin estimator to the ML estimator *increased* the limiting value of the estimates for hExhPoDM (cf. figures 6.20a and 6.20b), while for hExhNBDM switching from the plugin estimator to the ML estimator *decreased* the limiting value for the estimates (cf. figures 6.22a, and 6.22b). For hExhNBDM the effect of switching from the plugin estimator to the ML estimator for the density concentration was opposite the effect for the compositional concentration; the limiting value of the estimates *increased* (cf. figures 6.21a and 6.21b). That the plugin and ML estimators converge to different values from each other under these distributions reflects the model misspecification.

6.5 Discussion

Section 6.5.1 discusses why the results from section 6.4 are important for showing that estimates from the estimators will be useful in practice. Section 6.5.2 discusses how the results from section 6.4 suggest that future work might be able formulate density and compositional heterogeneities in a nonparametric context. Section 6.6 discusses how the results from section 6.4 bring us one step closer to the final goal of this thesis.

6.5.1 The Estimators are Good Enough for Use in Practice

The results show that the proposed estimators both (1) converge towards correct answers as the data size increases and (2) have decreasing variance as the data size increases. Hence they satisfy the minimal requirements one expects of "adequate" statistical estimators.

The estimators were already guaranteed to be consistent asymptotically when the statistical model was correctly specified. Nevertheless, that theoretical guarantee does not necessarily translate in practice into estimates usefully approximating the truth for realistic data sizes. (In this case realistic data sizes correspond to the medium or large batches.) Such a failure of "practical consistency" occurred, for example, for the version of the plugin compositional concentration estimator that equally weighted the estimates resulting from all pairs of strains. Consistency of an estimator need not imply efficiency of an estimator. The above results show that the proposed estimators are in fact "practically consistent", not merely asymptotically consistent. Not only that, but they often still managed to give useful results even for unrealistically small data sizes (the small batches).

---

[15] Cf. footnote 13 above.

6.5.2 Heterogeneity Might be Nonparametric

The plugin estimators also gave sensible results in the cases of model misspecification. This suggests that functions of moments, which are at least similar to the "effective density concentration" (6.2.12) and "effective compositional concentration" (6.2.27) that these estimators converge to, may be useful for describing qualitative notions of "density heterogeneity" and "compositional heterogeneity" for general[16] count-categorical distributions. Exploring how to define such notions precisely for more general distributions outside of the ghNBDM family would be a worthwhile goal for future work. This idea resembles the argument from [Nak97, section 4] that the negative binomial plugin estimator is "semi-parametric".

It is worth noting however that, at least for hExhPoDM and hExhNBDM, the ML estimators were also "unreasonably effective" at giving sensible results even under model misspecification. Thus we might anticipate either that (1) a "nonparametric" or "semi-parametric" interpretation of their estimands also exists, or (2) that the ML estimators would not prove equally "unreasonably effective" for other misspecified models. [Nak97] appears to argue the latter, at least with respect to the density concentration estimator. Unfortunately this work does not investigate the behavior of these estimators under misspecification enough to distinguish between these two possibilities, cf. footnote 2.

6.6 Conclusion

Practical Implications This chapter shows that we have viable statistical estimators for density heterogeneity and compositional heterogeneity. Therefore, neither quantity needs to be treated as an "unfathomable unknown". Instead, we are able to actively account for them when simulating or analyzing the data from MOREI. Given real-world data, even of only moderate size, we can get useful estimates for concentration parameters describing realistic levels of heterogeneity. Using these estimates as the parameter values for our simulations, we can make simulations of MOREI more realistic. This in turn enhances our ability to find the methods that best recover true interaction networks from MOREI data.

Next Steps and Open Questions An obvious next step would be to clarify the framework for the non-parametric definitions of heterogeneity. From there one could investigate these heterogeneity estimands for a more diverse class of multivariate count distributions than I have considered herein, and then evaluate the performance of various estimators for these within this more general context. For example, giving these estimands a clear definition as pathwise differentiable functionals (or, if this is not possible, then explaining why) would facilitate implementing the targeted learning framework for them.

---

[16]It probably makes sense only to attempt to define such notions for distributions with finite second moments, or to at least automatically say that any distribution lacking finite second moments is "infinitely heterogeneous" without exploring further.

Chapter 6.   Appendices

Appendix 6.A   Plugin Estimators for Strain Population Frequencies

When $\mathbb{E}[\vec{\mathbf{N}}(0) \mid N(0)]$ is Multinomial or Dirichlet-Multinomial distributed, one has that for all $n \geq 1$ and all strains $s \in [S]$:

$$f^{(s)} = \frac{\mathbb{E}[N^{(s)}(0) \mid N(0) = n]}{n}. \tag{6.A.1}$$

Averaging over all values of $n \geq 1$ leads to

$$f^{(s)} = \sum_{n \geq 1} \frac{\mathbb{E}[N^{(s)}(0) \mid N(0) = n]}{n} \mathbb{P}(N(0) = n \mid N(0) \geq 1) =: (f^{(s)})_{eff}. \tag{6.A.2}$$

This leads to the following plugin estimators for the frequencies:

$$\hat{f}^{(s)} = \sum_{n \geq 1} \frac{\hat{\boldsymbol{\mu}}[N^{(s)}(0) \mid N(0) = n]}{n} \cdot \frac{\sum_d \mathbb{I}\{N_d(0) = n\}}{\sum_d \mathbb{I}\{N_d(0) \geq 1\}}. \tag{6.A.3}$$

6.A.1   Effective Frequencies

Even when the left and right hand sides of (6.A.2) are not equal, the plugin estimator (6.A.3) is still consistent for the right hand side of (6.A.2), the "effective frequency" $(f^{(s)})_{eff}$.

Note that for all of the distributions used in the simulations, the true frequencies *do* equal the effective frequencies. Even for hExhNBDM and hExhPoDM this is true, because

$$f^{(s)} = \int_{\xi > 0} f^{(s)} \mathbb{P}(\zeta_C = \xi) d\xi = \int_{\xi > 0} (f^{(s)})_{eff} \mathbb{P}(\zeta_C = \xi) d\xi = (f^{(s)})_{eff}. \tag{6.A.4}$$

In other words, (6.A.2) holds for any individual value of ζ_C and will therefore still hold after averaging over any choice of "prior" for ζ_C.

For distributions such that $f^{(s)} \neq (f^{(s)})_{eff}$, the estimators (6.A.3) will no longer be consistent, and a different choice of estimators $\hat{f}^{(s)}$ which actually are consistent for the true

frequencies $f^{(s)}$ will need to be used inside of (6.2.26) to guarantee that (6.2.25) remains consistent for the effective compositional concentration (6.2.27).

Alternatively, one could instead change the definition of the effective compositional concentration inside of (6.2.28) to use the effective frequencies $(f^{(s)})_{eff}$ in place of the true frequencies $f^{(s)}$, in which case using (6.A.3) inside of (6.2.26) will continue to ensure that (6.2.25) is consistent for (6.2.27). Cf. section 6.2.3.6.

Appendix 6.B   Convergence Sometimes Disappointing for Small Batches

The variance of the distribution of estimates was sometimes much wider for the small batches than for the medium or large batches. This is unimportant in practice since the small batches represent an unrealistically small problem size.
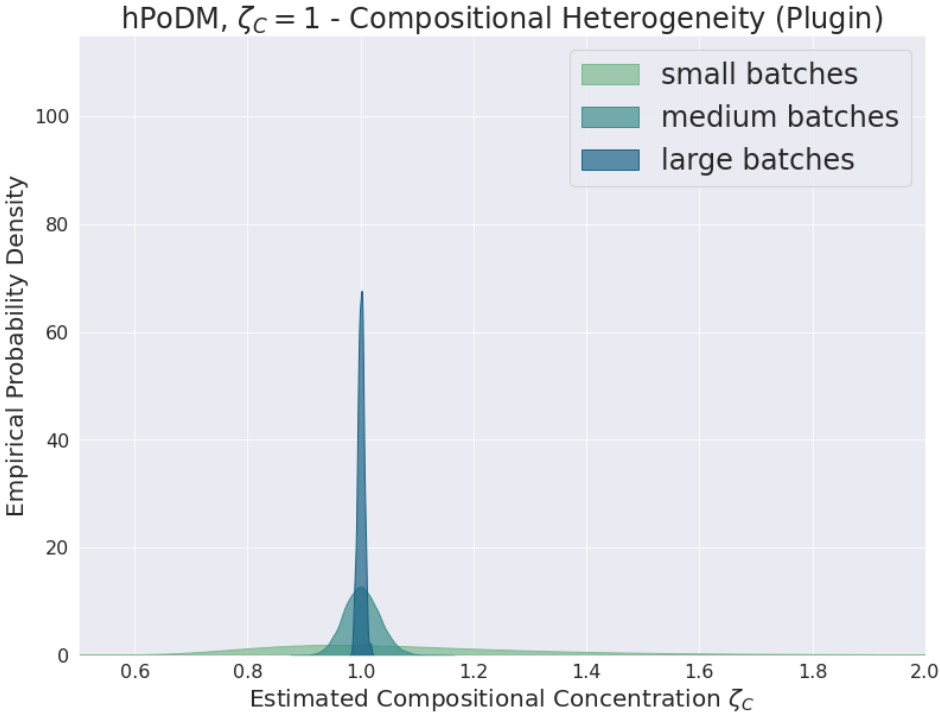
(a)
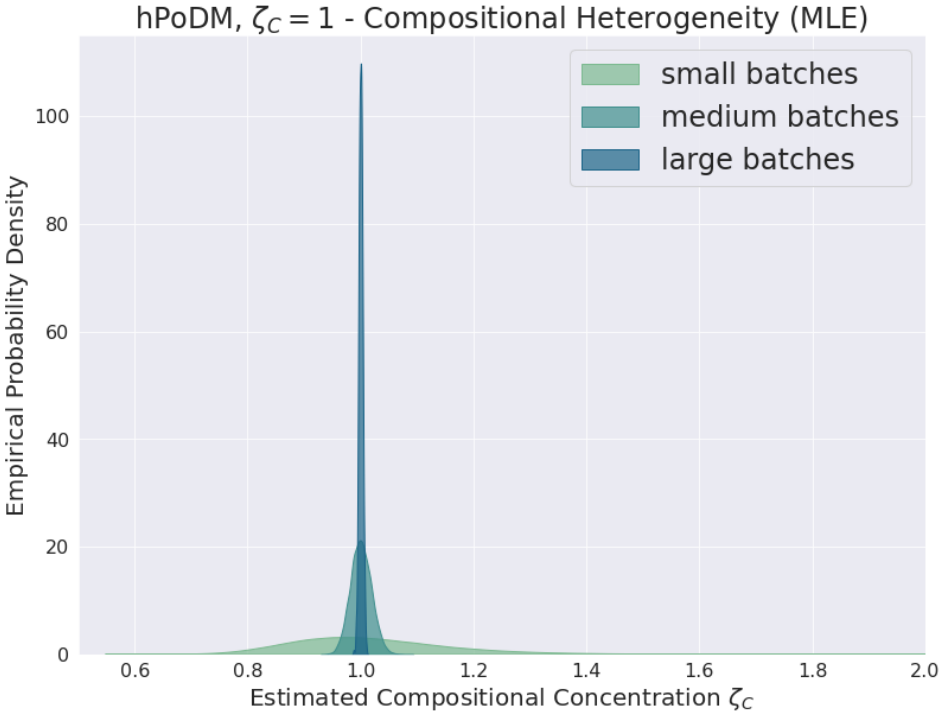
(b)

Figure 6.B.1: Same as figure 6.17, except now including small batches.
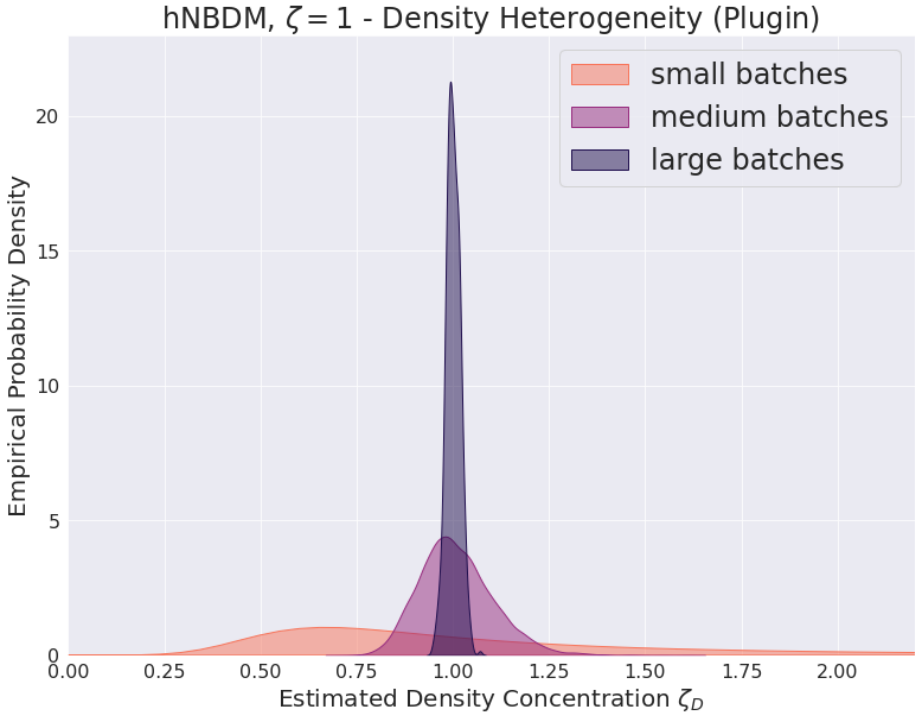
(a)
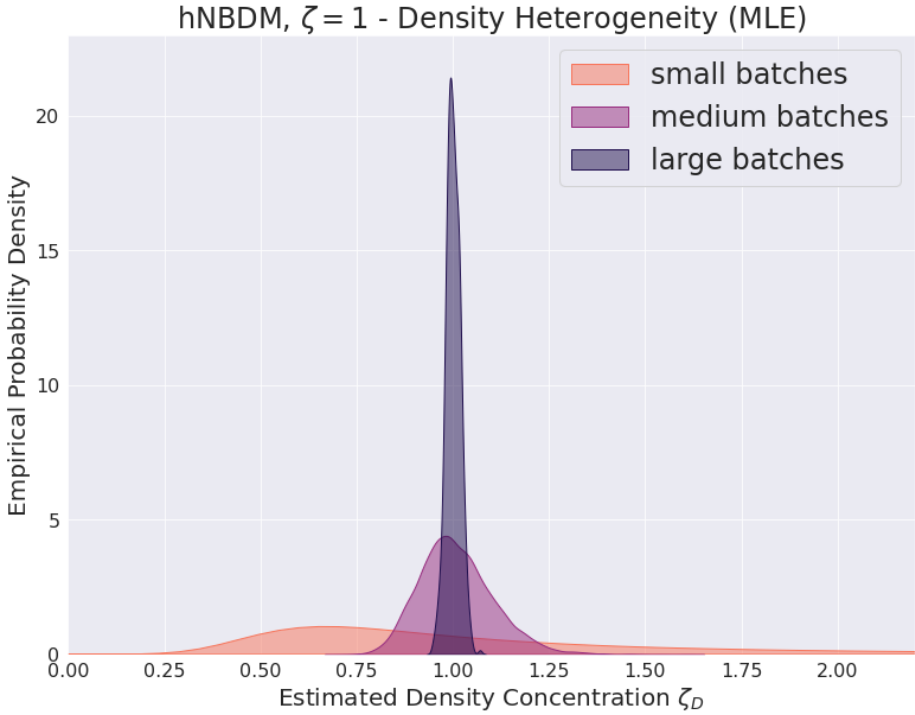
(b)

Figure 6.B.2: Same as figure 6.18 , except now including small batches.
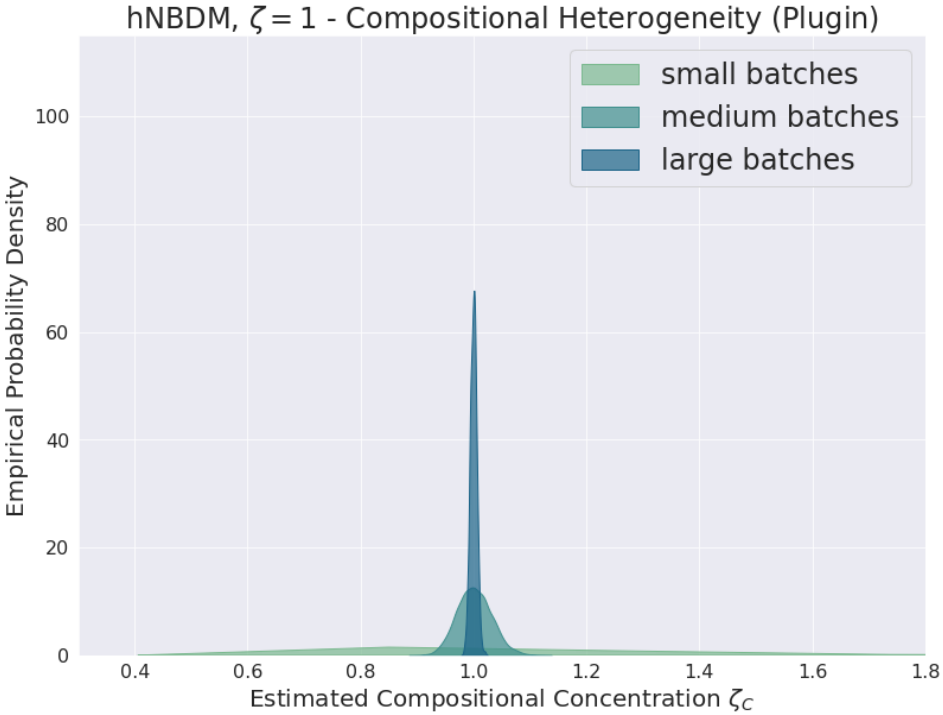
(a)
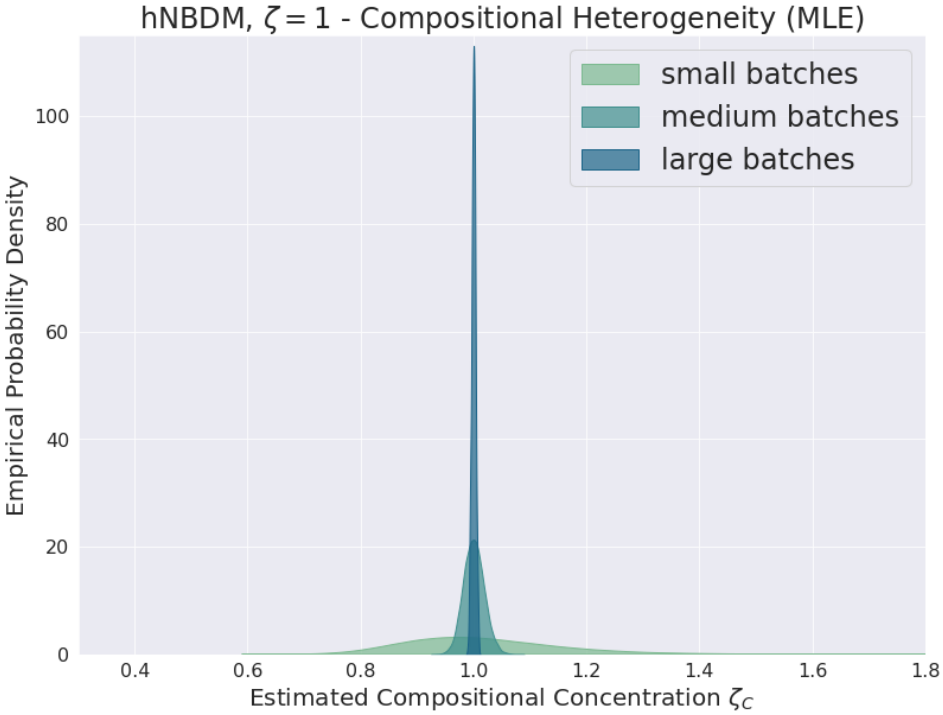
(b)

Figure 6.B.3: Same as figure 6.19 , except now including small batches.
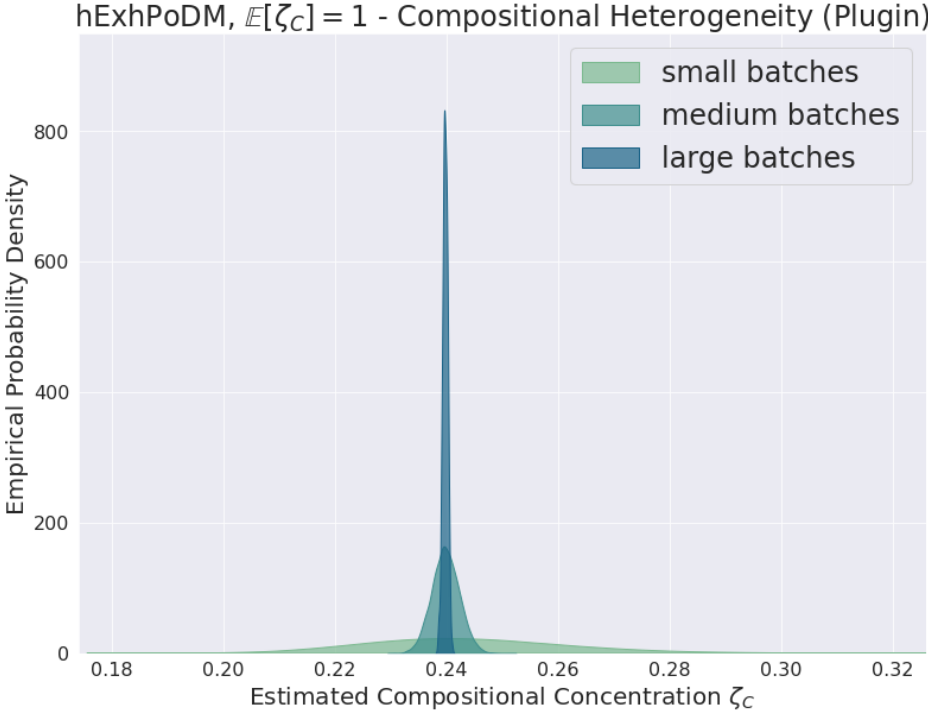
(a)
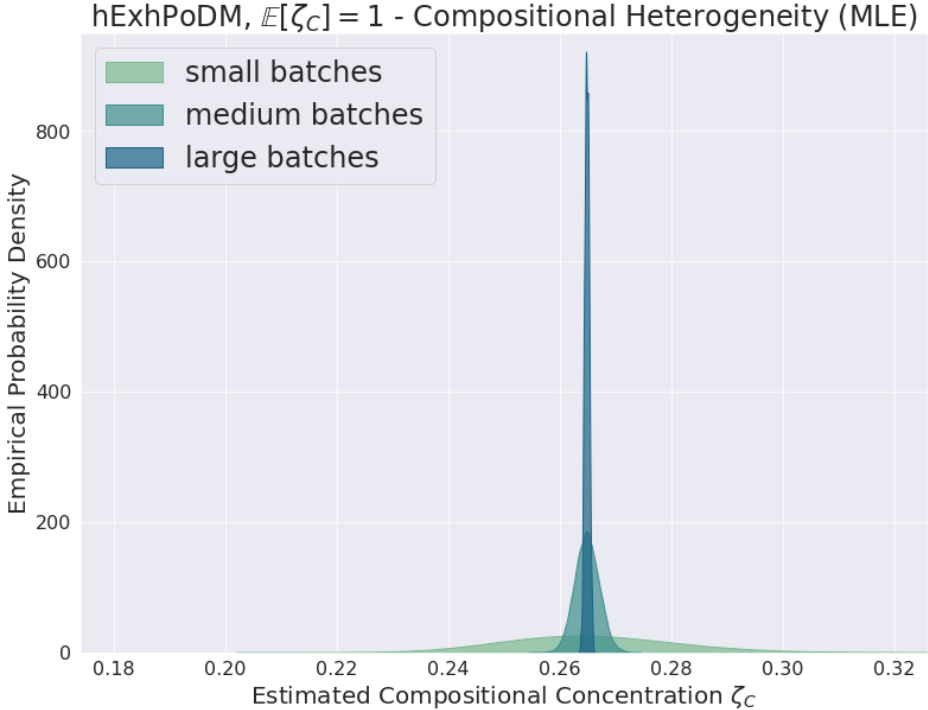
(b)

Figure 6.B.4: Same as figure 6.20, except now including small batches.
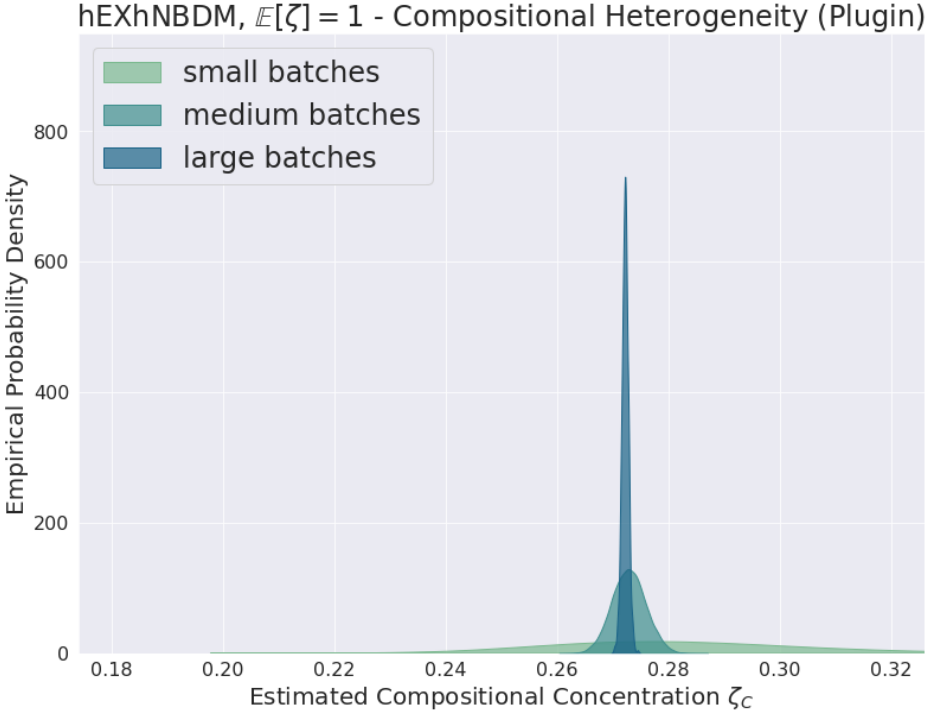
(a)
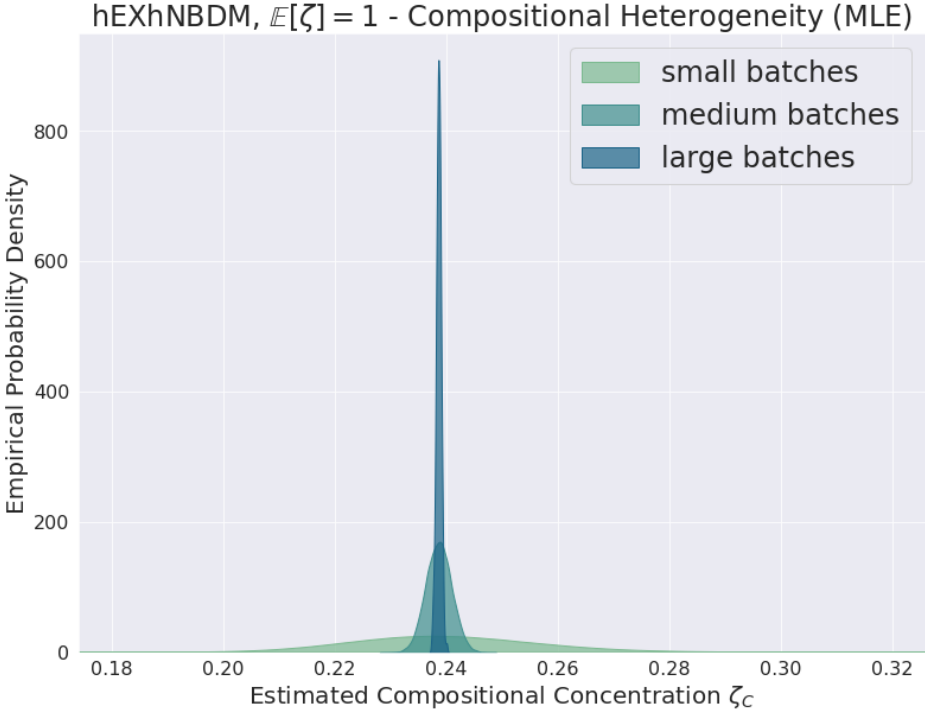
(b)

Figure 6.B.5: Same as figure 6.22, except now including small batches.

Part III

Inference of Ecological Networks

The broader field of Part III, and an important problem in microbial ecology, is inference of ecological models of microbial communities [AGH+17]. To infer ecological models of microbial communities, we need to understand the interactions between different microbes. (Cf. again the discussions from before in sections 2.2.5, 1.2, and 1.1.) It is commonly understood that networks are useful models of ecological interactions[1] [FR12] [AML+17]. Likewise, it is also commonly understood that dynamical systems are useful models of ecological interactions [MXS17] [XAF+17] [AML19]. Finally, at least as early as [Lid79], it has been commonly understood that directed ecological interactions can be categorized using signs, with "+" for beneficial effect, "0" for neutral effect, and "−" for detrimental effect on the growth of the recipient organism [GHG+18] [KOK+20]. By combining all of these observations, it is straightforward to conclude that signed networks [Har53], networks whose edge weights can be positive or negative, are natural models[2] of ecological interactions [XAF+17].

Hence Part III investigates the important subfield of how to infer ecological models that can be characterized as signed networks. The signs of edge weights have the same interpretation that signs have above [Lid79] [GHG+18] [KOK+20]. Namely, the signs of edge weights categorize the qualitative effects that ecological interactions have on the growth of recipient organisms (and the magnitudes of the edge weights encode how strong these effects are) [XAF+17]. (Cf. again section 1.2.) Competition assays [WL15] and related coculture experiments[3] are ideal for measuring how different microbes affect each other's growth.

Therefore in microbial ecology the problem of inferring signed networks that encode interactions often reduces to the problem of inferring signed networks from the data produced by competition assays and related coculture experiments. This is why the specific problem that Part III investigates is the inference of signed networks quantifying how microbes affect each other's growth. This problem corresponds to the "unpartitioned interaction network" competition assays from the introduction to Part I, or the equivalently to the "genes × genes interaction problem" from item (4) of section 2.1.1. The reference [GFP14] gives an overview (as of 2014) of several of the technologies used for implementing such experiments. Even experiments such as [VCF+18] that use fairly straightforward cell plating have made progress in understanding the ecological interactions of microbes.

Yet even if for no other reasons than those of throughput (cf. again section 1.3), the future state-of-the-art experiments for inferring microbial interactions (via their effects on each other's growth) are likely to use droplet-based microfluidics technology. See the articles [KvVDH10] or [GRHW12] for reviews (as of 2012) of high-throughput biological experiments using such droplet-based microfluidics technology. Some work using droplet-based

---

[1] Of *pairwise* ecological interactions. Cf. [BKK16] or [MXS17] regarding "higher-order" interactions.

[2] In the case one starts with a dynamical system model, one can define the edge weights of a corresponding signed network model as coefficients of the dynamical system [GCLF18], or as the values of its Jacobian at some point in time [XAF+17] [AML19]. Cf. the discussion towards the end of section 1.1.

[3] Such experiments (especially those involving cell plating or fluorescence imaging) usually produce absolute abundance data, but cf. also e.g. [BEG+18] and [GHG+18] for ideas on how competition assays might still be used to infer microbial interactions even when using relative abundance data. See also sections 1.3 and 1.4.3 for more on the controversies surrounding absolute vs. relative abundance data.

microfluidics to study microbial interactions has already been published, cf. e.g. [KKO$^+$19] [KOK$^+$20] or [HCT$^+$19], but it is expected (cf. again section 1.4) that in the future the data produced by such experiments for studying microbial interactions will improve even further.

Chapter 7 The broader field chapter 7 belongs to is the inference of ecological models for microbial interactions. The subfield that chapter 7 considers is the use of relative fitness measurements to quantify microbial interactions. The chapter investigates how a particular measure of relative fitness can be recast into the statistical framework of average treatment effects. The chapter explains how violations of positivity assumptions are inevitable for this problem, making controlling for confounding difficult. Finally, chapter 7 also gives explicit assumptions under which the estimands are identifiable from the observed data produced by incubated droplets, even though initial states of the droplets are unobserved.

Chapter 8 The broader field of chapter 8 is the use of signed networks as models for ecological interactions. In particular, chapter 8 is motivated by the important issue of how to identify "best-performing" estimators of such ecological models. This chapter investigates how comparisons of ecological interactions can be recast into the statistical framework of loss functions for signed networks. The chapter explains how avoiding unexpected behavior requires loss functions for signed networks to satisfy what I call herein "the double penalization principle". Starting from examples of loss functions for **un**signed networks, the chapter derives several examples of loss functions for signed networks that satisfy this property.

Chapter 7

Average Treatment Effects (ATEs) for Quantifying Microbial Interactions

Herein I show how a particular measure of relative fitness can be recast into the statistical framework of average treatment effects. See sections 7.3 and 7.5. I explain how (for this problem) violations of positivity assumptions are inevitable in practice, making controlling for confounding difficult. See section 7.7. I give explicit assumptions under which the corresponding estimands are identifiable from the observed data produced by incubated droplets, despite that initial states of the droplets are unobserved. See section 7.8.

Section 7.2 clarifies what the statistical model for the full data is assumed to be. Section 7.3 then clarifies which parameters of the full data statistical model we can choose to target for quantifying the strength and directionality of microbial interactions. Section 7.4 establishes what the statistical model corresponding to the observed data is, highlighting the difference from the full data statistical model that is most important for our choice of target parameters. Section 7.5 expands on this by giving the target parameters for the observed data statistical model that we would like to be able to identify with the target parameters of the full data statistical model. It also briefly states sufficient conditions for such an identifiability result to hold. All in all, the goal of these four sections is to clarify what research question we are choosing to answer, thus firmly positioning our approach to answering our given scientific question (quantifying the strength and directionality of microbial interactions) within the first stage of the targeted learning roadmap [vdLR11].

Section 7.6 briefly clarifies what the definitions of the relevant plugin estimators for the target parameters of the observed data model would be, and offers an initial explanation of how, starting from the previously established research question, we can proceed further to the estimation stage of the targeted learning roadmap using SuperLearner [vdLPH07] and targeted minimum-loss based estimation (TMLE) [vdLR11]. Section 7.7 clarifies which issues with confounding we might potentially encounter with our observed data, and indicates some possible obstacles and preliminary solutions for addressing those issues. Finally, section 7.8 gives a detailed clarification of the assumptions being used to establish sufficient conditions for identifiability of the full data target parameters from the observed data. It then provides

proofs demonstrating why these conditions are indeed sufficient for the intended results. The goal of these last sections is to demonstrate the feasibility of the remaining two stages of the targeted learning roadmap [vdLR11], estimation and inference, for this research question.

To the extent it is necessary either as a mathematical convenience, or for peace of mind, throughout what follows it can be assumed that the logarithm has had its domain extended from $(0, \infty)$ to $[0, \infty)$ by defining $\log(0) = 0$.

Chapter 7 Contents

7.1   Background and Significance

Broader field   The broader field chapter 7 belongs to is the inference of ecological models for microbial interactions. The goal is to use data produced by the MOREI experiment to infer these. See [FR12] for a review of how networks are often used as such ecological models, and [XAF+17] (or section 1.2) for how signed networks in particular can be used to describe microbial interactions. Cf. the discussion earlier from the introduction to Part III.

The subfield that chapter 7 considers is the use of relative fitness measurements to quantify microbial interactions. (The ecological model is then a signed network whose weights are the corresponding relative fitness values. Cf. section 1.3 or the discussion of "competition assays" [WL15] from the introduction of Part I.) There are many distinct estimands used as quantitative measures of "(relative) fitness" found throughout the literature. Cf. the definitions found in [WPW+15], [WL15], [BEG+18], or [HRM+19] for several distinct examples. Definitions are often stated implicitly or imprecisely, contributing to the difficulty of comparing definitions across papers and of analyzing associated statistical problems.

Specific problem   Herein I investigate how a particular measure of relative fitness can be recast into the statistical framework of average treatment effects. The underlying idea behind this measure is easy to motivate heuristically and very simple (cf. section 7.C). Similar features of other definitions of relative fitness perhaps encourage their primarily informal treatment in the literature. In any case, by making this definition precise (cf. section 7.3), analysis of the corresponding statistical problems becomes much clearer. In particular, one can readily see that this definition admits an interpretation as an average treatment effect estimand. Cf. [vdLR11, chapter 10] for more about the general theory of average treatment effects, which is substantially well-developed.

At least two papers in the virology literature that study[1] bipartite interactions, [MKBDB00] and [HdlTCD91], appear to use the same (or a sufficiently analogous) definition of relative fitness as that used in this chapter (cf. section 7.B.5.1 for the latter.) This is such a straightforward definition that most likely other examples of its use exist in the literature, including those studying non-viral microbes and/or unpartitioned interactions. However, I have not yet been able to find additional examples, in part due to how definitions of "fitness" in the literature are usually discussed only imprecisely or even implicitly (as in e.g. [HdlTCD91]).

---

[1] As opposed to the unpartitioned interactions that we are interested in, cf. the introduction to Part I for terminology.

Particular approach Being able to formulate this problem in terms of average treatment effects suggests it should be possible to apply targeted minimum loss based estimation (TMLE) [vdLR11] to this problem. This chapter demonstrates how one can follow the first two stages of the targeted learning roadmap [vdLR11, section 1.5] with this choice of estimand: (1) defining the research question, and (2) estimation. Hence this chapter can be thought of as a "proof of concept" for applying TMLE, but future work will fill out all of the details of completing the targeted learning roadmap from start to finish.

Because PhenoPath [CY18] was designed for a problem involving fairly high-throughput data produced using microfluidics, and which is not strictly longitudinal, attempting to modify it to apply to the problem of inferring microbial interactions from MOREI data seems like a natural approach, and more sophisticated than using these simple relative fitness estimates. However, although PhenoPath was suitably modified to apply to this problem, the results of preliminary work suggest that PhenoPath does not work well for this problem.

One of the reasons the method from [CY18] failed to work well on (simulated) data for this problem in preliminary work is because it was designed for "gene × environment" interactions, not "gene × gene" interactions. (Cf. again item (4) from section 2.1.1 for terminology.) Its computational performance for this latter, more computationally demanding, problem proved inadequate. (If n is the number of genes, the latter problem corresponds to n instances of the former problem, cf. footnote 2 from section 2.1.1.)

The method from [CY18] also seemed unable to handle the sparsity of the data well, much less exploit it. One reason seems to be that the model implemented by [CY18] inherently treats 0 values as "biological zeros", when in fact for this problem they are almost all "sampling zeros" (cf. the terminology from section 7.A and quoted from [PGB21]). More often than not, when a given strain is not found within a given droplet, the reason is because the strain was never in the droplet in the first place. Yet PhenoPath appears to always interpret this the same way as if the strain had been initially present in the droplet and later killed off by those strains that were found in the droplet.

7.2   Full Data Model

The full data model $\mathcal{M}^F$ corresponds to the likelihoods from equations (2.2.5), (2.2.8), and (2.2.12). This can be interpreted as the marginalization of a continuous time Markov process describing the growth and death of cells over time. The full data model $\mathcal{M}^F$ assumes for any given droplet not only that $\vec{\mathbf{N}}(t)$ is known, but also that $\vec{\mathbf{N}}(0)$ is known as well.

7.3 Full Data Average Treatment Effect (ATE) Definitions

For the full data model $\mathcal{M}^F$, denote any average treatment effect (ATE) full data estimand (i.e. assuming that $\vec{\mathbf{N}}(0)$ is known) corresponding to the effect of strain[2] s_1 on s_2 at time t by $\Psi^F(t)(s_1, s_2)$. In these definitions of "average treatment effect", the term corresponding to the treatment group can be interpreted heuristically as the "numerator" of a log-ratio, whereas the term corresponding to the control group can be interpreted heuristically as the "denominator" of the same log-ratio. Cf. section 7.C. This heuristic interpretation influences the choice of notation used below. Also note that it is implicit that any choice of treatment group and control group do not overlap. Thus the averages corresponding to the two terms of the ATE definition are always taken over two disjoint sets of droplets.

In defining the average treatment effects, note that (as before) we have at least 2 choices ("picky" or "gluttonous") of treatment groups and at least 2 choices ("picky" or "gluttonous") of control groups. Cf. again section 2.3.4, section 1.3, or section 5.2.1. ATE estimands using a "picky" treatment group will be denoted $\Psi^F_{p/.}(t)(s_1, s_2)$ ("p" for "**picky**"), while ATE estimands using a "gluttonous" treatment group will be denoted $\Psi^F_{g/.}(t)(s_1, s_2)$ ("g" for "**gluttonous**"). Likewise, ATE estimands using a "picky" control group will be denoted $\Psi^F_{./p}(t)(s_1, s_2)$ ("p" for "**picky**"), while ATE estimands using a "gluttonous" control group will be denoted $\Psi^F_{./g}(t)(s_1, s_2)$ ("g" for "**gluttonous**"). To summarize, within the above framework there are two ways to define the treatment group term of the ATE, and two ways to define the control group term of the ATE, leading to $2 \times 2 = 4$ total possible distinct definitions of ATE (average treatment effect) under consideration. Explicit definitions for all 4 of these possible choices of full data ATE estimand are given below.

(Full Data) Picky treatment, picky control:

$$
\begin{aligned}
\Psi^F_{p/p}(t)(s_1, s_2): \quad & \mathcal{M}^F \;\;\rightarrow\;\; \mathbb{R} \\
\Psi^F_{p/p}(t)(s_1, s_2): \quad & \mathcal{P}_F \;\;\mapsto\;\; \mathbb{E}_{\mathcal{P}_F}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \;\middle|\; \text{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\right] \\
& \qquad\quad - \mathbb{E}_{\mathcal{P}_F}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \;\middle|\; \text{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\right]
\end{aligned}
\tag{7.3.1}
$$

---

[2]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.

(Full Data) Picky treatment, gluttonous control:

$$
\begin{aligned}
\Psi_{p/g}^{F}(t)(s_1, s_2) &: \quad \mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi_{p/g}^{F}(t)(s_1, s_2) &: \quad \mathcal{P}_F \;\mapsto\; \mathbb{E}_{\mathcal{P}_F}\!\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \,\Bigg|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\right] \\
&\qquad\qquad -\mathbb{E}_{\mathcal{P}_F}\!\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \,\Bigg|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\},\, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\right]
\end{aligned}
$$

$$(7.3.2)$$

(Full Data) Gluttonous treatment, picky control:

$$
\begin{aligned}
\Psi_{g/p}^{F}(t)(s_1, s_2) &: \quad \mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi_{g/p}^{F}(t)(s_1, s_2) &: \quad \mathcal{P}_F \;\mapsto\; \mathbb{E}_{\mathcal{P}_F}\!\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \,\Bigg|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right] \\
&\qquad\qquad -\mathbb{E}_{\mathcal{P}_F}\!\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \,\Bigg|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\right]
\end{aligned}
$$

$$(7.3.3)$$

(Full Data) Gluttonous treatment, gluttonous control:

$$
\begin{aligned}
\Psi_{g/g}^{F}(t)(s_1, s_2) &: \quad \mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi_{g/g}^{F}(t)(s_1, s_2) &: \quad \mathcal{P}_F \;\mapsto\; \mathbb{E}_{\mathcal{P}_F}\!\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \,\Bigg|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right] \\
&\qquad\qquad -\mathbb{E}_{\mathcal{P}_F}\!\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \,\Bigg|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\},\, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\right]
\end{aligned}
$$

$$(7.3.4)$$

It turns out (cf. section 7.8) that these estimands are identifiable from the observed data provided we make the following assumptions:

- we always have that $\mathrm{supp}(\vec{\mathbf{N}}(0)) = \mathrm{supp}(\vec{\mathbf{N}}(t))$,

- the marginal distributions of $\vec{\mathbf{N}}(0)$, i.e. the distributions of $N^{(1)}(0)$, $N^{(2)}(0)$, ..., $N^{(S)}(0)$, are all mutually independent.

These conditions are sufficient, but they may not be necessary.

7.3.1 Stratified Full Data Average Treatment Effect

The motivation for using the following definition is explained in section 7.7.

$$
\begin{aligned}
\Psi^F(t)(s_1, s_2) : \quad & \mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi^F(t)(s_1, s_2) : \quad & \mathcal{P}_F \;\mapsto\; \mathbb{E}_{\mathcal{P}_F}\left[\mathbb{E}_{\mathcal{P}_F}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \right) \;\middle|\; \mathbb{S}^{(s_1)}(0) = 1, \vec{\mathbb{S}}^{\{s_1, s_2\}^c}(0) \right] \right. \\
& \qquad\qquad \left. - \mathbb{E}_{\mathcal{P}_F}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \right) \;\middle|\; \mathbb{S}^{(s_1)}(0) = 0, \vec{\mathbb{S}}^{\{s_1, s_2\}^c}(0) \right] \right] \\
=\;& \sum_{\vec{\mathbf{b}} \in \{0,1\}^{S-2}} \left[\mathbb{E}_{\mathcal{P}_F}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \right) \;\middle|\; \mathbb{S}^{(s_1)}(0) = 1, \vec{\mathbb{S}}^{\{s_1, s_2\}^c}(0) = \vec{\mathbf{b}} \right] \right. \\
& \left. - \mathbb{E}_{\mathcal{P}_F}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \right) \;\middle|\; \mathbb{S}^{(s_1)}(0) = 0, \vec{\mathbb{S}}^{\{s_1, s_2\}^c}(0) = \vec{\mathbf{b}} \right] \right] \cdot \mathbb{P}_{\mathcal{P}_F}\left(\vec{\mathbb{S}}^{\{s_1, s_2\}^c}(0) = \vec{\mathbf{b}} \right).
\end{aligned}
$$
$$(7.3.5)$$

Observe how this definition implicitly uses a definition of treatment and covariates (cf. section 7.7) that ignores the multiple representatives problem (cf. section 5.C). The analogous estimand that would correspond to not ignoring the multiple representatives problem appears to not (even in principle) be identifiable from the observed data. By ignoring the multiple representatives problem, we allow ourselves to use the binary strain presence/absence vector $\vec{\mathbb{S}}(0)$ (cf. section 5.2.1.1), instead of the full strain count vector $\vec{\mathbf{N}}(0)$, to define the treatment and covariates. Considering that already fairly strong assumptions (cf. section 7.8.1) are required to conclude that $\vec{\mathbb{S}}(0)$ can be determined from the observed data, the assumptions that would be required to determine (even "approximately" or "on average") $\vec{\mathbf{N}}(0)$ seem to be hopelessly strong. Considering how the strata of the analogous estimand that does not ignore the multiple representatives problem would be defined in terms of $\vec{\mathbf{N}}(0)$, this seems to be an insurmountable problem for (directly) estimating such an estimand.

One might try to sidestep this limitation by screening for "suspicious" outliers in the identified sample that might correspond to values of $N^{(s_1)}(0)$ (or $N^{(\sigma)}(0)$ for $\sigma \neq s_1, s_2$) other than 1. This would be analogous and closely related to the hypothetical strategy discussed at length in the comments at the end of section 7.8.1.2, for how to possibly weaken the "no censoring" assumption. Thus this strategy would suffer from the same problems, namely requiring a very strong "signal to noise ratio", and being very difficult to implement successfully even given a strong "signal to noise ratio". Thus it is left to future work to implement and validate such a strategy, if it is even possible.

7.4 Observed Data Model

The observed data model $\mathcal{M}^O$ corresponds to the likelihoods from equations (2.2.6), (2.2.9), and (2.2.13). The map $\phi^{\mathcal{M}} : \mathcal{M}^F \rightarrow \mathcal{M}^O$ sending distributions in the full data model to distributions in the observed data model marginalizes over all of the unobserved time points. Recall from section 1.4 that this means all cell counts at all times are censored, except for the final cell counts. The difference between the full data model and the observed data model

that is salient for the definitions of the ATE estimands is that the observed data model assumes that $\vec{\mathbf{N}}(0)/\vec{\mathbb{S}}(0)$ is unknown and that only $\vec{\mathbf{N}}(t)$ is known.

7.5 Observed Data ATE Definitions, Identifiability

What has so far been glibly ignored is the obvious issue that the droplets are observed at only one time point. In particular, the droplets are not observed upon formation at $t = 0$. There is no way to observe $N_d^{(s)}(0)$ neither for any strain s nor for any droplet d. Without further assumptions, our targeted estimands, defined in terms of the full data in section 7.3 above, cannot possibly be defined in terms of the observed data alone.

By omitting expressions not belonging to the observed data in the definitions from section 7.3 of the full data ATE estimands, we arrive at the following definitions for observed data ATE estimands. These estimands, rather than "normalizing" by the (unobservable) initial counts of cells to get a proper "growth factor", instead consider only the final counts of cells. Moreover, instead of defining the treatment and control groups in terms of the unobservable initial presence/absence of strains, these estimands define the treatment and control groups in terms of the observable presence/absence of strains at time t.

(Observed Data) Picky treatment, picky control:

$$
\begin{aligned}
\Psi_{p/p}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\;\rightarrow\;\; \mathbb{R} \\
\Psi_{p/p}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\;\mapsto\;\; \mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_1, s_2\}\Big] \\
&\qquad\qquad -\mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\Big]
\end{aligned}
\tag{7.5.1}
$$

(Observed Data) Picky treatment, gluttonous control:

$$
\begin{aligned}
\Psi_{p/g}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\;\rightarrow\;\; \mathbb{R} \\
\Psi_{p/g}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\;\mapsto\;\; \mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_1, s_2\}\Big] \\
&\qquad\qquad -\mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\Big]
\end{aligned}
\tag{7.5.2}
$$

(Observed Data) Gluttonous treatment, picky control:

$$
\begin{aligned}
\Psi_{g/p}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\;\rightarrow\;\; \mathbb{R} \\
\Psi_{g/p}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\;\mapsto\;\; \mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\Big] \\
&\qquad\qquad -\mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\Big]
\end{aligned}
\tag{7.5.3}
$$

(Observed Data) Gluttonous treatment, gluttonous control:

$$\Psi_{g/g}(t)(s_1, s_2): \quad \mathcal{M}^O \quad \to \quad \mathbb{R}$$

$$\Psi_{g/g}(t)(s_1, s_2): \quad \mathcal{P}_O \quad \mapsto \quad \mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big)\ \Big|\ \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\Big]$$

$$-\mathbb{E}_{\mathcal{P}_O}\Big[\log\big(N^{(s_2)}(t)\big)\ \Big|\ \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\Big]$$

$$(7.5.4)$$

Under certain assumptions (that are satisfied e.g. by the hPoMu and hNBDM working models), the full data ATE estimands (which assume $\vec{\mathbf{N}}(0)$ is known) would equal the observed data ATE estimands. This means the full data ATE estimands would be identifiable from the observed data. The assumptions are:

- we always have that $\mathrm{supp}(\vec{\mathbf{N}}(0)) = \mathrm{supp}(\vec{\mathbf{N}}(t))$,

- the marginal distributions of $\vec{\mathbf{N}}(0)$, i.e. the distributions of $N^{(1)}(0)$, $N^{(2)}(0)$, …, $N^{(S)}(0)$, are all mutually independent.

Cf. section 7.8 and figure 7.1 below.
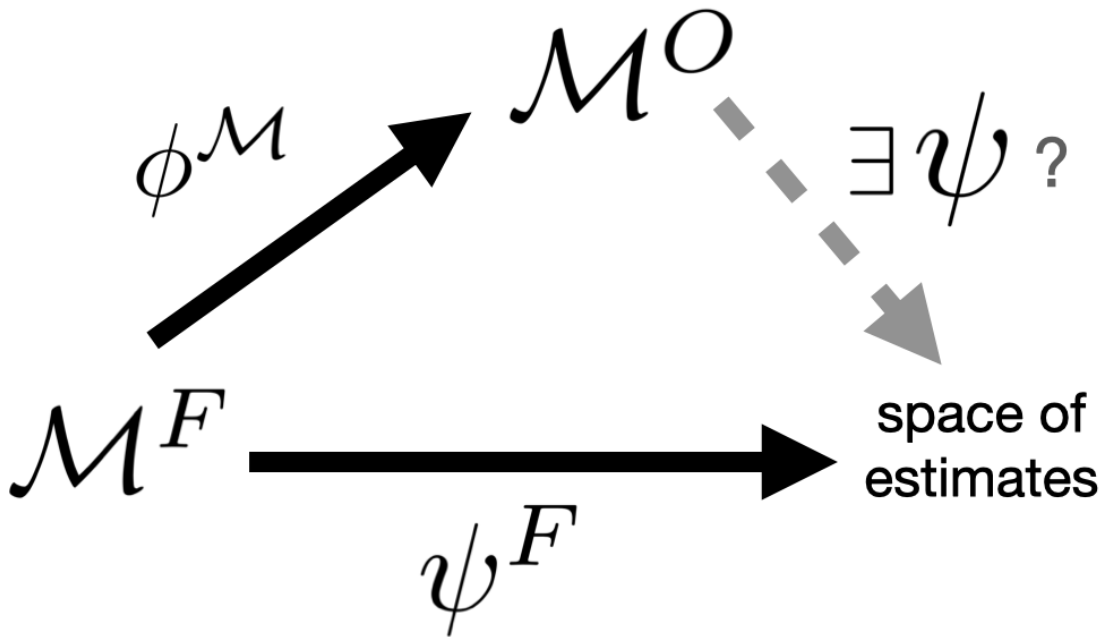
Figure 7.1: $\mathcal{M}^F$ is the full statistical model, $\mathcal{M}^O$ is the observed statistical model (i.e. corresponding to the observed data and observed likelihoods), $\phi^{\mathcal{M}}$ sends each distribution in the full statistical model to the corresponding distribution of observed data, ψ^F is the estimand mapping defined on the full statistical model answering the scientific question we are interested in. This estimand is called identifiable if there exists a corresponding estimand defined on the observed statistical model such that the above commutative diagram commutes. (For an introduction to commutative diagrams, cf. e.g. [LS09].)

Another way to interpret the observed data ATE estimands is as pretending that $N^{(s)}(0) = 1$ always (whenever it is nonzero). This may not be entirely unreasonable. The experiment can be set up so that 1 is the highest probability nonzero count of cells belonging to any given strain at a droplet's formation. (Cf. the discussion in section 5.C.) Hence, even when we do not believe the assumptions made in section 7.8 to be valid, the experiment can still be calibrated to minimize the "identification gap" between the full data ATE estimands and the observed data ATE estimands, which makes the estimands useful.

7.5.1 Stratified Observed Data Average Treatment Effect

The motivation for using the following definition is explained in section 7.7.

$$
\begin{aligned}
\Psi(t)(s_1, s_2): \quad & \mathcal{M}^O \;\rightarrow\; \mathbb{R} \\
\Psi(t)(s_1, s_2): \quad & \mathcal{P}_O \;\mapsto\; \mathbb{E}_{\mathcal{P}_O}\left[\mathbb{E}_{\mathcal{P}_O}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathbb{S}^{(s_1)}(t)=1, \vec{\mathbb{S}}^{\{s_1,s_2\}^c}(t)\right]\right. \\
& \hspace{3.5cm} \left. -\, \mathbb{E}_{\mathcal{P}_O}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathbb{S}^{(s_1)}(t)=0, \vec{\mathbb{S}}^{\{s_1,s_2\}^c}(t)\right]\right] \\
& \hspace{1.2cm} =\; \sum_{\vec{\mathbf{b}}\in\{0,1\}^{S-2}}\left[\mathbb{E}_{\mathcal{P}_O}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathbb{S}^{(s_1)}(t)=1, \vec{\mathbb{S}}^{\{s_1,s_2\}^c}(t)=\vec{\mathbf{b}}\right]\right. \\
& \hspace{1.0cm} \left. -\mathbb{E}_{\mathcal{P}_O}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathbb{S}^{(s_1)}(t)=0, \vec{\mathbb{S}}^{\{s_1,s_2\}^c}(t)=\vec{\mathbf{b}}\right]\right]\cdot\mathbb{P}_{\mathcal{P}_F}\left(\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(t)=\vec{\mathbf{b}}\right).
\end{aligned}
$$
(7.5.5)

Like definition (7.3.5), this definition also implicitly uses a definition of treatment and covariates (cf. section 7.7) that ignores the multiple representatives problem (cf. section 5.C). For a discussion of many of the issues that would be entailed by trying to use an estimand analogous to (7.5.5) that addresses the multiple representatives problem, cf. section 7.3.1.

7.6 Estimation

For a given observed data ATE estimand Ψ, let its corresponding (plugin) estimator be denoted $\hat{\boldsymbol{\Psi}}$. The latter are defined explicitly below.

(Estimator) Picky treatment, picky control:

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{p/p}(t)(s_1, s_2) :=& \hat{\boldsymbol{\mu}}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(t))=\{s_1, s_2\}\right] \\
& -\,\hat{\boldsymbol{\mu}}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(t))=\{s_2\}\right]
\end{aligned}
$$
(7.6.1)

(Estimator) Picky treatment, gluttonous control:

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{p/g}(t)(s_1, s_2) :=& \hat{\boldsymbol{\mu}}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(t))=\{s_1, s_2\}\right] \\
& -\,\hat{\boldsymbol{\mu}}\left[\log\left(N^{(s_2)}(t)\right)\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(t))\supseteq\{s_2\}, s_1\notin\mathrm{supp}(\vec{\mathbf{N}}(t))\right]
\end{aligned}
$$
(7.6.2)

(Estimator) Gluttonous treatment, picky control:

$$\hat{\boldsymbol{\Psi}}_{g/p}(t)(s_1, s_2) := \hat{\boldsymbol{\mu}}\Big[\log\big(N^{(s_2)}(t)\big) \mid \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\Big]$$
$$- \hat{\boldsymbol{\mu}}\Big[\log\big(N^{(s_2)}(t)\big) \mid \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\Big]$$

(7.6.3)

(Estimator) Gluttonous treatment, gluttonous control:

$$\hat{\boldsymbol{\Psi}}_{g/g}(t)(s_1, s_2) := \hat{\boldsymbol{\mu}}\Big[\log\big(N^{(s_2)}(t)\big) \mid \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\Big]$$
$$- \hat{\boldsymbol{\mu}}\Big[\log\big(N^{(s_2)}(t)\big) \mid \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\Big]$$

(7.6.4)

7.6.1 Stratified Estimator Definition and TMLE

The motivation for using the following definition is explained in section 7.7.

$$\hat{\boldsymbol{\Psi}}(t)(s_1, s_2) := \sum_{\vec{\mathbf{b}} \in \{0,1\}^{S-2}} \Big[\hat{\boldsymbol{\mu}}\Big[\log\big(N^{(s_2)}(t)\big) \mid \mathbb{S}^{(s_1)}(t) = 1, \vec{\mathbb{S}}^{\{s_1, s_2\}^c}(t) = \vec{\mathbf{b}}\Big]$$
$$- \hat{\boldsymbol{\mu}}\Big[\log\big(N^{(s_2)}(t)\big) \mid \mathbb{S}^{(s_1)}(t) = 0, \vec{\mathbb{S}}^{\{s_1, s_2\}^c}(t) = \vec{\mathbf{b}}\Big]\Big] \cdot \frac{\displaystyle\sum_{d \in [D]} \mathbb{I}\Big\{\vec{\mathbb{S}}_d^{\{s_1, s_2\}^c}(t) = \vec{\mathbf{b}}\Big\}}{D}.$$

(7.6.5)

Note that we can do much better than this naïve plugin estimator when it comes to estimating a stratified ATE estimand of the form (7.3.5) or (7.5.5). Even when it is not possible in finite samples to stratify over all possible values of the covariates (cf. sections 7.7.1 and 7.7.3), for our initial estimate it is still possible to use SuperLearner (see [vdLPH07] and [vdLR11, Chapter 3]) to implement an ensemble machine learning approach that does not require us to make additional assumptions that we do not actually believe. For example, we could use different settings for GLMnet [FHT10] regression on the binary treatment and covariates as some of the learners in the ensemble considered by SuperLearner.

Afterwards, we can use the targeting step as implemented in e.g. the R package 'tmle' [GL12] for targeted minimum-loss based estimation (TMLE). This adjusts for the high-dimensional covariates in a way that makes the initial estimate produced by SuperLearner unbiased. Targeted learning leads to the most efficient and effective use of the available data possible, because our efforts are focused only on estimating the portion $\mathcal{Q}_O$ of the observed probability distribution $\mathcal{P}_O$ that is relevant for determining our chosen target parameter (7.5.5). The targeted learning story has already been developed in detail for this kind of estimand, cf. [vdLR11] and [vdLR18]. The first five chapters of [vdLR11] are particularly relevant. In this preliminary exposition I do not go into the full details of implementing the targeted learning roadmap for this class of problem. However that is the goal of future work.

7.7 Confounder Adjustment and Positivity

Although not mentioned explicitly in sections 2.3.4 and 5.2.1, the gluttonous and picky groupings are two strategies for adjusting for possible confounders. These two strategies lie at opposite ends of a tradeoff. The tradeoff is between, on the one hand, using all droplets that might possibly be informative for inferring the effect of one strain on the other, and, on the other hand, completely controlling for the presence of third strains. To not give up on being able to use all possibly informative droplets, the gluttonous groups give up on being able to completely control for the presence of third strains. In order to not give up on being able to completely control for the presence of third strains, the picky groups give up on being able to use all possibly informative droplets.

As discussed in section 1.3, one reason for choosing to analyze data from MOREI, rather than from e.g. an observational metagenome study, is that the encapsulation in droplets should already control for confounding due to other strains $\sigma \neq s_1, s_2$ as much as possible. In particular, one of the goals is so that the data is as comparable as possible to the "gold standard" of manually plating cell cultures (cf. e.g. [VCF+18]), for which[3] $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0) = \vec{\mathbf{0}}_{S-2}$ always. Our goal is that, all other things being equal[4], in general $\vec{\mathbf{0}}_{S-2}$ will be the value of $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0)$ that occurs with the highest probability. As long as we have enough data to make a reliable, estimate using the picky groups only, that is what we always prefer to do. Whenever we use any droplets for which $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0) \neq \vec{\mathbf{0}}_{S-2}$ while making estimates, it is (almost) *always* due to necessity induced by "data starvation" (cf. again the discussion in section 2.3.4). This last point is important to keep in mind for later.

7.7.1 Outcomes, Treatments, Covariates

The problem of confounder adjustment or control can be understood within the framework of splitting the observed data ("O") between (i) outcomes/effects ("Y"), (ii) treatments/exposures ("A"), and (iii) covariates/potential confounders ("W") as is done e.g. in the references [vdLR11], [vdLR18], and [PPG+10]. Recall that throughout this chapter we are using ATEs to study the effect of strain s_1 on the growth of strain s_2.

The definitions in section 7.5 correctly suggest that the outcome Y we are interested in is $\log\left(N^{(s_2)}(t)\right)$. When we assume the full data is available and known, then the definitions in section 7.3 correctly suggest that the outcome Y we are interested in becomes $\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right)$.

The treatment A we are interested in is the presence/absence $\mathbb{S}^{(s_1)}(0)$ of strain s_1. Observe how, by using the binary $\mathbb{S}^{(s_1)}(0)$ as the treatment instead of $N^{(s_1)}(0)$, we are effectively choosing to ignore the multiple representatives problem (cf. section 5.C). It is clear that the value of $\mathbb{S}^{(s_1)}(0)$ is the treatment variable, because (given a fixed choice between "gluttonous" or "picky") a value of $\mathbb{S}^{(s_1)}(0) = 1$, which is equivalent to the condition $\{s_1\} \in \text{supp}(\vec{\mathbf{N}}(0))$,

---

[3] Here $\vec{\mathbf{0}}_{S-2}$ is used to denote the all-0's vector of length $(S-2)$.

[4] E.g. the experiment is set up reasonably so that the expected number of cells per droplet λ is not too large, in particular not a number much larger than 2. Cf. again section 1.4.1.

is always what distinguishes the treatment group from the control group, the latter corresponding to $\mathbb{S}^{(s_1)}(0) = 0$ or equivalently $\{s_1\} \notin \text{supp}(\vec{\mathbf{N}}(0))$.

Finally this leaves us to address the covariates, or potential confounders, W. These correspond of course to strains $\sigma \neq s_1, s_2$ which are not directly relevant to the effect of strain s_1 on the growth of s_2 (although of course they might potentially mediate that effect, cf. e.g. section 1.2, [MXS17], or [BKK16]). With the length S binary random vector $\vec{\mathbb{S}}(0)$ representing the presence/absence of all S strains at time 0, let $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0)$ denote the length $(S-2)$ binary random vector containing the $(S-2)$ entries of $\vec{\mathbb{S}}(0)$ corresponding to the $S-2$ strains $\sigma \neq s_1, s_2$, i.e. all indices in $\{s_1, s_2\}^c := [S] \setminus \{s_1, s_2\}$. Reviewing the definitions of (7.3.5), (7.5.5), and (7.6.5), and because the definitions of the gluttonous groups differ from those of their picky counterparts only[5] in the presence/absence of strains $\sigma \neq s_1, s_2$, it is hopefully clear that the relevant choice of covariates, and thus potential confounders, is $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0)$. Observe how considering entries of the vector $\vec{\mathbb{S}}(0)$, rather than entries of the vector $\vec{\mathbf{N}}(0)$, in defining our covariates W again implies that we are choosing to ignore the multiple representatives problem (cf. section 5.C). Notice also that the picky definitions amount to requiring that we only consider droplets for which the value of *all* entries of $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0)$ is 0, whereas the gluttonous definitions allow us to consider droplets for which $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0)$ has *any* of its 2^{S-2} possible values in $\{0,1\}^{S-2}$.

7.7.2 Stratify to Attempt to Escape Tradeoff

To avoid the tradeoff for which the picky groups and gluttonous groups are positioned at opposite ends, we might like to be able to use all possibly informative droplets while still controlling for the presence of strains $\sigma \neq s_1, s_2$ *in some way*, even if not completely controlling for them by excluding all droplets where they are present. We can achieve such a compromise by stratifying over the effects that occur for all possible values $\vec{\mathbf{b}} \in \{0,1\}^{S-2}$ of $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0)$ and weighting the effects according to the probabilities for $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0) = \vec{\mathbf{b}}$. Cf. the definitions (7.3.5), (7.5.5), and (7.6.5). Such an approach in fact corresponds to the standard definition of average treatment effect (cf. chapters 1 and 2 of [vdLR11]).

Observe how the conditional expectation estimands for strata $\vec{\mathbf{b}}$ implying the presence of 3 or more strains in the definitions (7.3.5) and (7.5.5) amount to quantifications of the effects of "higher-order interactions", cf. section 1.2. Thus the stratifications in the definitions (7.3.5), (7.5.5), and (7.6.5) can legitimately be interpreted as "stratifying over all relevant higher-order interactions". Hence being able to use (7.6.5) to effectively estimate (7.3.5) or (7.5.5) requires being able to effectively estimate all higher-order interactions, which as discussed before in sections 2.3.5 and 1.2 is where we begin to encounter serious problems.

In particular, the usefulness of the stratified ATE requires that the experimental treatment assignment (ETA) assumption, also known as the positivity assumption, be satisfied [vdLR11, p. 35]. Quoting from [PPG+10]:

---

[5] I.e. irrespective of the particular values of the nonzero entries of $\vec{\mathbf{N}}(0)$.

Positivity violations can arise for two reasons. First, it may be theoretically impossible for individuals with certain covariate values to receive a given exposure of interest... Second, violations or near violations of positivity can arise in finite samples due to chance.

The first kind of violation, the theoretical impossibility, corresponds to some of the conditional expectation estimands that are summands in the expression (7.3.5) not being well-defined. The second kind of violation or near violation, occurring for finite samples, can correspond to some of the plugin estimators for conditional expectations that are summands in the expression (7.6.5) either not being well-defined in the finite sample[6] or being unreliable/having high variance in the finite sample[7]. Although the first kind of violation is not relevant here, I argue in the next section that in practice it is impossible for us to avoid the second kind of violation or near violation of the positivity assumption.

7.7.3 Why Positivity Violations are Likely Unavoidable in Practice, if not in Principle

To emphasize, the following bounds are not meant to be precise consequences proven from certain fixed assumptions, they are intended to be entirely heuristic. The basic ideas are that (1) the probability of a given strain appearing in a droplet should be roughly proportional to its relative abundance in the sampling population, and that (2) the probabilities of multiple strains appearing in the same droplet should be ("to 0th or 1st order") roughly multiplicative.

The number of treatment samples for the effect of strain s_1 on the growth of strain s_2 can be thought of as roughly[8] proportional to $\frac{1}{S^2}$ or[9] $O(\frac{1}{S^2})$ in probability, whereas the number of control samples for the effect of strain s_1 on the growth of s_2 can be thought of as roughly proportional to $\frac{1}{S}$ or $O(\frac{1}{S})$ in probability. Cf. section 2.C for essentially the same idea.

In particular, the number of samples that contain at least one of each of I given strains s_1, s_2, ..., s_I would be roughly proportional to $\frac{1}{S^I}$ or $O(\frac{1}{S^I})$ in probability. Thus, even as the number of strata $\vec{\mathbf{b}} \in \{0,1\}^{S-2}$ in the definitions (7.3.5), (7.5.5), and (7.6.5) ***increases*** exponentially with increasing S, at the same time the number of samples available for most strata $\vec{\mathbf{b}}$ roughly ***decreases*** exponentially (in probability) with increasing S. Thus, even in the absence of any theoretical positivity violations, i.e. even if $\mathbb{P}(\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0) = \vec{\mathbf{b}}) > 0$ for all strata $\vec{\mathbf{b}}$, in practice we are almost always[10] confronted with massive finite sample positivity

---

[6] With 0 observations in the finite sample of at least one of the treatment or control group for that stratum, the corresponding sample conditional mean(s) would not be well-defined.

[7] Due to the number of observations in the finite sample of at least one of the treatment or control group for that stratum being very small, even if the number of observations for both is technically nonzero. The corresponding sample conditional mean(s) would be unreliable.

[8]E.g. assuming all strains have approximately the same relative abundance $\frac{1}{S}$.

[9] Strictly speaking this should be considered a blatant misuse of Landau Big-Oh notation.

[10] It should probably go without saying, but this is because it is not physically possible to indefinitely increase the total number of D exponentially with respect to increasing S.

violations. This is especially the case in the regime where S is large, which is the regime of greatest scientific novelty and thus interest. Cf. the similar problem motivating [TRKS$^+$17].

Cf. also section 2.3.5, which discusses a closely related issue. Note again that conditional expectation estimands for strata $\vec{\mathbf{b}}$ implying the presence of 3 or more strains effectively correspond to quantifying the effects of "higher-order interactions". While not explicitly targeting these higher-order interactions may possibly limit the usefulness of ecological interpretations of any results (cf. e.g. [BKK16] or [MXS17] for arguments to this effect), as discussed before already in section 1.2 such a point is irrelevant to the extent that we simply lack the data required to reliably estimate these effects. (All of this also ignores the inevitable computational/algorithmic difficulties associated with attempting to estimate exponentially many parameters, as was mentioned before in section 2.3.5.)

7.7.4   Chosen Approaches to Positivity Violations

Section 6.5 of [PPG$^+$10] describes any strategy for approaching positivity violations as a "tradeoff" between "altering the target parameter" and "improved identifiability". Earlier I described both the picky and gluttonous approaches as opposite ends of another tradeoff. I explain below how the tradeoff mentioned earlier can be considered a special case of the tradeoff described in [PPG$^+$10].

In particular, the picky approach is a special case of the "trimming" strategy described in section 6.3 of [PPG$^+$10]. By restricting the sample to only include the stratum for which $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0) = \vec{\mathbf{0}}_{S-2}$, we improve identifiability by removing the possibility of confounding due to the presence of strains $\sigma \neq s_1, s_2$.

The gluttonous approach does not seem to correspond directly to any of the strategies mentioned explicitly in [PPG$^+$10] or chapter 8 of [vdLR11], although it may be analogous to the "projection" approach mentioned therein for marginal structural models. The idea of the gluttonous approach is to alter the target parameter by "collapsing all of the strata into a single stratum". Because there is now only a single stratum, there is no longer any need to restrict samples according to their stratum membership, which in some sense improves the finite-sample identifiability from observed data. This is, of course, at the cost of substantially altering the research question being asked.

This leads to the question of how one chooses between the two approaches in practice. As implied above, in the introduction to 7.7, and sections 2.3.4 and 5.2.1, the main criterion is "data starvation". While it is true in general that the stratum corresponding to $\vec{\mathbb{S}}^{\{s_1,s_2\}^c}(0) = \vec{\mathbf{0}}_{S-2}$ is usually the stratum with the highest probability of occurrence, especially for smaller finite samples this is often still not enough to guarantee that both sample conditional means in definition (7.6.1) are even well-defined[11]. Thus the "trimming" strategy only works when we have sufficient data from our largest stratum. When we lack sufficient data from our largest stratum, then it is perhaps unsurprising that our only remaining choice is to collapse all strata, i.e. adopt the gluttonous approach. Even then that sometimes is still not enough.

---

[11]Much less calculated from a large enough sample to have manageable variance.

Recall from earlier (e.g. sections 2.C and 7.7.3) how in general we are much more likely to suffer "data starvation" for the treatment groups than for the control groups. With that in mind, preliminary results (not shown) from simulated data appear to support the "data starvation" criterion for choosing between the picky and gluttonous approaches. In particular, for combinations involving strains with smaller relative abundances ($\approx 2\%$ or less) and thus having smaller finite-sample sizes, the "gluttonous treatment and gluttonous control" approach, cf. equation (7.6.4), appeared to be the most effective. Due to the small sample sizes, "data starvation" was evidently occurring for both the treatment and control groups. However, for combinations involving strains with larger relative abundances ($\approx 5-10\%$) the "gluttonous treatment and picky control" approach appeared to be the most effective, cf. equation (7.6.3). Due to the slightly larger sample sizes, "data starvation" evidently was not a factor for the control groups, allowing the picky approach's stricter controlling for confounding to reap benefits. However, because treatment group sizes are much smaller than control group sizes, even given the larger sample sizes the treatment group sizes were still sufficiently small that the correction for "data starvation" implemented by the gluttonous approach remained useful. These preliminary results were not only for average treatment effect estimators, but also for other kinds of estimands (e.g. conditional correlations or rank-based coefficients). That the same pattern occurred for all of these different kinds of estimands strongly suggests that their shared tradeoff, between available sample size and strictness of controlling for confounders, most plausibly explains the pattern.

One might object to the collapsing of all strata implicit in the gluttonous approach as being too odious. However, in choosing to ignore the multiple representatives problem (cf. section 5.C) by using binary variables $\vec{\mathbb{S}}(0)$ as our exposure and covariates, rather than the count variables $\vec{\mathbf{N}}(0)$, we were also effectively already choosing to collapse strata. Cf. again section 7.7.1. For example, a stratum corresponding to $\mathbb{S}^{(s_1)}(0) = 1$ can be considered collapsing strata corresponding to $N^{(s_1)}(0) = 1$, $N^{(s_1)}(0) = 2$, $N^{(s_1)}(0) = 3$, ... Therefore I argue that the gluttonous approach should not be considered any more odious than our previous choice to ignore the multiple representatives problem.

It is true that there is at least one important difference. Collapsing the strata corresponding to $\vec{\mathbb{S}}(0)$ is necessitated by a lack of identifiability that in principle may hold for arbitrary, even infinite, sample sizes. In contrast, the further collapse of strata employed by the gluttonous approach is only due to finite-sample constraints. However, as argued before in section 7.7.3, given that the sample sizes necessary to overcome those finite-sample constraints would be exponential, and thus usually not even physically possible, I argue that in practice this difference amounts to an unimportant distinction. Thus even if the gluttonous approach is odious, I argue that in practical terms it is no less of a necessary evil than choosing to ignore the multiple representatives problem.

7.8 Sufficient Conditions for Identifiability

Section 7.8.1 in effect addresses the question: why (given that we may be able to ignore the multiple representatives problem, cf. section 5.C) is the presence/absence of strains at the final time t sufficient to completely characterize the presence/absence of strains initially at time 0? Section 7.8.2 in effect addresses the question: why can we ignore the multiple representatives problem?

Comment: Recall from section 5.2.1.1 how $\text{supp}(\vec{\mathbf{N}}(0)) = \text{supp}(\vec{\mathbb{S}}(0))$. Yet at the same time $\text{supp}(\vec{\mathbf{N}}(0)) = \text{supp}(\vec{\mathbf{N}}(t))$ if and only if $\vec{\mathbb{S}}(0) = \vec{\mathbb{S}}(t)$. Hence claims about the identifiability of $\text{supp}(\vec{\mathbf{N}}(0))$ can equally be interpreted as claims about the identifiability of $\vec{\mathbb{S}}(0)$. In particular, even in the presence of an identification gap between the full data estimand $\Psi^F(t)(s_1, s_2)$ and the observed data estimand $\Psi(t)(s_1, s_2)$, because the assumptions for identifiability of $\text{supp}(\vec{\mathbf{N}}(0))$ from the observed data are (arguably) much milder, it is still appealing to define treatment and control groups in terms of $\vec{\mathbb{S}}(0)$ instead of $\vec{\mathbf{N}}(0)$ and thereby neglect the multiple representatives problem (cf. section 5.C).

7.8.1 Determination of $\text{supp}(\vec{\mathbf{N}}(0))$ from Observed Data

The assumption that

$$\text{supp}(\vec{\mathbf{N}}(0)) = \text{supp}(\vec{\mathbf{N}}(t)), \tag{7.8.1}$$

can be broken down into two assumptions.

7.8.1.1 No Spontaneous Generation

This assumption says that the number of strains within any given droplet can not increase over time, throughout the course of the experiment:

$$\text{supp}(\vec{\mathbf{N}}(0)) \supseteq \text{supp}(\vec{\mathbf{N}}(t)). \tag{7.8.2}$$

In other words, strains that were not present during the droplet's formation can not later spontaneously appear within the droplet.

This is certainly true if we consider only the incubation of the droplets (cf. section 1.4.2). As Louis Pasteur already demonstrated in the 19th century using a swan neck flask, the creation of cells of a given strain of microbe requires the prior presence of cells of that strain. The theory of "spontaneous generation" is false, so microbes of a given strain can not spontaneously arise from non-living matter, preventing the number of strains from spontaneously increasing. Thus prima facie this assumption appears to be entirely unquestionable.

However, the name of the assumption is somewhat misleading. In practice, the implausible occurrence of spontaneous generation may not actually be necessary for this assumption to be violated. Indeed, experimental error that causes some small percentage of the droplets to be merged incorrectly, or to otherwise incorrectly receive the same barcode (cf. sections

1.4.3 and A.2.2) would cause a similar effect to this assumption being violated. Preliminary results suggested that this effect may be negligible for inference. The simulation found little to no deleterious effect on inference caused by the rate of merging error assumed by the simulation. My collaborator also said that MOREI protocol has subsequently been improved to drastically further reduce the rate of merging error. Hence, although this assumption may not be *entirely* unquestionable in practice, it still appears to be a safe assumption to make.

7.8.1.2 No Censoring

This corresponds to the other direction, namely the assumption that the number of strains cannot decrease over time, throughout the course of the experiment:

$$\operatorname{supp}(\vec{\mathbf{N}}(0)) \subseteq \operatorname{supp}(\vec{\mathbf{N}}(t)). \tag{7.8.3}$$

In other words, once a given strain is present in a droplet, it can never disappear. This assumption is actually much less realistic in practice than that from section 7.8.1.1. In particular, in the case of antagonistic interactions, where one strain harms another, we would expect the antagonized strain to quite often go extinct.

Thus the failure of this identifiability assumption poses particular danger to the inference of antagonistic microbial interactions (negative edge weights in the signed network representation, cf. section 1.2). However, as discussed in sections 1.4.3, A.2.2, and A.2.3, given that the data we observe does not actually correspond exactly to the counts of living cells $\vec{\mathbf{N}}(t)$ at the time of sequencing t, there are reasons to believe that such censoring concerns could be mitigated by phenomena that occur in practice[12], e.g. "relic" genetic material[13]. In other words, we might expect the sequencing results to be more roughly proportional to the total number of cells that have lived *ever*, rather than necessarily be roughly proportional to the current number of living cells. Cf. section A.2.2 and figure A.2.3 in particular.

Whether censoring occurs depends not only on the living or dying of cells during the second (incubation) phase of the experiment as described in section 1.4, but also on possible effects of PCR amplification that occurs during the third (sequencing) phase. This presents several possible complications. One is the possible "swamping out" in the sequencing results of the signal from strains with small relative abundances. Cf. footnote 13 from section 1.4.3. Regardless of whether the genetic material is "relic" or from living cells, if the proportion corresponding tor a given strain is small, then its sequence might not show up in the sequencing results after PCR amplification. Cf. the discussion of "detection limits" as a possible cause of "zero inflation" in [MGSD19, section 1.1]. Thus a strain might not be

---

[12] Preliminary results suggest that, while avoiding censoring this way could increase our ability to infer antagonistic interactions, it could also decrease our ability to infer the magnitude of beneficial interactions (which correspond to positive edges in the signed network representation, cf. again section 1.2). This is because overcounting the number of living benefit-inducing cells required to induce a certain growth advantage (by including "relic" genetic material in the count) causes us to underestimate the strength of the beneficial growth effect. Therefore "relic" genetic material is potentially a double-edged sword.

[13] Cf. section 1.4.3 and appendix A.2.2 for more discussion of "relic" genetic material.

registered as having been in the droplet even though the droplet contained genetic material from living and/or dead cells of the strain. (Such effects could be called "technical zeros" using the terminology of [PGB21].) This is related to "jackpot effects", cf. e.g. [MSH⁺19].

In particular, the number of reads after PCR amplification is not necessarily a linear function of the initial number of reads, not even "on average". In the case of such "non-linear PCR amplification", the relative proportions of reads for each of the strains before PCR amplification may not still be the same after PCR amplification. This is in contrast to the predictions of the simple, naïve model mentioned in appendix A.2.3, for which the number of amplified reads is a constant linear factor times some multiplicative noise. Besides potential non-linearity or multiplicative noise of PCR amplification, a further factor that might exacerbate "technical zeros" resulting from "swamping out" are sequence-specific PCR amplification biases. See [MWC19] or appendix A.2.2 for a discussion.

Ultimately though, as stated before at the beginning of chapter 2:

> The problem of choosing good statistical models of PCR amplification is applicable to an extremely wide range of biological experiments. It should be its own project and fully addressing it is outside of the scope of this work. Previous work that has more substantially considered similar problems includes [YRCS06] [KMP⁺07] [YWSJ08] [DH15][MWC19][MSH⁺19].

Future research directions could include (1) gauging the extent to which experimental artifacts actually do blunt the effects of censoring in practice, (2) developing methods to adjust the analysis to account for the possibility of censoring, (3) gauging how often censoring might occur in practice, or (4) gauging to what extent failure of the no censoring assumption might bias the results (i.e. the size of the identification gap) in practice.

Comments: This assumption also appears to effectively be necessary to sidestep the issue of whether it makes sense to use the convention that "$\log(0) = 0$". Namely because any droplets for which $N^{(s_2)}(t) = 0$ are always de facto excluded from the sample of droplets for which we assume that $N^{(s_2)}(0) \geq 1$. This is because the no censoring assumption implies that such droplets do not exist, and because de facto we have no way to reliably identify them anyway, given that our only observations of the droplets are at time t, and not time 0. Any future work which claimed to be able to identify such droplets would then have to put more thought into what convention should be used for "$\log(0)$". Such "zero inflation" problems however are not inherently unique to the analysis of MOREI-like data, cf. for example [MGSD19] or [PGB21]. (However the problems posed by zeros are in many ways different for the compositional datasets studied in [MGSD19] or [PGB21], whereas the data from MOREI is inherently not compositional, cf. again the discussion of spikein genes from section 1.4.3. Cf. e.g. [CHL⁺20] or [SWMD17] for subtleties of using compositional data.)

One might try to circumvent the limitations of this assumption as follows, at least for censoring caused by antagonistic interactions. Namely, given an appropriate "neutral model" predicting which percentage of droplets should contain certain combinations of strains based

on the strains' relative abundances, we could look at whether any combinations have "suspiciously" small numbers of observations. From that we might conclude that in fact some of the droplets containing that combination at time 0 were not counted due to one of the relevant strains becoming extinguished by time t. For example, if strain s_1 harms strain s_2, and we see "suspiciously" few droplets containing both s_1 and s_2, we might conclude that several of the droplets seeming to only contain strain s_1 most likely originally also contained strain s_2 but that subsequently strain s_2 became extinct in those droplets by time t.

Even ignoring the obvious (and potentially insurmountable) difficulty of identifying which droplets seeming to contain only s_1 actually originally also contained strain s_2 at time 0, there are at least two issues with such a proposal.

First, the expected number of droplets containing a given combination of strains might be very small and thus subject to large variance (or at least large "noise to signal ratio", cf. the discussion of a similar issue in footnote 13 from section 1.4.3). For example, even in a toy neutral model where the observed count of any given combination of strains varies usually by no more than ± 2 from its expected count, if the expected count is 5 and the observed count is 3 or 2, there is very little "signal" available with which to distinguish the two possibilities of random variation and censoring due to antagonistic interaction. Cf. the related discussion about practical violations of the positivity assumption in section 7.7. This is in contrast to e.g. a situation where the expected count is 50 and the observed count is 20 (instead of e.g. 48 or 47).

Second, even if we somehow manage to create a large enough total number of droplets D such that all of the combinations of strains have large enough expected counts to have amenable "signal to noise ratios", we are still left with the problem of identifying an appropriate neutral model to compare against in order to flag "suspicious" observed counts. Given previous work using the Poisson distribution as the default neutral model for total cell counts in droplets for single-cell sequencing experiments (e.g. [CNd+15]), it is tempting to accept the hPoMu working model as "the unquestionably appropriate neutral working model". However, even retaining the assumption that the total number of cells in each droplet is Poisson distributed (e.g. as is also the case for the hPoDM working model), one of the major conclusions of chapter 5 is that in a neutral model the observed counts can still be "statistically significantly" lower than the expected counts predicted by hPoMu. Preliminary work (not shown) also suggested that trying to identify "suspicious" low observed counts using hPoMu as a neutral model is most likely doomed to failure, with almost all examples identified as potential instances of censoring due to antagonism being false positives caused by the bias introduced by the misspecified choice of neutral model.

Thus choosing an appropriate neutral model requires either fitting to some relevant data, or establishing a substantial body of research that identifies what is in most scenarios an appropriate neutral model (assuming that such an appropriate "universal" choice even exists, which it may not). One reason for recommending the "$t = 0$ pre-experiment" described in section 2.3 is because, by not incubating the droplets, it allows us to control for the possibility that "suspicious" observed counts are due to antagonistic interactions, rather than an incorrectly chosen neutral model. Therefore the "$t = 0$ pre-experiment" is ideally

suited for producing data with which to fit an optimal neutral model.

7.8.2 Identifiability of $\Psi^F(t)(s_1, s_2)$ from Observed Data

With respect to the notation used in the following lemmas below, one has that the picky treatment ATE estimands correspond to the choice:

$$\mathscr{O}_\mathcal{T} := \{\Omega \subseteq [S] : \Omega = \{s_1, s_2\}\} \,, \tag{7.8.4}$$

while the gluttonous treatment ATE estimands correspond to the choice:

$$\mathscr{O}_\mathcal{T} := \{\Omega \subseteq [S] : \Omega \supseteq \{s_1, s_2\}\} \,. \tag{7.8.5}$$

Similarly, the picky control ATE estimands correspond to the choice:

$$\mathscr{O}_\mathcal{C} := \{\Omega \subseteq [S] : \Omega = \{s_2\}\} \,, \tag{7.8.6}$$

and the gluttonous control ATE estimands correspond to the choice:

$$\mathscr{O}_\mathcal{C} := \{\Omega \subseteq [S] : \Omega \supseteq \{s_2\}, s_1 \notin \Omega\} \,. \tag{7.8.7}$$

Lemma 7.1. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

- *that the distributions of the droplets are independent,*

- *we always have that* $\mathrm{supp}(\vec{\mathbf{N}}(t)) = \mathrm{supp}(\vec{\mathbf{N}}(0))$ *(cf. section 7.8.1),*

then for any two subsets $\mathscr{O}_\mathcal{T}, \mathscr{O}_\mathcal{C} \subseteq \mathcal{P}([S])$ of the power set of $[S]$ such that $s_2 \in \Omega_\mathcal{T}$ for all $\Omega_\mathcal{T} \in \mathscr{O}_\mathcal{T}$ and $s_2 \in \Omega_\mathcal{C}$ for all $\Omega_\mathcal{C} \in \mathscr{O}_\mathcal{C}$, then the estimator

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}(t)(s_1, s_2) :=& \hat{\boldsymbol{\mu}} \left[\log \left(N^{(s_2)}(t) \right) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_\mathcal{T} \right] \\
& - \hat{\boldsymbol{\mu}} \left[\log \left(N^{(s_2)}(t) \right) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_\mathcal{C} \right]
\end{aligned}
\tag{7.8.8}
$$

is statistically consistent for the full data estimand assuming that $\vec{\mathbf{N}}(0)$ is known

$$
\begin{aligned}
\Psi^F(t)(s_1, s_2) :=& \mathbb{E} \left[\log \left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \right) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T} \right] \\
& - \mathbb{E} \left[\log \left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \right) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C} \right].
\end{aligned}
\tag{7.8.9}
$$

Proof of Lemma 7.1:

Proof of identifiability of $\Psi^F(t)(s_1, s_2)$ from the observed data:
Using the properties of logarithms and the linearity of expectation one has that the expression
(7.8.9) defining $\Psi^F(t)(s_1, s_2)$ above equals

$$
\mathbb{E}\left[\log(N^{(s_2)}(t)) \mid \text{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right] - \mathbb{E}\left[\log(N^{(s_2)}(t)) \mid \text{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{C}}\right]
$$
$$
- \left(\mathbb{E}\left[\log(N^{(s_2)}(0)) \mid \text{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right] - \mathbb{E}\left[\log(N^{(s_2)}(0)) \mid \text{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{C}}\right]\right) . \tag{7.8.10}
$$

By Lemma 7.2 using $\phi(n) = \log(n)$, the expression inside of parentheses in (7.8.10) vanishes.
Thus in this case $\Psi^F(t)(s_1, s_2) = \Psi(t)(s_1, s_2)$:

$$
\begin{aligned}
\Psi^F(t)(s_1, s_2) &= \mathbb{E}\left[\log(N^{(s_2)}(t)) \mid \text{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right] - \mathbb{E}\left[\log(N^{(s_2)}(t)) \mid \text{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{C}}\right] \\
&= \mathbb{E}\left[\log(N^{(s_2)}(t)) \mid \text{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{T}}\right] - \mathbb{E}\left[\log(N^{(s_2)}(t)) \mid \text{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{C}}\right] \\
&= \Psi(t)(s_1, s_2) ,
\end{aligned}
$$
$$\tag{7.8.11}$$

where we used the assumption that $\text{supp}(\vec{\mathbf{N}}(t)) = \text{supp}(\vec{\mathbf{N}}(0))$ always.

Proof of consistency of $\hat{\mathbf{\Psi}}(t)(s_1, s_2)$ for $\Psi(t)(s_1, s_2)$:
Apply the Weak Law of Large Numbers to the first terms of (7.8.8) and (7.8.9) and the
second terms of (7.8.8) and (7.8.9) respectively, and then use the fact $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$ and $\mathbf{Y}_n \xrightarrow{P} \mathbf{Y}$
implies $(\mathbf{X}_n - \mathbf{Y}_n) \xrightarrow{P} \mathbf{X} - \mathbf{Y}$. □

7.8.2.1 Practical Limitations of Identifiability Result

Let $\hat{\mathbf{\Psi}}^F$ denote the "oracle plugin estimator" for $\Psi^F(t)(s_1, s_2)$. (This can not be computed
from the observed data, but an "all-knowing oracle" knows the observed data and the un-
observed data from time 0, and thus can compute plugin estimators that are functions of
$\vec{\mathbf{N}}(0)$.) By the law of large numbers and the continuous mapping theorem we have that the
oracle plugin estimator $\hat{\mathbf{\Psi}}^F$ converges in probability to Ψ^F, basically for the same reason
that the observed data plugin estimator $\hat{\mathbf{\Psi}}$ converges in probability to Ψ.

   Now under the assumptions of Lemma 7.1 we have that *both* the oracle plugin estimator
$\hat{\mathbf{\Psi}}^F$ and the observed data plugin estimator $\hat{\mathbf{\Psi}}$ converge in probability to the Ψ^F (because
$\Psi^F = \Psi$, which again is not true in general). However, this does *not* mean that we also have
that the oracle plugin estimator equals the observed data plugin estimator. Indeed, although
$(\hat{\mathbf{\Psi}}^F - \hat{\mathbf{\Psi}})$ approaches 0 in probability, it is still the case that $(\hat{\mathbf{\Psi}}^F - \hat{\mathbf{\Psi}}) \neq 0$. In particular,
for finite samples the difference between the oracle and observed data plugin estimators will
usually not be zero[14], and it may not even converge *quickly* to 0 in probability.

---

[14]This is related to the multiple representatives problem that is described in section 5.C. Figure 5.C.1
illustrates clearly why the oracle and observed data plugin estimators can differ.

Thus the usefulness of the result of Lemma 7.1 in practice depends on how quickly the random variable $(\hat{\boldsymbol{\Psi}}^F - \hat{\boldsymbol{\Psi}})$ converges to zero in probability. (We would like a finite-sample "with high probability" bound, but an asymptotic convergence rate in probability would not be bad either.) Because the size of the treatment group $\mathbf{T}^{(s_1,s_2)}$ sample will typically grow much more slowly than the size of the control group $\mathbf{C}^{(s_1,s_2)}$ sample, the rate at which difference between the oracle and observed data plugin estimators converges to 0 in probability should probably be bounded by the rate at which the difference[15] $\hat{\boldsymbol{\mu}}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \Big| \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right] - \hat{\boldsymbol{\mu}}\left[\log(N^{(s_2)}(t)) \Big| \operatorname{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{T}}\right]$ converges to 0 in probability. The latter difference only approaches zero in probability as the size of the treatment group sample grows $\mathbf{T}^{(s_1,s_2)}$ without bound. Hence to the extent that $\mathbf{T}^{(s_1,s_2)}$ grows slowly in probability, we should not necessarily expect the difference between the oracle and observed data plugin estimators to converge quickly to 0 in probability.

7.8.2.2 Proofs of Supporting Lemmas

Lemma 7.2. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

- *that $\mathscr{O} \subseteq \mathcal{P}\left([S]\right)$ is such that $s_2 \in \Omega$ for all $\Omega \in \mathscr{O}$,*

then for any measurable $\phi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ one has

$$\mathbb{E}\left[\phi\left(N^{(s_2)}(0)\right) \Big| \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right] = \mathbb{E}\left[\phi\left(N^{(s_2)}(0)\right) \Big| \operatorname{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\right]. \qquad (7.8.12)$$

In particular, for any two subsets $\mathscr{O}_{\mathcal{T}}, \mathscr{O}_{\mathcal{C}} \subseteq \mathcal{P}\left([S]\right)$ of the power set of $[S]$ such that $s_2 \in \Omega_{\mathcal{T}}$ for all $\Omega_{\mathcal{T}} \in \mathscr{O}_{\mathcal{T}}$ and $s_2 \in \Omega_{\mathcal{C}}$ for all $\Omega_{\mathcal{C}} \in \mathscr{O}_{\mathcal{C}}$ one has that

$$\mathbb{E}\left[\phi\left(N^{(s_2)}(0)\right) \Big| \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right] = \mathbb{E}\left[\phi\left(N^{(s_2)}(0)\right) \Big| \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{C}}\right], \qquad (7.8.13)$$

I.e. the value does not depend on the particular subset of the power set, as long as that subset of the power set is such that s_2 is in every set belonging to it.

It may be unnecessary to assume that ϕ is non-negative. I want to preclude potential pathological situations where e.g. the infinite summation is convergent but not absolutely convergent, and ensure that tools like the monotone convergence theorem are available.

---

[15] Note that the quantity $\hat{\boldsymbol{\mu}}\left[\log\left(\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\right) \Big| \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right]$ is also an "oracle estimator" that cannot be computed from the observed data. In fact it is the first term of $\hat{\boldsymbol{\Psi}}^F$.

Proof of Lemma 7.2: Using the law of total expectation, one has that

$$
\begin{aligned}
&\mathbb{E}\Big[\phi\left(N^{(s_2)}(0)\right)\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))\in\mathscr{O}\Big]\\
&=\sum_{\Omega\in\mathscr{O}}\mathbb{E}\Big[\phi(N^{(s_2)}(0))\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\Big]\mathbb{P}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))\in\mathscr{O}\Big)\\
&=\sum_{\Omega\in\mathscr{O}}\mathbb{E}\Big[\phi(N^{(s_2)}(0))\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\{s_2\}\Big]\cdot\mathbb{P}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))\in\mathscr{O}\Big)\\
&=\mathbb{E}\Big[\phi(N^{(s_2)}(0))\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\{s_2\}\Big]\cdot\bigg(\sum_{\Omega\in\mathscr{O}}\mathbb{P}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))\in\mathscr{O}\Big)\bigg)\\
&=\mathbb{E}\Big[\phi\left(N^{(s_2)}(0)\right)\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\{s_2\}\Big]\cdot(1),
\end{aligned}
\tag{7.8.14}
$$

where the second equality follows from Lemma 7.3, the third equality is the distributive property, and the last equality follows because the events are disjoint and exhaustive, thus their probabilities sum to 1. □

Lemma 7.3. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

then for any $\Omega\subseteq[S]$ such that $s\in\Omega$ one has for measurable $\phi:\mathbb{N}\to\mathbb{R}_{\geq0}$ that

$$
\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right)\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\Big]=\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right)\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\{s\}\Big].
\tag{7.8.15}
$$

Proof of Lemma 7.3: Applying the law of total expectation, one has

$$
\begin{aligned}
&\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right)\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\Big]\\
&=\sum_{n=1}^{\infty}\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right)\,\big|\,N^{(s)}(0)=n\Big]\cdot\mathbb{P}\Big(N^{(s)}(0)=n\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\Omega\Big)\\
&=\sum_{n=1}^{\infty}\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right)\,\big|\,N^{(s)}(0)=n\Big]\cdot\mathbb{P}\Big(N^{(s)}(0)=n\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\{s\}\Big)\\
&=\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right)\,\Big|\,\mathrm{supp}(\vec{\mathbf{N}}(0))=\{s\}\Big],
\end{aligned}
\tag{7.8.16}
$$

where the second equality was an application of Lemma 7.4, and the last equality another use of the law of total expectation. □

Lemma 7.4. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

then for any $\Omega \subseteq [S]$ such that $s \in \Omega$, and for any $n \in \mathbb{N}$ (i.e. $n > 0$) one has

$$\mathbb{P}\Big(N^{(s)}(0) = n \;\Big|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega\Big) = \mathbb{P}\Big(N^{(s)}(0) = n \;\Big|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s\}\Big). \qquad (7.8.17)$$

In particular the value does not depend on the particular Ω as long as $s \in \Omega$.

Proof of Lemma 7.4: Noting that the empty product is defined (assumed) to equal 1, e.g. in the case when $\Omega \setminus \{s\} = \emptyset$:

$$
\begin{aligned}
&\mathbb{P}\Big(N^{(s)}(0) \;\Big|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega\Big) \\[4pt]
&= \frac{\mathbb{P}\big(N^{(s)}(0) = n\big) \cdot \displaystyle\prod_{\sigma \in \Omega \setminus \{s\}} \mathbb{P}\big(N^{(\sigma)}(0) \geq 1\big) \cdot \displaystyle\prod_{\sigma \notin \Omega} \mathbb{P}\big(N^{(\sigma)}(0) = 0\big)}{\mathbb{P}\Big(\mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega\Big)} \\[8pt]
&= \frac{\mathbb{P}\big(N^{(s)}(0) = n\big) \cdot \displaystyle\prod_{\sigma \in \Omega \setminus \{s\}} \mathbb{P}\big(N^{(\sigma)}(0) \geq 1\big) \cdot \displaystyle\prod_{\sigma \notin \Omega} \mathbb{P}\big(N^{(\sigma)}(0) = 0\big)}{\mathbb{P}\big(N^{(s)}(0) \geq 1\big) \cdot \displaystyle\prod_{\sigma \in \Omega \setminus \{s\}} \mathbb{P}\big(N^{(\sigma)}(0) \geq 1\big) \cdot \displaystyle\prod_{\sigma \notin \Omega} \mathbb{P}\big(N^{(\sigma)}(0) = 0\big)} \\[8pt]
&= \frac{\mathbb{P}\big(N^{(s)}(0) = n\big)}{\mathbb{P}\big(N^{(s)}(0) \geq 1\big)} \,.
\end{aligned}
\qquad (7.8.18)
$$

The last equality can be seen to clearly not depend on the specific Ω, thus the value is the same as long as $s \in \Omega$. $\square$

7.9 Conclusion

Findings and Contributions I showed how a particular measure of relative fitness can be recast into the statistical framework of average treatment effects. I explained how (for this problem) violations of positivity assumptions are inevitable in practice, making controlling for confounding difficult. I gave explicit assumptions under which the corresponding estimands are identifiable from the observed data produced by incubated droplets.

Practical Implications This chapter demonstrates that this problem has a difficult balance between controlling for confounding and avoiding "data starvation". This makes it particularly important to be able to predict in advance how much data will be available, using the results from part II. By explicitly defining the intended estimands, the specific estimators, and the assumptions required for identifiability from the observed data, this chapter clarifies the goals and limitations of using these estimators.

Next Steps and Open Questions   This chapter points the way forward to formalizing many other statistical estimators that are informally defined in the microbiological literature, even in the context of other experiments. Doing so will hopefully clarify the properties and performance of such estimators. For the analysis of MOREI data specifically, it would be helpful to identify other possible estimands besides those which were described in this chapter. Simulation studies comparing the performance of different estimators would be particularly useful. The next chapter, chapter 8, addresses an important issue that we need to clarify when comparing the performance of different estimators in a simulation study.

Sparsification   An issue with the estimators proposed in this chapter is their "lack of sparsity". The values of the estimators from section 7.6 will almost never be exactly 0 in finite samples, even if they might approach 0 asymptotically. Hence the corresponding estimated interaction network will almost always be a complete graph, even if the magnitudes of many, or even most, of the edges might be small. One approach to address this "lack of sparsity" issue would be to impose an arbitrary magnitude threshold, below which any estimate values would automatically be set to 0. That would be problematic for many reasons, not least of which would include deciding on a choice of threshold value and justifying that choice after it has been made. Another approach would be to throw out the average treatment effects framework altogether and use some modification of the sparse regression and variable selection methods mentioned in e.g. [LR20] to apply to the "gene $\times$ gene" interactions problem we have here. Cf. again item (4) from section 2.1.1. A third approach begins with thinking of sparsity as false discovery control. Then one could attempt to modify the non-parametric method for false discovery control proposed in [GCS$^+$21], originally designed for the "gene $\times$ environment" interactions problem, to apply to the "gene $\times$ gene" interactions problem. This modified method could then be used on the ATE estimators.

Chapter 7.   Appendices

Appendix 7.A   Pre-Existing Notions of Positivity Violations in Microbiology Literature

The contrast between the two kinds of positivity violations mentioned in [PPG$^+$10] appears to be analogous, and perhaps even equivalent, to a similar dichotomy often presented in the existing microbiology literature. When discussing missing data in compositional (i.e. relative abundance) datasets, the authors of [PGB21] and other sources distinguish between "structural zeros" and "sampling zeros". It appears that the "structural zeros" (also called "biological zeros") of [PGB21] correspond to the situation from [PPG$^+$10] of when "it may be theoretically impossible for individuals with certain covariate values to receive a given exposure of interest". Likewise, it appears that the "sampling zeros" (and the related "technical zeros") of [PGB21] correspond to the situation from [PPG$^+$10] of when "violations or near violations of positivity can arise in finite samples due to chance". This suggests that causal inference is relevant to both of these problems previously documented in the microbiology literature. (Although cf. [MGSD19] for a different interpretation of "technical zeros".)

It is also mentioned in [PGB21] how "methods that consider the conditional dependency structure of microbial networks (like probabilistic graphical models) should perform better... because [they have] the power to resolve direct from indirect associations (e.g. associations between two species mediated through a third species)". This appears to correspond to the same motivation underlying confounder adjustment in a causal analysis. The main difference here is that [PGB21] refers to methods using undirected graphical models and outside of any explicit causal framework, whereas causal inference (more typically) relies on methods using directed graphical models (Bayesian networks). Nevertheless, the strength of the analogies here suggests the potential for future work that more directly applies a causal inference framework to make a major, positive impact on the study of microbial interactions.

Appendix 7.B   Combining Estimates from Different Batches

An immediately obvious issue with using the ATE estimands is that they result in different estimates for each batch. Note that, since droplets from different batches were grown for different amounts of time, inherently they are not directly comparable. Especially to the extent that it might be expected that the nature of the interactions could change with time, this could potentially be a useful feature rather than a liability. (See e.g. [HFH21] for previous work tackling such issues head-on in the case of time series[16] i.e. longitudinal data.) Collaborators have expressed interest in possibly estimating such functions of time. Readers interested in such approaches should look into the field of functional data analysis, cf. e.g. the monographs [RHG09], [KR21], [RS05], [HK12], [HE15], or [Gre81] for an introduction. Either of the two reviews [WCM16] or [UF13] may also be helpful, if less thorough.

7.B.1   Criteria for Choosing Aggregation Method

In this preliminary work the goal will be to choose an aggregate summary of these estimates which

1. is structurally and conceptually simple,

2. is easy to implement,

3. sensibly incorporates information from all batches.

This is so that any chosen summary method can be and is likely to be used in practice, even by practitioners without extensive quantitative training.

7.B.2   Why Incorporate Information from All Batches

Given that information from different batches will generally correspond to different microbial growth phases (lag, exponential, stationary), incorporating information from all batches is questionable. After all, arguably from a biological standpoint the growth behavior during exponential phase is what we would most want to quantify. Nevertheless, I still suggest that information from all batches be incorporated, regardless of whether they may be in lag or stationary phases. This is for at least two reasons.

First, on theoretical or "philosophical" grounds, any statistical estimator which is unable to utilize potentially informative data (thus requiring that said data essentially be "thrown out") just because said data may be in some sense "imperfect" for learning about the problem at hand (e.g. potentially lag or stationary phase batches) is "inefficient". In other words, any method that (in general) requires discarding potentially informative data should be avoided on principle if viable alternatives exist.

---

[16] See [CHL+20] for an introduction to time series geared towards microbial ecologists.

Second, on more practical grounds, it could be very inconvenient, frustrating, and time-consuming for the experimenter to run additional statistical analyses on each of the batches to attempt to discern which batches are mostly in what growth phase, as well as to need to repeatedly look up arbitrary rules of thumb for deciding thresholds for e.g. when a batch should be considered "too far" into stationary growth phase and have its corresponding data discarded. For the experimenter the simplest and easiest to use possible answer to the question "Which batches should be used for the data analysis?" will always be "All of them". This is true all the more so when one recalls how there are $O(S^2)$ interactions to consider. Thus, as S becomes large, any ad hoc approach quickly becomes untenable, even if it would have been feasible for a single interaction.

7.B.3   Weighting Contributions from Distinct Batches

Note that the obvious approach of directly averaging over all of the droplets from all of the batches simultaneously does not satisfy the above criteria. Because droplets from different batches are not directly comparable, averaging droplets over all batches to get a single estimate would not be sensible.

While obviously there cannot be any universal or conclusive answer to the question of which frameworks "best" satisfy the aforementioned subjective criteria, another obvious choice would be some convex combination of some fixed function f of the estimands from distinct batches, i.e.

$$\sum_{b=1}^{B} c_b f\left(\hat{\boldsymbol{\Psi}}(t_b)\right), \qquad c_b \geq 0 \text{ for all } b \in [B], \sum_{b=1}^{B} c_b = 1. \tag{7.B.1}$$

Because there are legitimate biological reasons to believe that some batches (those in exponential growth phase) would be more informative than other batches in stationary or lag growth phases, in principle the coefficients c_b in (7.B.1) should be unequally weighted.

However, without an automated and scientifically well-motivated framework for deciding the values of the weights, from a practical perspective this leads to an even worse generalization of the aforementioned problem with ad hoc thresholds for including or removing batches from the data analysis. Specifically, the problem from section 7.B.2 would corresponding to the special case of deciding a subset of the coefficients c_b to set to zero and then assuming the remaining non-zero coefficients correspond to equally informative batches and thus be given equal weights. Again, we have $O(S^2)$ interactions to consider, which means one needs to decide the values of the weights $O(S^2)$ times.

Given the constraint of practicality the only viable way forward[17] appears to be to brazenly assume that the data from all batches is equally informative for describing the microbial interactions. This means equally weighting all of the coefficients in the convex combination (7.B.1). The result of doing so is simply the arithmetic mean of the $f\left(\hat{\boldsymbol{\Psi}}(t_b)\right)$.

---

[17]At least in this preliminary work, where deriving such an automated and scientifically well-motivated framework for deciding the values of the weights is outside of scope.

Given the discussion in section 7.E, this might prompt the question of whether the geometric mean would be a more suitable summary statistic. Note that the estimators $\hat{\boldsymbol{\Psi}}$ correspond effectively to log-transformed data. The primary purpose of the logarithm is argabuly to transform multiplicative relationships in $(0, \infty)$ to additive relationships in $(-\infty, \infty)$. Therefore, provided that the "link function" f is in some way "additive" itself, it seems that an additive summary, as expressed by the arithmetic mean, is more appropriate than the multiplicative summary represented by the geometric mean.

7.B.4   Choice of "Link Function"

This is of course still leaves open the choice of the "link function" f, for which two alternatives are proposed

$$f_1\left(\hat{\boldsymbol{\Psi}}(t_b)\right) = \hat{\boldsymbol{\Psi}}(t_b), \qquad f_2\left(\hat{\boldsymbol{\Psi}}(t_b)\right) = \frac{\hat{\boldsymbol{\Psi}}(t_b)}{||\hat{\boldsymbol{\Psi}}(t_b)||_1}, \tag{7.B.2}$$

where $|| \cdot ||_1$ denotes the entrywise L_1 norm (i.e. equivalent to vectorizing the matrix and then taking the L_1 norm of the resulting vector), and does not denote an operator norm. Cf. section 8.4.1. The choices have been reduced to the (unweighted) average of the unnormalized ATE matrices, or the (unweighted) average of the ATE matrices after they have been normalized.

Starting from the assumption that each batch should be considered equally potentially informative, it follows that the average of the normalized batches is to be preferred. Normalizing the batches beforehand ensures that all of their entries are "on comparable length scales" and thus that every batch has "equal contribution potential". Taking the average of unnormalized batches, and thus implicitly assuming that the data from all batches are directly comparable, also introduces the question of why different batches are comparable but droplets from different batches are not. (Recall that the assumption that droplets from different batches are not comparable is the motivation for these "hierarchical" averaging schemes, rather than just averaging over all droplets from all batches.)

As for the choice of norm, the $\hat{\boldsymbol{\Psi}}(t_b)$ matrices do not represent linear functions, so operator norms are not appropriate. Instead the matrices are (in physics parlance) entirely "contravariant". As for the choice between two common norms for entirely "contravariant" data, namely entrywise-L_1 and entrywise-L_2 (a.k.a. Frobenius), entrywise-L_1 seems more suitable. The L_1 norm is more "robust" to "extreme" entries, and it is additive for non-negative data, thus more "compatible" with the additive summary represented by the arithmetic mean. The geometric intuition related to the Pythagorean theorem that might be used to justify the use of the L_2 norm does not seem applicable here, and in fact could even be used to argue against its use.

7.B.5   Dividing Out Time Heuristic Motivation

Normalization in this context also has the biological interpretation of "revealing information about growth rates" by "dividing out time". For an explanation of this heuristic, first imagine B batches, with each batch b grown until time t_b and containing S droplets, each droplet s starting a time zero with exactly one cell of strain s.

Under the simple exponential growth model (A.1.2.2), for each batch b and strain s one will have for the final number of cells:

$$\log\left(N^{(s)}(t_b)\right) = r_s t_b \,, \tag{7.B.3}$$

where r_s is the "intrinsic growth rate" of strain s. Without first normalizing the values from each batch, the average log count of cells for strain s across all batches is

$$r_s \cdot \frac{t_1 + \cdots + t_B}{B} \,, \tag{7.B.4}$$

which says as much, or possibly more, about the amount of time each of the batches were grown as it does about the growth rate of strain s. In particular, rate estimates from batches which were allowed to grow for longer amounts of time will disproportionately influence the final result. However, if one normalized the entries from each batch first, the value recorded for strain s from any batch b will be

$$\frac{r_s \cdot t_b}{\sum_{\sigma=1}^{S} (r_\sigma \cdot t_b)} = \frac{r_s \cdot t_b}{\left(\sum_{\sigma=1}^{S} r_\sigma\right) \cdot t_b} = \frac{r_s}{\sum_{\sigma=1}^{S} r_\sigma} \,, \tag{7.B.5}$$

which clearly will then also be the average value across all of the batches.

Thus, under this simple (and blatantly inaccurate) model, first normalizing the batches literally "divides out" the amount of time each batch was grown and leaves only information about the intrinsic growth rates from each strain. Moreover, since the resulting value is the same from every batch, the final result corresponds to equal contributions from all batches, rather than the values from batches grown for longer amounts of time contributing disproportionately. The information about the growth rates in (7.B.5) of course doesn't say anything about their absolute scale, but in this context scientifically what is most interesting is "qualitative" information about the growth rates of the strains relative to each other, which clearly is contained in the information from (7.B.5) above.

Even under the more complicated model of bacterial growth (A.1.4.1), and with potentially multiple strains in each droplet, the above argument is still true "to zeroth order". At time t the number of cells of strain s in a given droplet will be a "perturbation" of the amount predicted (7.B.3) by the exponential growth model (A.1.2.2):

$$\log\left(N^{(s)}(t)\right) = r_s t + \delta(t) \,, \quad \delta(t) := \int_0^t \left[\sum_{\sigma=1}^{S} \alpha_{\sigma s} N^{(s)}(\tau) - \frac{G_s(\tau) H_s(\tau)}{K}\right] d\tau \,, \tag{7.B.6}$$

where the notation from equation (A.1.5.1) was used for the purposes of simplification. Arguing using the linearity of expectation and that the ATE estimators and simplified estimands are linear combinations of log counts like those in (7.B.3) and (7.B.6) above, heuristically speaking it still seems reasonable to expect that the averages of the normalized values should usually give at least roughly better insight into the microbial interactions, as defined by their effects on (relative) growth rates, than the average of the unnormalized values. Thus what seems to have been borne out empirically is also something which could have been anticipated a priori on theoretical grounds.

7.B.5.1 Comparison with Previous Work

The argument in this section appears to be very similar to the ideas in [MKBDB00], in particular compare equation (4) and the Appendix of that paper. So although derived independently, the ideas from this section have precedent in the literature. I am unsure whether the definition of fitness that they propose as being better is applicable to this problem. In particular, with pseudolongitudinal data it is much more difficult or even impossible to clearly identify timepoints corresponding to the (average) beginning and end of an exponential growth phase. (This was already alluded to in sections 7.B.2 and 7.B.3.) The methods suggested in [MKBDB00] and its predecessor [HdlTCD91] both seem to assume that one can clearly identify such timepoints. Valuable future work might investigate the feasibility of identifying such average timepoints from pseudolongitudinal data. That being said, as alluded to already in sections 7.B.2 and 7.B.3, even with such a method for pseudolongitudinal data implementation could still be difficult for the computational/combinatorial reasons that there are $O(S^2)$ interactions for which this method would need to be applied.

7.B.6 Treatment of Missing Data

The above discussion has so far neglected to consider how in many instances estimates for certain interactions will be missing from certain batches. When normalizing the matrices, the missing entries are treated as if they were found to exactly equal zero. Zero seems to be the most conservative possible choice to impute for missing interaction estimates. In the absence of any evidence for an interaction, the choice is made to prefer the possibility of making the false negative error of assuming no interaction exists when one might exist, rather than possibly making the false positive error of assuming an interaction exists when in fact none exists. Cf. the discussion from appendix A.3.

Afterwards, when computing the average of the normalized matrices, the "entrywise average" is taken, i.e. for each entry the sum of the non-missing values across all batches (equal to the sum of the values over all batches after imputing missing estimates as zero) is divided by the number of batches for which an estimate for the corresponding interaction was not missing (which of course is only strictly less than the total number of batches when an estimate for that interaction is missing from at least one batch). Cf. figure 7.B.1.
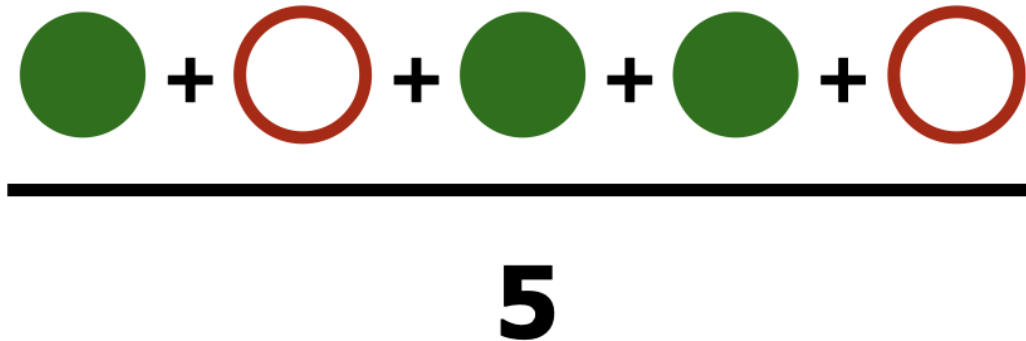
Figure 7.B.1: Comparison of "entrywise average" with taking the average of the corresponding matrices.

Compared to dividing the sum of the values (where again the missing values are being treated as the additive identity, zero) by the total number of batches, this should in principle avoid artificially "deflating" some entries in the aggregate matrix just because estimates for the corresponding interaction were missing from some batches. (Of course to the extent that such values could be considered more unreliable, one could plausibly also argue that such "artificial" deflation/shrinkage of those values would be a feature, not a bug. Without modification however such an argument neglects to consider how the estimates from each batch are themselves averages whose own reliabilities are themselves contingent upon the number of droplets over which the average was formed.) This corresponds again to the assumption that all batches have "equal potential to be informative". Thus e.g. in the extreme case that an estimate for a given interaction was available only from one batch, that single estimate should be used "as is" in the final aggregate matrix, rather than being deflated or otherwise modified.

Preliminary investigations of simulations (data not shown) seemed to suggest that this method of handling the missing estimates leads to results which are perhaps slightly more effective (at least in terms of Spearman correlation or relative error) than simply dividing all entries by the total number of batches regardless of the number of missing values. Any difference was difficult to notice however. Because the chosen approach would have because of the above arguments, regardless of empirical performance, any potential empirical performance differences were not examined in any further detail.

Appendix 7.C   Heuristic Derivation

Basically the scientific goal is to compare the "typical growth of strain s_2 in the presence of strain s_1" to the "typical growth of strain s_2 (in general)". The idea is that any effect strain s_1 has on the growth of strain s_2 via their interactions will be the only, or at least primary, latent variable explaining any discrepancy between the aforementioned two types of "typical growth", and that absence of evidence for any discrepancy can probably be construed as evidence of absence of any underlying interaction causing strain s_1 to affect strain s_2. Cf. again sections 2.3.4 and 1.3, or the discussion of "competition assays" from the introduction to Part I. As noted before, previous work such as [MXS17] or [BKK16] has raised concerns about potential problems with assuming all interactions are pairwise interactions.

One way to define "growth" is the multiplicative factor by which the number of cells of a given strain in a droplet has increased relative to its initial population after the predetermined time for its batch has elapsed, with comparison of the growth factors defined via a ratio:

$$\frac{\left(\text{``typical growth factor } \dfrac{N^{(s_2)}(t_b)}{N^{(s_2)}(0)} \text{ of cells of strain } s_2 \text{ in the presence of strain } s_1 \text{ for batch } b\text{''}\right)}{\left(\text{``typical growth factor } \dfrac{N^{(s_2)}(t_b)}{N^{(s_2)}(0)} \text{ of cells of strain } s_2 \text{ in general for batch } b\text{''}\right)}. \tag{7.C.1}$$

However there are several reasons to define "growth" in terms of the logarithms of the aforementioned quantities instead. First, doing so accounts for the fact that in the eutrophic regime the number of cells will grow more rapidly the more cells there already are present simply by virtue of the fact that there are more cells present to divide, whereas the logarithm of the number of cells will continue to grow at a constant rate (compare equation A.1.2.2 from appendix A.1). Second, the logarithm corresponds to the natural assumption that e.g. "causing to grow twice as many cells" should be considered the opposite of "causing to grow half as many cells" as a consequence of the identity $\log\left(\frac{1}{x}\right) = -\log(x)$. Finally, the logarithm transforms multiplicative relationships in $(0, \infty)$, like (7.C.1) above, to additive relationships on $(-\infty, \infty)$, the latter being easier to work with in general for numerous reasons. At the same time the logarithm "preserves order", in the sense that the logarithm of one number is larger than that of another number if and only if the original number was larger than the other original number, in other words $\log(y) > \log(x)$ if and only if $y > x$, thus log-transforming the data can be said to preserve "relationships of relative size" (i.e. which values are largest, which are smallest, etc.). After deciding to log transform, the scientific goal is reduced to the analysis of quantities of the form:

$$
\begin{pmatrix} \text{`` typical logarithm growth factor of cells of strain } s_2 \\ \text{in the presence of strain } s_1 \text{ for batch } b\text{''} \end{pmatrix}
$$
$$
- \left(\text{`` typical logarithm growth factor of cells of strain } s_2 \text{ in general for batch } b\text{''} \right) ,
$$
(7.C.2)

so-called "log-ratios" (because they are logarithms of ratios, while in contrast they are not ratios of logarithms). Note that values of the quantities from (7.C.2) greater than zero correspond to interactions where the presence of s_1 promotes the growth of s_2, values less than zero correspond to interactions where the presence of s_1 inhibits the growth of s_2, and values (near) zero correspond to "null interactions" where the growth of s_2 is the same regardless of the presence of s_1. Cf. the coefficients $\alpha_{s_1 s_2}$ from equation (2.2.15) of section 2.2.5 as well as the related discussion from sections 1.2 and 1.3.

7.C.1   Issues with Heuristic Definition

Even after reducing the scientific goal to the analysis of quantities of the form (7.C.2), there are still many issues with trying to make this actually precise enough to use as a basis for data analysis. Beginning with definitional issues, i.e. how to rephrase (7.C.2) so that it may be able to define a concrete statistical estimand, there are at least three such issues:

1. Does "typical logarithm" correspond to the logarithm of the expectation or to the expectation of the logarithm?

2. Which droplets correspond to "in the presence of s_1"?

3. Which droplets correspond to "the typical growth of s_2 in general"?

7.C.2   Discussion

Referring to the beginning of section 7.C.1, note that the way the first question ('Does "typical logarithm" correspond to the logarithm of the expectation or the expectation of the logarithm?') is phrased already implicitly assumes that one has decided that "typical" signifies the use of expected value rather than any other possible distributional summary (e.g. the median or mode). This assumption was made largely, as one might expect, for the sake of simplicity of analysis, i.e. so that it is possible to appeal directly to a law of large numbers (combined with the continuous mapping theorem in the case of the logarithm of the expectation) to guarantee the (theoretical) statistical consistency of the corresponding plug-in estimators (i.e. the values of the estimands with respect to the empirical distribution). Of course in relying on (the standard) laws of large numbers there is also the implicit use of the assumption that all of the droplets in the experiment are mutually statistically independent, an assumption which is most likely to be very questionable at best in practice. In the absence of obviously better alternative models to use (i.e. as opposed to alternative models which would be more complicated, much more difficult to analyze theoretically, and likely even less realistic than assuming the droplets are independent), and since the goal of this work is to introduce this problem for the first time, such simplifying assumptions seemed forgivable if not exactly warranted. Nevertheless worthwhile possible goals for future work would be to examine, from an empirical and/or theoretical perspective, the consequences of either choosing a distributional summary besides the expected value to define "typical", or the consequences of relaxing the assumption that all of the individual droplets are mutually statistically independent.

Appendix 7.D   Calculations for hPoMu Working Model

Lemma 7.4 allows us to perform some potentially useful calculations in the case of the particularly simple hPoMu working model.

Corollary 1. *For any two subsets $\Omega_1, \Omega_2 \subseteq [S]$ such that $s \in \Omega_1 \cap \Omega_2$, for all $n \in \mathbb{N}$ one has under the hPoMu distribution that*

$$
\begin{aligned}
&\mathbb{P}\left(N^{(s)}(0) = n \,\middle|\, \text{supp}(\vec{\mathbf{N}}(0)) = \Omega_1 \right) \\
&= \frac{\dfrac{(f_s\lambda)^n}{n!}}{e^{f_s\lambda} - 1} \\
&= \mathbb{P}\left(N^{(s)}(0) = n \,\middle|\, \text{supp}(\vec{\mathbf{N}}(0)) = \Omega_2 \right).
\end{aligned}
\tag{7.D.1}
$$

Proof of Corollary 1: Because $N^{(s)}(0) \sim \text{Pois}(f_s\lambda)$ and from e.g. the law of total proba-

bility:

$$\frac{\mathbb{P}\big(N^{(s)}(0) = n\big)}{\mathbb{P}(N^{(s)}(0) \geq 1)} = \frac{e^{-f_s\lambda}\left(\frac{(f_s\lambda)^n}{n!}\right)}{e^{-f_s\lambda}\left(\sum\limits_{m=1}^{\infty} \frac{(f_s\lambda)^m}{m!}\right)} = \frac{\left(\frac{(f_s\lambda)^n}{n!}\right)}{e^{f_s\lambda} - 1} . \quad \square \tag{7.D.2}$$

Corollary 2. *For any two subsets $\Omega_1, \Omega_2 \subseteq [S]$ with $s \in \Omega_1 \cap \Omega_2$, under the hPoMu distribution one has for any measurable $\phi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ the following formula:*

$$\mathbb{E}\Big[\phi\left(N^{(s)}(0)\right) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega_1\Big] = \frac{\sum\limits_{n=1}^{\infty} \phi(n) \frac{(f_s\lambda)^n}{n!}}{e^{f_s\lambda} - 1} \tag{7.D.3}$$

$$= \mathbb{E}\Big[\phi\left(N^{(s)}(0)\right) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega_2\Big].$$

Proof of Corollary 2: Using the law of total expectation and Corollary 1:

$$\sum_{n=1}^{\infty} \mathbb{E}\Big[\phi\left(N^{(s)}(0)\right) \,\Big|\, N^{(s)}(0) = n\Big] \cdot \mathbb{P}\Big(N^{(s)}(0) = n \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s\}\Big)$$

$$= \sum_{n=1}^{\infty} \phi(n) \cdot \mathbb{P}\Big(N^{(s)}(0) = n \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s\}\Big) \tag{7.D.4}$$

$$= \frac{\sum\limits_{n=1}^{\infty} \phi(n) \frac{(f_s\lambda)^n}{n!}}{e^{f_s\lambda} - 1} . \quad \square$$

Some special cases of Corollary 2 include the case where $\phi(n) = n$:

$$\mathbb{E}\Big[N^{(s)}(0) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega\Big] = \frac{f_s\lambda e^{f_s\lambda}}{(e^{f_s\lambda} - 1)} , \tag{7.D.5}$$

as well as the case where $\phi(n) = \frac{1}{n}$:

$$\mathbb{E}\Big[\frac{1}{N^{(s)}(0)} \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) = \Omega\Big] = \frac{\mathrm{Ein}\,(f_s\lambda)}{(e^{f_s\lambda} - 1)} , \tag{7.D.6}$$

where $\mathrm{Ein}\,(x)$ refers to a modified version of the exponential integral function, given by the power series:

$$\sum_{n=1}^{\infty} \frac{x^n}{n \cdot n!} . \tag{7.D.7}$$

The analytic function (7.D.7) can be calculated in terms of the usual exponential integral function and the Euler-Mascheroni constant, both of which are implemented for example in SciPy[VGO+20].

Appendix 7.E   Related (Less Useful) Estimands

Regarding the first issue from section 7.C.1, the method of taking expectations first followed by logarithms second will be called "arithmetic" estimands and denoted $\Psi^{(a)}(t)(s_1, s_2)$ ("(a)" for "**a**rithmetic"). The term "arithmetic" refers to how these estimands correspond to the logarithms of the arithmetic means of the relevant growth factors.

The method of taking logarithms first followed by expectations second corresponds to the ATE estimands. For purposes of comparing the ATEs with the "arithmetic" estimands, the ATEs can be called "geometric" estimands and denoted $\Psi^{(g)}(t)(s_1, s_2)$ ("(g)" for "**g**eometric"), although this will not be needed below. The term "geometric" refers to how these estimands correspond to the logarithms of the geometric means of the relevant growth factors.

The "arithmetic" estimands are less useful than the "geometric" estimands not only for theoretical reasons, such as how more stringent assumptions (compared to the "geometric" estimands) are required for identifiability from the observed data, or their lack of interpretation as an ATE. Preliminary simulation studies also found them to lead to less accurate or reliable estimators of microbial interactions. This was probably due not only to a larger "identification gap" but probably also to the general feature of geometric means typically being more robust to outliers than arithmetic means. However, more studies would be required to identify definitive reasons for what appears to be their poorer performance. Nevertheless, for all intents and purposes I would only recommend using the ATE estimands in practice.

If for no other reason than to demonstrate that it is hypothetically possible to use them, the "arithmetic" estimands are discussed in detail below. Like the ATEs, the "arithmetic" estimands also have versions corresponding to every possible combination of treatment and control groups. Thus, including the possibility of "arithmetic" estimands doubles the number of estimands to potentially consider from 4 to 8.

(Full Data) Picky treatment, picky control:

$$
\begin{aligned}
\Psi^{F^{(a)}_{p/p}}(t)(s_1, s_2) : \quad & \mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi^{F^{(a)}_{p/p}}(t)(s_1, s_2) : \quad & \mathcal{P}_F \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\tfrac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\right]\right) \\
& \qquad - \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\tfrac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\right]\right)
\end{aligned}
\tag{7.E.1}
$$

(Full Data) Picky treatment, gluttonous control:

$$
\begin{aligned}
\Psi^{F^{(a)}_{p/g}}(t)(s_1, s_2) : \quad & \mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi^{F^{(a)}_{p/g}}(t)(s_1, s_2) : \quad & \mathcal{P}_F \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\tfrac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_1, s_2\}\right]\right) \\
& \qquad - \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\tfrac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\middle|\;\mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\right]\right)
\end{aligned}
\tag{7.E.2}
$$

(Full Data) Gluttonous treatment, picky control:

$$
\begin{aligned}
\Psi^{F^{(a)}}_{g/p}(t)(s_1, s_2): \quad &\mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi^{F^{(a)}}_{g/p}(t)(s_1, s_2): \quad &\mathcal{P}_F \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\left.\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right]\right) \\
&\quad - \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\left.\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) = \{s_2\}\right]\right)
\end{aligned}
\tag{7.E.3}
$$

(Full Data) Gluttonous treatment, gluttonous control:

$$
\begin{aligned}
\Psi^{F^{(a)}}_{g/g}(t)(s_1, s_2): \quad &\mathcal{M}^F \;\rightarrow\; \mathbb{R} \\
\Psi^{F^{(a)}}_{g/g}(t)(s_1, s_2): \quad &\mathcal{P}_F \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\left.\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_1, s_2\}\right]\right) \\
&\quad - \log\left(\mathbb{E}_{\mathcal{P}_F}\left[\left.\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(0))\right]\right)
\end{aligned}
\tag{7.E.4}
$$

Note that due to the different signs of the treatment and control terms, Jensen's inequality does not enable any general comparisons or bounds between the "arithmetic" estimands and their "geometric" counterparts in general.

7.E.1 Observed Data Estimand and Estimator Definitions

Just like their "geometric" counterparts, the full data "arithmetic" estimands also possess observed data counterparts, in the obvious way. See below.

(Observed Data) Picky treatment, picky control:

$$
\begin{aligned}
\Psi^{(a)}_{p/p}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\rightarrow\; \mathbb{R} \\
\Psi^{(a)}_{p/p}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_O}\left[\left.N^{(s_2)}(t)\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_1, s_2\}\right]\right) \\
&\quad - \log\left(\mathbb{E}_{\mathcal{P}_O}\left[\left.N^{(s_2)}(t)\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\right]\right)
\end{aligned}
\tag{7.E.5}
$$

(Observed Data) Picky treatment, gluttonous control:

$$
\begin{aligned}
\Psi^{(a)}_{p/g}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\rightarrow\; \mathbb{R} \\
\Psi^{(a)}_{p/g}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_O}\left[\left.N^{(s_2)}(t)\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_1, s_2\}\right]\right) \\
&\quad - \log\left(\mathbb{E}_{\mathcal{P}_O}\left[\left.N^{(s_2)}(t)\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\right]\right)
\end{aligned}
\tag{7.E.6}
$$

(Observed Data) Gluttonous treatment, picky control:

$$
\begin{aligned}
\Psi^{(a)}_{g/p}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\rightarrow\; \mathbb{R} \\
\Psi^{(a)}_{g/p}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_O}\left[\left.N^{(s_2)}(t)\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\right]\right) \\
&\quad - \log\left(\mathbb{E}_{\mathcal{P}_O}\left[\left.N^{(s_2)}(t)\;\right|\; \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\right]\right)
\end{aligned}
\tag{7.E.7}
$$

(Observed Data) Gluttonous treatment, gluttonous control:

$$
\begin{aligned}
\Psi_{g/g}^{(a)}(t)(s_1, s_2): \quad &\mathcal{M}^O \;\to\; \mathbb{R} \\
\Psi_{g/g}^{(a)}(t)(s_1, s_2): \quad &\mathcal{P}_O \;\mapsto\; \log\left(\mathbb{E}_{\mathcal{P}_O}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\right]\right) \\
&\qquad - \log\left(\mathbb{E}_{\mathcal{P}_O}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\right]\right)
\end{aligned}
\tag{7.E.8}
$$

For a given observed data "arithmetic" estimand $\Psi^{(a)}$, let its corresponding estimator be denoted $\hat{\boldsymbol{\Psi}}^{(a)}$. Then the (plugin) estimators corresponding to each "arithmetic" estimand are explicitly defined below.

(Estimator) Picky treatment, picky control:

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{p/p}^{(a)}(t)(s_1, s_2) := &\log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_1, s_2\}\right]\right) \\
&- \log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\right]\right)
\end{aligned}
\tag{7.E.9}
$$

(Estimator) Picky treatment, gluttonous control:

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{p/g}^{(a)}(t)(s_1, s_2) := &\log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_1, s_2\}\right]\right) \\
&- \log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\right]\right)
\end{aligned}
\tag{7.E.10}
$$

(Estimator) Gluttonous treatment, picky control:

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{g/p}^{(a)}(t)(s_1, s_2) := &\log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\right]\right) \\
&- \log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) = \{s_2\}\right]\right)
\end{aligned}
\tag{7.E.11}
$$

(Estimator) Gluttonous treatment, gluttonous control:

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{g/g}^{(a)}(t)(s_1, s_2) := &\log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_1, s_2\}\right]\right) \\
&- \log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\Big|\, \mathrm{supp}(\vec{\mathbf{N}}(t)) \supseteq \{s_2\}, s_1 \notin \mathrm{supp}(\vec{\mathbf{N}}(t))\right]\right)
\end{aligned}
\tag{7.E.12}
$$

The full data "arithmetic" estimands can still be identifiable from the observed data, but only under assumptions which are stronger than those required for the full data ATE ("geometric") estimands. Cf. Lemma 7.6 or Lemma 7.7.

7.E.2 Sufficient Conditions for Identifiability

Observe how Lemma 7.5 below corresponds to the "arithmetic" estimands in exactly the same way that Lemma 7.1 above corresponds to the ATE estimands.

Lemma 7.5. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

- *that the distributions of the droplets are independent,*

- *the assumption (7.E.16) from Lemma 7.6 and/or the assumption (7.E.19) from Lemma 7.7,*

- *we always have that $\operatorname{supp}(\vec{\mathbf{N}}(t)) = \operatorname{supp}(\vec{\mathbf{N}}(0))$ (cf. section 7.8.1),*

then for any two subsets $\mathscr{O}_{\mathcal{T}}, \mathscr{O}_{\mathcal{C}} \subseteq \mathcal{P}\left([S]\right)$ of the power set of $[S]$ such that $s_2 \in \Omega_{\mathcal{T}}$ for all $\Omega_{\mathcal{T}} \in \mathscr{O}_{\mathcal{T}}$ and $s_2 \in \Omega_{\mathcal{C}}$ for all $\Omega_{\mathcal{C}} \in \mathscr{O}_{\mathcal{C}}$, then the estimator

$$\hat{\boldsymbol{\Psi}}^{(a)}(t)(s_1, s_2) := \log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{T}}\right]\right)$$
$$- \log\left(\hat{\boldsymbol{\mu}}\left[N^{(s_2)}(t) \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{C}}\right]\right) \tag{7.E.13}$$

is statistically consistent for the estimand

$$\Psi^{F^{(a)}}(t)(s_1, s_2) := \log\left(\mathbb{E}\left[\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right]\right)$$
$$- \log\left(\mathbb{E}\left[\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{C}}\right]\right). \tag{7.E.14}$$

Note that $\mathscr{O}_{\mathcal{T}}$ and $\mathscr{O}_{\mathcal{C}}$ have the same interpretations here as in section 7.8.

Proof of Lemma 7.5:

Proof of identifiability of $\Psi^{F^{(a)}}(t)(s_1, s_2)$ from the observed data:
Using the result of Lemma 7.6 and/or Lemma 7.7 below one immediately has as a consequence that the expression (7.E.14) equals

$$
\begin{aligned}
\Psi^{F^{(a)}}(t)(s_1, s_2) &= \log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{T}}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_{\mathcal{C}}\right]\right) \\
&= \log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{T}}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \operatorname{supp}(\vec{\mathbf{N}}(t)) \in \mathscr{O}_{\mathcal{C}}\right]\right) \\
&= \Psi^{(a)}(t)(s_1, s_2).
\end{aligned}
$$
$$\tag{7.E.15}$$

In other words, the full data "arithmetic" estimand equals its observed data counterpart, $\Psi^{F^{(a)}}(t)(s_1, s_2) = \Psi^{(a)}(t)(s_1, s_2)$. The penultimate equality used the assumption that we always have $\operatorname{supp}(\vec{\mathbf{N}}(t)) = \operatorname{supp}(\vec{\mathbf{N}}(0))$.

Proof of consistency of $\hat{\Psi}^{(a)}(t)(s_1, s_2)$ for $\Psi^{(a)}(t)(s_1, s_2)$:
Then using the Weak Law of Large Numbers, the expression inside the logarithm for the first term of (7.E.13) converges in probability to the expression inside of the logarithm for the first term of (7.E.14), and then applying the Continuous Mapping Theorem with the continuous function log allows one to conclude that the first term of (7.E.13) converges in probability to the first term of (7.E.14). Completely analogous reasoning leads one to also conclude that the second term of (7.E.13) also converges in probability to the second term of (7.E.14). The result then follows using the fact $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$ and $\mathbf{Y}_n \xrightarrow{P} \mathbf{Y}$ implies $(\mathbf{X}_n - \mathbf{Y}_n) \xrightarrow{P} \mathbf{X} - \mathbf{Y}$. $\square$

Lemma 7.6. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

- *that for any $\mathscr{O} \subseteq \mathcal{P}([S])$ such that $s_2 \in \Omega$ for all $\Omega \in \mathscr{O}$, one has*

$$
\mathbb{E}\left[\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right]
$$
$$
= \mathbb{E}\left[N^{(s_2)}(t) \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right] \cdot \mathbb{E}\left[\frac{1}{N^{(s_2)}(0)} \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right]
$$

$$(7.E.16)$$

then for any two subsets $\mathscr{O}_\mathcal{T}, \mathscr{O}_\mathcal{C} \subseteq \mathcal{P}([S])$ of the power set of $[S]$ such that $s_2 \in \Omega_\mathcal{T}$ for all $\Omega_\mathcal{T} \in \mathscr{O}_\mathcal{T}$ and $s_2 \in \Omega_\mathcal{C}$ for all $\Omega_\mathcal{C} \in \mathscr{O}_\mathcal{C}$ one has that the expression (7.E.14) equals

$$
\log\left(\mathbb{E}\left[N^{(s_2)}(t) \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(t) \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C}\right]\right). \quad (7.E.17)
$$

Proof of Lemma 7.6: Using assumption (7.E.16), as well as the properties of logarithms, one has that (7.E.14) above equals

$$
\log\left(\mathbb{E}\left[N^{(s_2)}(t) \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(t) \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C}\right]\right)
$$
$$
+ \left(\log\left(\mathbb{E}\left[\frac{1}{N^{(s_2)}(0)} \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T}\right]\right) - \log\left(\mathbb{E}\left[\frac{1}{N^{(s_2)}(0)} \;\middle|\; \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C}\right]\right)\right).
$$

$$(7.E.18)$$

By Lemma 7.2 using $\phi(n) = \frac{1}{n}$, the expression inside of parentheses in (7.E.18) vanishes, since of course $x = y$ also implies that $\log(x) = \log(y)$ (the analogous statement is true for literally any function), and thus the result follows. $\square$

The condition (7.E.16) is equivalent to stating that $N^{(s_2)}(t)$ and $N^{(s_2)}(0)^{-1}$ are conditionally uncorrelated given that $\mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}$. A sufficient (but obviously not necessary) condition for this hypothesis to be true is when $N^{(s_2)}(t)$ and $N^{(s_2)}(0)$ are conditionally independent when given that $\mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}$, since this in turn implies that $N^{(s_2)}(t)$ and $N^{(s_2)}(0)^{-1}$ are conditionally independent given $\mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}$, which in turn implies the

conditional uncorrelatedness corresponding to the hypothesis. (Of course, as is probably obvious to the reader, $N^{(s_2)}(t)$ and $N^{(s_2)}(0)$ being at all independent, even if only conditionally, seems like a rather implausible assumption.)

Lemma 7.7. *Assuming that*

- *the marginal distributions of $\vec{\mathbf{N}}(0)$ are mutually independent, e.g. as turns out to be true for the hPoMu and hNBDM distributions,*

- *that for any $\mathscr{O} \subseteq \mathcal{P}([S])$ such that $s_2 \in \Omega$ for all $\Omega \in \mathscr{O}$, one has*

$$\mathbb{E}\left[\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)} \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right] = \frac{\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right]}{\mathbb{E}\left[N^{(s_2)}(0) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}\right]} \qquad (7.E.19)$$

then for any two subsets $\mathscr{O}_\mathcal{T}, \mathscr{O}_\mathcal{C} \subseteq \mathcal{P}([S])$ of the power set of $[S]$ such that $s_2 \in \Omega_\mathcal{T}$ for all $\Omega_\mathcal{T} \in \mathscr{O}_\mathcal{T}$ and $s_2 \in \Omega_\mathcal{C}$ for all $\Omega_\mathcal{C} \in \mathscr{O}_\mathcal{C}$ one has that the expression (7.E.14) equals

$$\log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C}\right]\right). \quad (7.E.20)$$

Proof of Lemma 7.7: Using assumption (7.E.19) and the properties of logarithms, one has that (7.E.14) above equals

$$\log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(t) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C}\right]\right)$$
$$- \left(\log\left(\mathbb{E}\left[N^{(s_2)}(0) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{T}\right]\right) - \log\left(\mathbb{E}\left[N^{(s_2)}(0) \,\middle|\, \mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}_\mathcal{C}\right]\right)\right).$$
$$(7.E.21)$$

Using Lemma 7.2 with $\phi(n) = n$, it follows as before that the expression inside of parentheses in (7.E.21) vanishes, and thus the result follows. $\square$

As one can see by multiplying both sides of (7.E.19) by the denominator of the right side of (7.E.19), the condition (7.E.19) is equivalent to assuming the conditional uncorrelatedness of the multiplicative growth factor $\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}$ and the initial cell count $N^{(s_2)}(0)$ when given that $\mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}$. A sufficient, but again obviously not necessary, condition for this to be true is when the factor by which the population has grown by time t, $\frac{N^{(s_2)}(t)}{N^{(s_2)}(0)}$, and the initial cell count $N^{(s_2)}(0)$ are conditionally independent when given that $\mathrm{supp}(\vec{\mathbf{N}}(0)) \in \mathscr{O}$. While this still sounds implausible, it is perhaps less so than the above sufficient condition for (7.E.16).

Chapter 8

Loss Functions for Ecological Networks

Dissimilarity and Similarity Functions for Quantitatively

Comparing Networks with Mixed-Sign Edge Weights

Herein I discuss how comparisons of ecological interactions can be recast into the statistical framework of loss functions for signed networks. See sections 8.2 and 8.B.1. I explain how avoiding unexpected behavior requires loss functions for signed networks to satisfy what I call "the double penalization principle". See, for example, section 8.2.5, section 8.6, or even section 8.B.2. Starting from loss functions of unsigned networks, I derive several examples of loss functions for signed networks that satisfy this property. See section 8.4.

Section 8.2 clarifies the problem and identifies principles to judge whether a method sensibly quantifies the (dis)similarity of two networks with mixed-sign edge weights. Section 8.3 provides technical definitions which facilitate the analysis in the rest of the chapter. In section 8.4 I propose several examples of methods that satisfy the principles described in section 8.2. Section 8.6 demonstrates empirically how methods that satisfy these principles behave compared to methods that do not satisfy these principles. Finally section 8.7 interprets the results and explains which facets of the problem still require further attention.

Chapter 8 Contents

8.1   Background and Significance

Broader field   The broader field of chapter 8 is the use of signed networks [Har53] as models for ecological interactions. See [FR12] for a review of network models of interactions in microbial ecology, and [Lid79] or [GHG+18] for how signs (of network edge weights) can be used to categorize ecological interactions. Positive edges are for interactions that promote the growth of the recipient, and negative edges are for interactions that suppress the growth of the recipient, cf. again section 1.2. Cf. also the discussion of this earlier from the introduction to Part III. These networks would be inferred using the data from MOREI.

In particular, chapter 8 is motivated by the important issue of how to identify "best-performing" or "most accurate" estimators of such ecological models. Our statistical estimation problem has a network with mixed-sign edge weights as its estimand (cf. again section 1.2), the microbial interaction network [XAF+17] [AML19]. The issue of what it means for an estimator to be "best-performing" is already subtle in the thoroughly studied

field of point estimation (cf. e.g. [LC98] or [Kee10]), where the estimand is a much simpler data structure than a signed network. Hence this issue deserves some thought.

Specific problem   Herein I investigate how comparisons of ecological interactions can be recast into the statistical framework of loss functions for signed networks. Given multiple choices of estimator, or "algorithm", for characterizing signed networks of ecological interactions from MOREI data, we want to know how to decide which choice is best or "optimal". Chapter 7 established that at least one such choice exists, so the problem is not "vacuous". To ensure that the estimates produced by any "optimal" choice are useful, the process for deciding which estimator is "optimal" must itself be a "reasonable" process.

To decide between estimators, we need to compare their performance. The performance can be defined as how (dis)similar the estimates produced by the estimators are to a reference "ground truth" signed network. (Such a "ground truth" network is available, for example, when running a simulation and evaluating the performance of the estimators on the simulated data[1]. One can also compare with the results inferred from data produced by a more established experiment, such as cell plating, defining those results operationally as the "ground truth".) A function that quantifies the dissimilarity of an estimate to a "ground truth" estimand is called a loss function. (Cf. the discussion of terminology in section 8.B.1.) An "optimal" choice of estimator is then one whose estimates (typically) minimize the value of the loss function with respect to the "ground truth". Therefore, making "reasonable" choices of "optimal" estimators requires us to identify and use a "reasonable" loss function.

Reviews in the literature exist discussing possible choices of such functions in the "classical" case of *un*signed networks. The review [TITP19] considers functions that are applicable to directed[2] (unsigned) networks, whereas the review [WM20] only considers functions that are applicable for undirected (unsigned) networks. However, no previous literature appears to exist discussing loss functions for *signed* networks. If it does, I was unable to find it.

Particular approach   Because loss functions for signed networks do not seem to be studied in the literature, the approach of this chapter is to provide a baseline for future work in this area. Principles are defined to characterize the behavior of "reasonable" loss functions for signed networks, and concrete examples of loss functions that obey these principles are given.

To some extent, how to optimally choose such an algorithm is already obvious: as explained in [vdLR11, chapter 3], given multiple possible choices of algorithm, discrete SuperLearner can identify a best possible choice, and SuperLearner [vdLPH07] can give us a weighted combination of algorithms that performs even better. However, to make any such choice even SuperLearner needs a specified choice of loss function [vdLR11, section 3.5].

---

[1] This is another reason why the work in Part II is important. If our simulations do not accurately characterize how much data will be available to infer the edges of the interaction network, whatever conclusions we draw based on that simulation about the effectiveness of various estimators are likely to be distorted.

[2] And thus by extension are also applicable to undirected networks, because undirected networks can be formulated as a special case of directed networks in a standard way.

Indeed, even regardless of whether one uses SuperLearner, choosing an "optimal" estimator requires us to first specify our preferences through the choice of a loss function. If the chosen loss function behaves "unreasonably" in some way, then the resulting choice of "optimal" estimator will be similarly "unreasonable". Quoting from [APRL$^+$20]:

> The criteria to define what is optimal depends on what quantity is to be estimated. Hence, it is first necessary to define an objective or utility function, i.e. the measure by which we evaluate the performance of any given design.

Thus the challenge for this chapter is to do this explicitly for a type of estimand, signed networks, that appears to not have been studied for this purpose before in the literature.

8.1.1 Applications of Signed Networks

Observe how a framework using signed networks has the flexibility to model all possible kinds of ecological interactions, not just trophic interactions[3] where the growth of one organism necessarily comes at the expense of another organism. Previous work [LPR19] has discussed comparing such trophic interaction networks[4]. In such a network signed edge weights are unnecessary because an edge from s_1 to s_2 indicates s_1 promotes the growth of s_2 and s_2 inhibits the growth of s_1 (necessarily because s_2 consumes s_1), while swapping the direction of the edge to be from s_2 to s_1 indicates the opposite relationship, that s_2 promotes the growth of s_1 and s_1 inhibits the growth of s_2 (due to s_1 consuming s_2). See [MJS20] for further discussion of trophic interaction networks as unsigned, directed networks, analysis of the underlying theory, and comparison with structurally similar problems.

In the signed network framework (cf. [Lid79], [GHG$^+$18], or even [KOK$^+$20]), the former corresponds to $s_1 \overset{+}{\to} s_2$ and $s_2 \overset{-}{\to} s_1$, while the latter corresponds to $s_2 \overset{+}{\to} s_1$ and $s_1 \overset{-}{\to} s_2$. Note also that the signed network framework does not require the assumption that the cause of this asymmetric interaction is one organism consuming the other. For example, humans promote the growth of rats because human waste serves as a food source for rats, while rats can inhibit the growth of humans by serving as a vector for human diseases. The signed network framework also allows modelling mutualistic/mutually supporting interactions (pairs of interactions of the form $s_1 \overset{+}{\to} s_2$ and $s_2 \overset{+}{\to} s_1$), competitive/mutually antagonistic interactions (pairs of interactions of the form $s_1 \overset{-}{\to} s_2$ and $s_2 \overset{-}{\to} s_1$), commensalistic interactions (pairs of interactions of the form $s_{1/2} \overset{+}{\to} s_{2/1}$ and $s_{2/1} \overset{0}{\to} s_{1/2}$), and amensalistic interactions (pairs of interactions of the form $s_{1/2} \overset{-}{\to} s_{2/1}$ and $s_{2/1} \overset{0}{\to} s_{1/2}$). None of these other kinds of interactions can be described by the unsigned trophic network framework.

---

[3] At the population level this includes predation and parasitoidism, but in terms of temporally persistent symbioses of individual organisms, this only includes parasitism.

[4] As pointed out in [LPR19], (unsigned) directed networks can also describe other types of ecological interactions, not just the trophic interactions mentioned above. However, without using signed edge weights, only one type of ecological interaction can be described by a single network, whereas in real ecology all possible types of interaction can occur in the same ecosystem.

Previous work has looked at e.g. comparing "local" neighborhoods of nodes in (un-weighted)[5] signed networks [ZW17] or diffusion kernels in signed networks [QSyL+08]. How-ever, overall the existing literature on signed networks seems underdeveloped. This is surprising because such networks could presumably model an extremely wide range of phenomena, not just ecological interaction networks. For example, the[6] "gene interaction networks" defined implicitly in [MSH+08] could be considered (undirected) networks with mixed-sign edge weights if the edge weights were taken to be the differences ($\varepsilon = W_{xy} - E(W_{xy})$ in the notation of the paper) between the observed values of double mutant fitnesses and those expected under a null model of no interactions, with "positive" ($\varepsilon > 0$) edges corresponding to "alleviating" interactions ($W_{xy} > E(W_{xy})$), and "negative" ($\varepsilon < 0$) edges corresponding to "synergistic" interactions ($W_{xy} < E(W_{xy})$). More generally, any dynamical system with a multivariate state space that has incomplete connectivity could potentially be summarized using a network with mixed-sign edge weights. (Cf. section 2.1.1 for more on this idea.) In particular, it seems that no previous published work discusses what is called, in the framework of [TITP19], the known node correspondence comparison problem for networks with mixed-sign edge weights. That is surprising because the problem seems like an obvious question to ask about a class of objects with extremely broad potential applicability. If prior literature exists then most likely I was unable to find it due to it using different terminology.

8.2   Preliminaries

Section 8.2.1 clarifies the type of problem considered herein. Section 8.2.2 introduces the main issue with trying to compare signed networks the same way as unsigned networks. Section 8.2.3 gives some basic definitions, but cf. section 8.3 for more details. Sections 8.2.4, 8.2.5, and 8.2.6 provide criteria we might like comparison methods to satisfy in order to address these concerns. Section 8.2.7 provides an additional criterion which is also relevant for unsigned networks. Finally section 8.2.8 overviews at a very high level the kinds of comparisons of networks which we might want to make herein.

8.2.1   Known vs. Unknown Node Correspondence

There are at least two types of network comparison problem which we should distinguish. Following the framework from [TITP19], we can consider

- the known node correspondence problem ("fixed nodes, variable edges"), or

---

[5] In the framework of [TITP19] what [ZW17] addresses is the unknown node correspondence comparison problem. In contrast, this work addresses the "known node correspondence" comparison problem. Cf. section 8.2.1. Other differences are that this work considers "global" comparisons of entire networks and allows edge weights to have magnitudes other than 1.

[6] These correspond to "gene × gene" interaction networks in the terminology of item (4) of section 2.1.1, or equivalently to "unpartitioned interaction networks" from the terminology of the introduction to Part I.

- the unknown node correspondence problem ("variable everything").

For the known node correspondence problem, the identity of the nodes is essential and fixed. The two compared networks must either have the same set of labeled nodes, or a known correspondence between their node sets must be given. The known-node-correspondence problem asks whether the two networks describe similar relationships for the fixed set of nodes. An example of the known node correspondence problem would be to ask how much airline flight routes in Europe changed after the onset of the pandemic. This amounts to comparing two networks (one for before the pandemic and one for after) with the same labeled nodes (the airports), but for which the edges (the flight routes) between those nodes may be different. Cf. section 8.B.1, that attempts to make these ideas more precise.

For the unknown node correspondence problem, the identity of the nodes does not matter. The identity of the nodes may differ between the two networks. The unknown node correspondence asks essentially whether the relationships described by the two networks are "analogous". An example of the unknown node correspondence problem would be to ask whether the structure of airports and flight routes is analogous between Europe and China. The set of nodes is obviously different, and a priori there's no reason to believe that e.g. the Shanghai airport should be identified with the London airport rather than the Paris airport or vice versa. Nevertheless one could still sensibly ask e.g. whether it's possible to identify analogous subsets of highly connected "central hub" airports in both networks.

The network comparison problem described in section 8.1 clearly is known node correspondence. We want to quantitatively assess the ability of statistical estimators to recover the correct set of edges (interactions) for a given set of nodes (microbes). It makes no difference if the estimated network is "analogous" to the true network. If the ecological roles of the various microbes are misidentified by ascribing certain relationships to the incorrect pairs of microbes, the estimated network is still a terrible estimate of the truth. Only the known node correspondence problem is relevant for comparing the estimates produced by an estimator with the true value of an estimand.

Unfortunately, the fact that our network comparison problem is known node correspondence also means that most of the methods discussed in [TITP19] would be inapplicable and irrelevant even if we were considering unsigned networks. While comparing distinct microbial ecosystems for analogous network structures is a valid scientific problem and would also require extension of methods for the unknown node correspondence problem to signed networks, it is outside of scope here and left to future work. Herein I try to modify and extend (as necessary) most of the computationally tractable methods[7] for the known node correspondence problem discussed in [TITP19] to apply also to signed networks.

8.2.2 Positive and Negative Edges Qualitatively Distinct

In microbial ecology, positive interactions and negative interactions correspond to categorically and qualitatively distinct phenomena. For example, if all growth-promoting (positive)

---

[7]I.e. not the cut distance, which is not really computationally tractable.

interactions were changed to growth-suppressing (negative) interactions, ecologically speaking the result would be quite different. This is true even if, or in some cases even especially if, the magnitudes of the interactions were left unchanged. Translated to a network representation, this implies that the positive and negative edges of networks with mixed-sign edge weights should be treated as qualitatively distinct.

   Standard network comparison methods assume all edge weights have the same (positive) sign. Therefore applying or extending such methods to networks with mixed-sign edge weights requires caution. We need to avoid situations where, for example, positive and negative weights "cancel" in a scientifically meaningless way, or two compared edges with different signs are not penalized due to having similar magnitudes. Section 8.2.5 addresses this caution by providing a principle to "sanity check" whether a network comparison method may potentially give meaningful results for networks with mixed-sign edge weights.

8.2.3   Definition of Networks with Mixed-Sign Edge Weights

These definitions are discussed in more detail in section 8.3.

   A "**network with mixed-sign edge weights**" $\mathcal{G}$ with $S \in \mathbb{N}$ nodes is

(i) a "**node set**" $\mathcal{N}_\mathcal{G}$ (which we can always assume equals $[S] := \{1, \ldots, S\}$ without loss of generality, cf. the discussion below),

(ii) an "**edge set**" $\mathcal{E}_\mathcal{G} \subseteq \mathcal{N}_\mathcal{G} \times \mathcal{N}_\mathcal{G}$ such that $(s_1, s_2) \in \mathcal{E}_\mathcal{G}$ if and only if there is an edge directed from node s_1 to node s_2, and

(iii) an "**edge weight function**" $A_\mathcal{G} : \mathcal{N}_\mathcal{G} \times \mathcal{N}_\mathcal{G} \to \mathbb{R}$ such that

$$A_\mathcal{G}(s_1, s_2) = \begin{cases} w \in (-\infty, 0) \cup (0, \infty) & (s_1, s_2) \in \mathcal{E}(\mathcal{G}) \\ 0 & (s_1, s_2) \notin \mathcal{E}(\mathcal{G}) \end{cases} . \tag{8.2.1}$$

For each edge $(s_1, s_2) \in \mathcal{E}_\mathcal{G}$, the value $A_\mathcal{G}(s_1, s_2)$ is the "**weight**" of the edge, the value $\mathrm{sign}(A_\mathcal{G}(s_1, s_2)) = \pm 1$ is the "**sign**" of the edge, and the value $|A_\mathcal{G}(s_1, s_2)|$ is the "**magnitude**" of the edge. Given an explicit identification between $\mathcal{N}_\mathcal{G}$ and $[S]$ (cf. below), there always exists a unique $S \times S$ matrix $\mathbf{A}_\mathcal{G}$, which is called the "**adjacency matrix**" of the network $\mathcal{G}$, that corresponds to the edge weight function $A_\mathcal{G}$.

   For every pair $s_1, s_2 \in \mathcal{N}_\mathcal{G}$ such that $s_1 \neq s_2$, the definition allows for at most one edge directed from s_1 to s_2, and for at most one edge directed from s_2 to s_1. Likewise, for every $s \in \mathcal{N}_\mathcal{G}$, the definition allows for at most one edge ("self-loop") directed from s to s. I.e. "multigraphs" are not considered.

8.2.4   Monotonicity Principle

Because literature already exists on comparing unsigned networks, a practical paradigm for comparing signed networks focuses on how to extend comparison methods for unsigned

networks to also apply to signed networks. Herein we will focus only on extending comparison methods for unsigned networks that satisfy a mild "well-behavedness" criterion, termed the "monotonicity principle". Cf. figure 8.1.
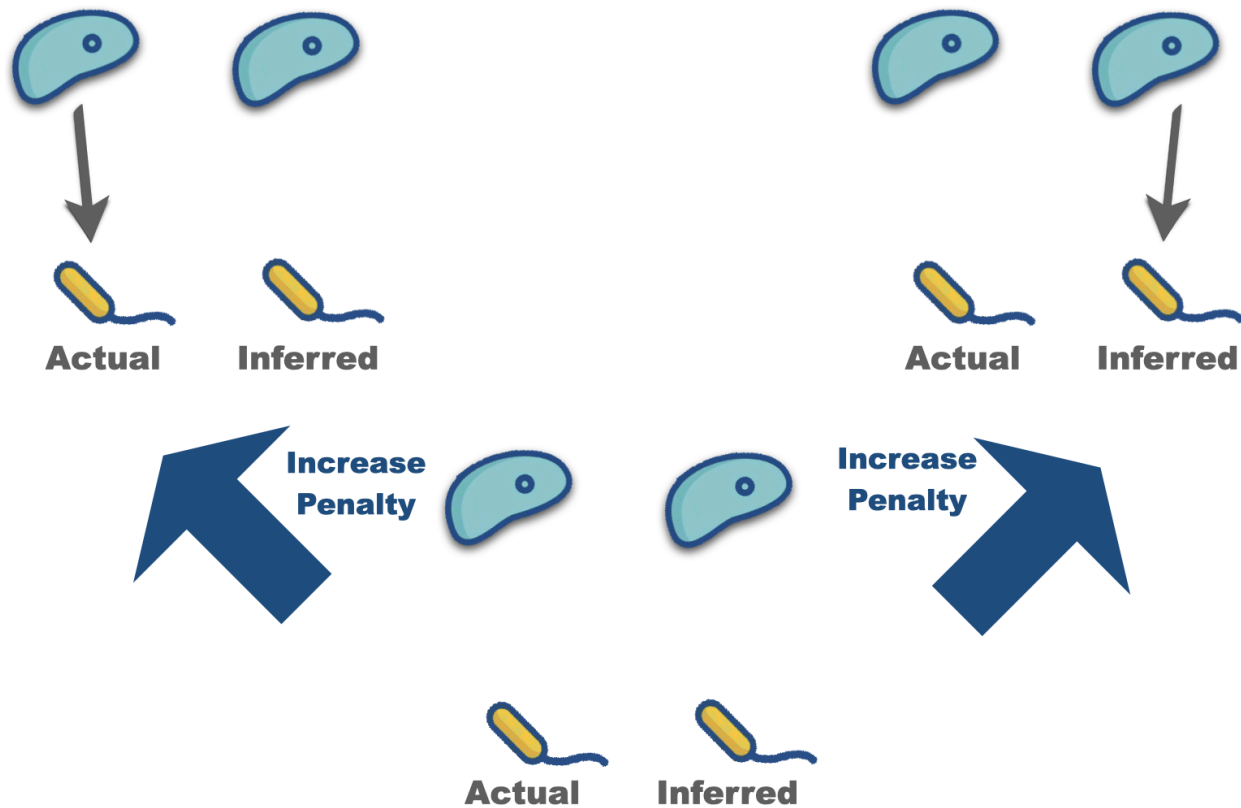
Figure 8.1: Schematic illustration of the monotonicity principle for well-behaved comparison functions of unsigned networks.

At a high level, the idea of the monotonicity principle is that "reasonable" comparison methods for unsigned networks should penalize existence errors no less than they penalize "true misses", i.e. situations where an edge is missing in both networks. Cf. section 8.B.8 for an attempted precise definition. One can also ask for a "continuous" version of the monotonicity principle whereby the size of the penalty for an existence error must never decrease as the magnitude of the weight of the unmatched edge increases. Cf. section 8.B.9.

Because the monotonicity principle is a very intuitive property that one would most often tacitly assume to be true by default, we can consider comparison methods for unsigned networks that satisfy the monotonicity principle to be "reasonable". Therefore, to implement our paradigm of extending such comparison methods to also apply to signed networks, we need to be able to check whether any proposed extension satisfies the analogous intuitive property for signed networks. This of course first requires us to identify what the analogous intuitive property for signed networks is, which we do below.

8.2.5   The Double Penalization Principle

Confusing two (non-zero) edges with different signs is the only new kind of error that can occur when comparing signed networks that possesses no analogue when comparing unsigned networks. All other kinds of errors have analogues in the comparison of unsigned networks and hence can be treated using a "reasonable" comparison method for unsigned networks. Thus our main challenge, when trying to identify "reasonable" ways to extend comparison methods intended for unsigned networks to also apply to signed networks, is ensuring that errors confusing edges with different signs are treated in a "reasonable" way.

When interpreting positive and negative edge weights as corresponding to distinct ("equal but opposite") phenomena, a mistake confusing edges with different signs amounts to the composite of two separate mistakes: both (i) inferring a phenomenon which does not exist, and (ii) failing to infer a phenomenon which does exist. Being worse than either of those two separate mistakes when considered individually, such a composite mistake should therefore receive a penalty that is no smaller than either of the penalties given to the two separate mistakes. This observation leads directly to the double penalization principle.

Double penalization principle: *When the sign of an edge differs between two networks, the resulting penalty should be larger than the maximum of the two penalties that would occur if the edge was missing in either network.*

In other words, the penalty should equal the maximum penalty that could occur if the edge was missing from either graph, plus an additional (second) penalty. Both mistakes, of inferring the wrong sign and of failing to infer the correct sign, should be penalized. The "composite" mistake should therefore be "doubly penalized". Cf. figure 8.2. The double penalization principle is equivalent to requiring that the monotonicity principle is simultaneously satisfied for both the positive parts and for the negative parts of the signed networks being compared. (Cf. section 8.3.4 for definitions of "positive part" and "negative part" of a signed network. Both are unsigned networks, with the positive part having weights equal to the magnitudes of edges with positive sign, and analogously for the negative part.)
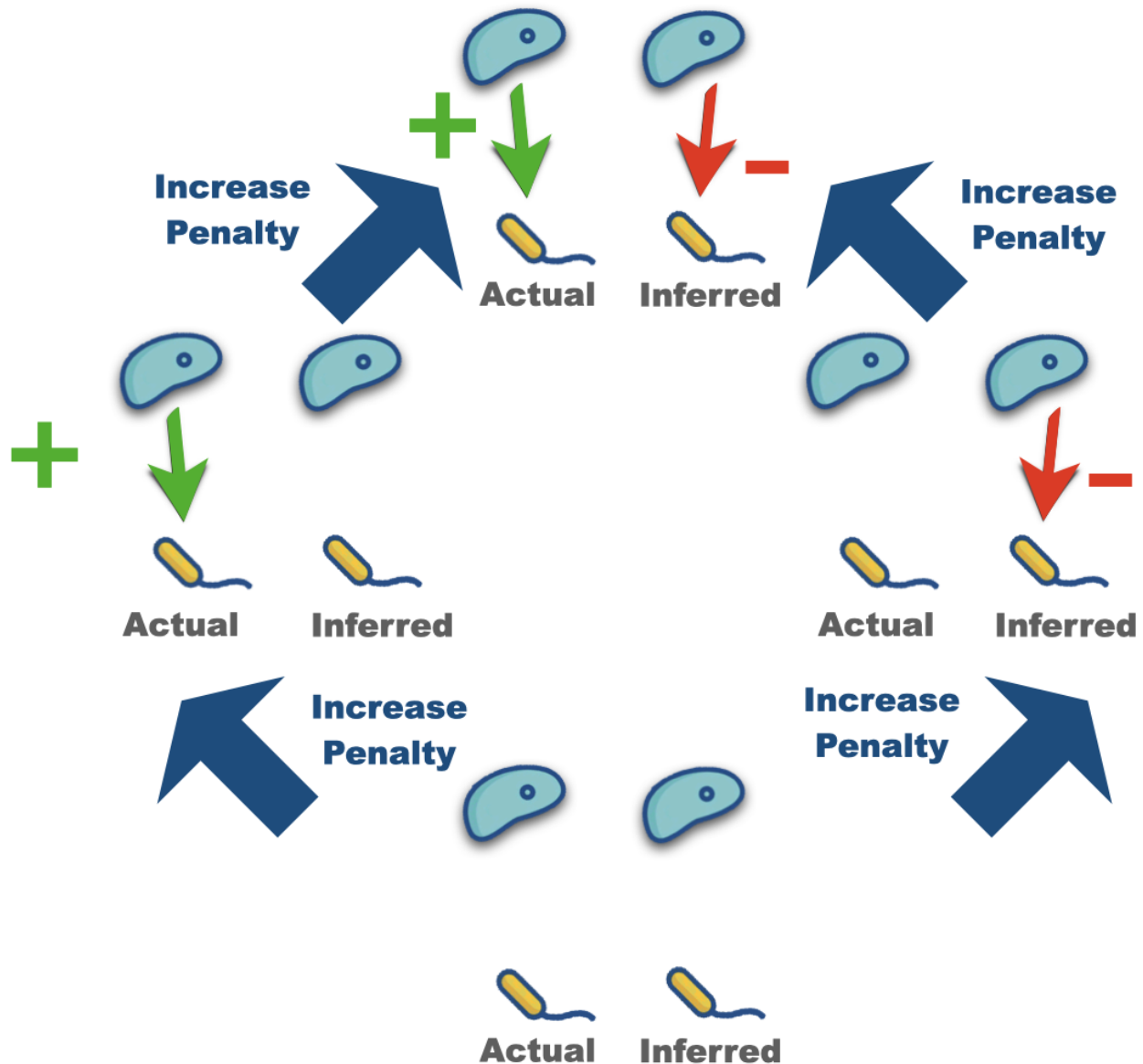
Figure 8.2: Double Penalization Principle

Note that the double penalization principle excludes naive comparison methods for signed networks that "project" the signed networks being compared onto a space of unsigned networks and then apply a comparison method for unsigned network to the "projections". This means that the most straightforward idea for extending comparison methods for unsigned networks to also apply to signed networks fails to create comparison methods for signed networks that either are "well-behaved" or "behave intuitively". This is true even when the original comparison method for unsigned networks is "well-behaved" and "behaves intuitively", in the sense that it satisfies the monotonicity principle.

8.2.6   Convex Combination Decomposition Property

A sufficient, *but not necessary*, condition for satisfying the double penalization principle is

- When the sign of an edge differs between two graphs, the resulting penalty equals the sum of the two penalties that would occur if the edge was set to zero ("removed") in either graph.

An equivalent formulation of the double penalization principle is that the penalty for the edge should equal the maximum of the two penalties that occur when considering either the networks' positive parts only or negative parts only (see section 8.3.4), plus an additional penalty. This leads to the following equivalent formulation of the above sufficient, *but not necessary*, criterion:

- When the sign of an edge differs between two graphs, the resulting penalty equals the sum of the penalty from comparing the positive parts of the graphs with the penalty from comparing the negative parts of the graphs.

When the penalties are additive across edges, the above criterion means the total penalty is the (weighted) sum of the penalties for the positive parts and negative parts of the graphs. Such (dis)similarity measures can be modified by suitable normalizations to produce (dis)similarity measures whose values equal a convex combination of (i) their value when applied to the positive parts of the networks only and (ii) their value when applied to the negative parts of the networks only. The resulting (dis)similarity measures are said to have the **convex combination decomposition property**. It follows from the above that the convex combination decomposition property is sufficient, *but not necessary*, to satisfy the double penalization principle.

   The convex combination decomposition property is potentially desirable not only because it guarantees that the double penalization principle will be satisfied. (Dis)similarity measures with this property are easy to interpret because understanding how the networks' positive edges vs. their negative edges contribute to the final value is straightforward.

8.2.7   Sparsity-Savviness Principle

This principle is applicable to general networks, not just signed networks. The fact that an edge is absent from two networks should not lead to an increase in their similarity score. Sparsity should not lead to artifactual similarity. In other words, "don't reward true misses". Otherwise, two networks which are very sparse (have a small percentage of all possible edges that *could* exist) could potentially be scored as very similar, even when the structure of the edges they *do* have is very different.

8.2.8   Aspects of Networks to Compare

There are many aspects of networks we might seek to capture in our comparisons.

"Numerical" comparisons describe the correspondence of the numerical values of the edge weights. This could refer to only their magnitudes, or to their signed values. Relative error using a given norm (e.g. entrywise L_1) is an example. Cf. section 8.4.1 for details.

There are also many ways we might seek to compare networks more "qualitatively".

We might seek comparisons that describe the relative ordering of the edge weights (which are largest, which are smallest), either in magnitude, signed value, or both. (Sparsity-adjusted) Spearman correlation is an example. Cf. section 8.4.2 for details.

We might also seek "qualitative" comparisons that describe features of network topology. For example, (unweighted) Jaccard similarity describes the presence/absence of edges. Cf. section 8.4.3 for details. Similarly, (unweighted) DeltaCon distance [KSV$^+$16] describes the correspondences of paths. Cf. section 8.4.4 for details. One might also want to define comparisons to describe the correspondences of other subgraph motifs besides paths.

We can also seek to capture both "numerical" and "qualitative" aspects of networks in a single comparison. Usually this is done by considering the correspondence of given qualitative features more or less important depending upon the weights of the edges in those features. Examples include weighted Jaccard similarity and weighted DeltaCon distance.

8.3 General Definitions

This section gives precise definitions which will be used throughout the rest of the chapter for describing or modifying networks. Section 8.3.1 gives definitions and notation which will be used interchangeably to specify network structure. Section 8.3.2 defines a related matrix which is important for section 8.4.4 later. Section 8.3.3 defines standard ways of "forgetting" network structure. Section 8.3.4 defines a standard way of decomposing a network, extending well-known decompositions applicable to the functions and matrices from section 8.3.1 which can be used to represent the network. Finally section 8.3.5 defines a standard way to join two networks together to create a new (disconnected) network.

8.3.1 Representations of Networks as Functions, Sets, and Matrices

The (i,j)th entry of any matrix $\mathbf{M}$ will be denoted $[\mathbf{M}]_{ij}$. S denotes the number of strains.

Any function $\epsilon : [S] \times [S] \to \mathbb{R}$ is equivalent to specifying a (mixed-sign) weighted and directed graph $\mathcal{G}$ with S nodes. The values of ϵ are the edges weights. Given any pair (s_1, s_2), $\epsilon(s_1, s_2)$ gives the weight of the edge directed from s_1 to s_2, or equals 0 if no such edge exists. Cf. figure 8.3. Any such **edge function** $\epsilon : [S] \times [S] \to \mathbb{R}$ can also be identified with an $S \times S$ **adjacency matrix** $\mathbf{A}$, where entry $[\mathbf{A}]_{s_1 s_2}$ of the matrix equals $\epsilon(s_1, s_2)$. Thus the (s_1, s_2)'th entry $[\mathbf{A}]_{s_1 s_2}$ of the adjacency matrix $\mathbf{A}$ gives the weight of the edge from s_1 to s_2, or is equal to 0 if no such edge exists.

Note that here we are only explicitly considering networks that are "static" in time, but if we allow the adjacency matrix to vary with time, then the above also applies to time-varying networks (for modelling e.g. time-varying interactions [HFH21]), cf. [AML$^+$17].

The set of edges $\mathcal{E}$ of the graph $\mathcal{G}$ is then defined by the set

$$\mathcal{E} := \mathrm{supp}(\epsilon) := \{(s_1, s_2) \in [S] \times [S] : \epsilon(s_1, s_2) \neq 0\} . \tag{8.3.1}$$
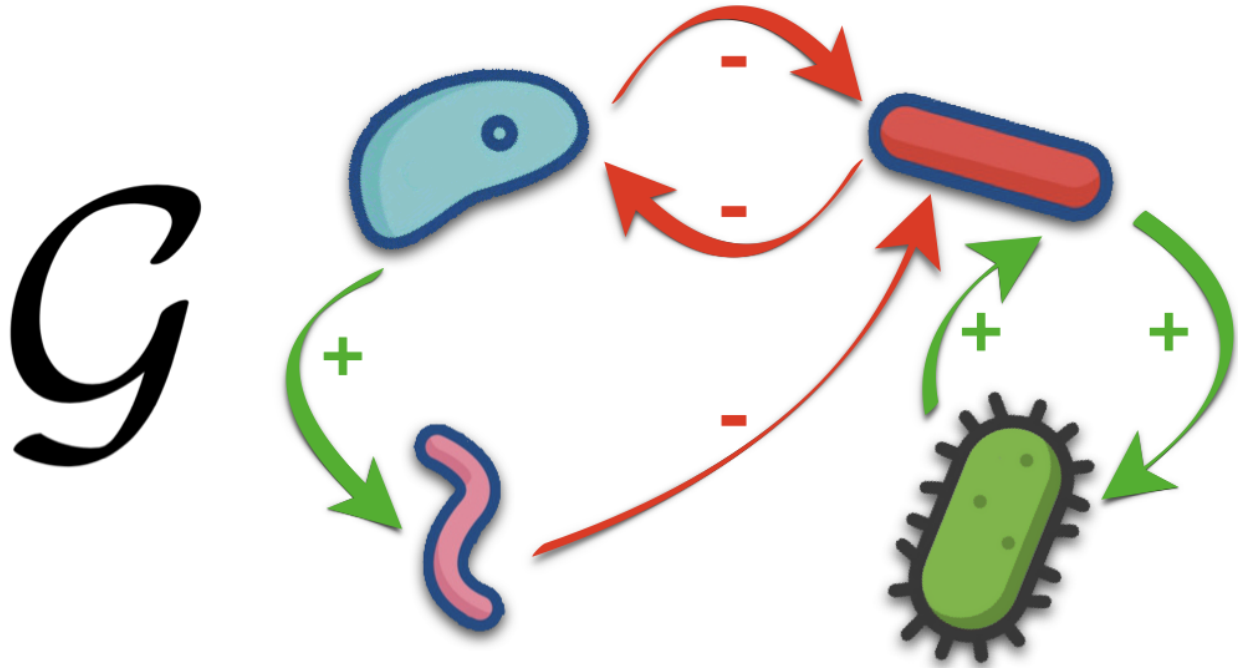
Figure 8.3: Directed network with signed edges. Bacteria are the nodes and identified with the set $[S] = [4]$. The arrows are the edges, identified with $\mathrm{supp}(\epsilon)$ where $\epsilon : [4] \times [4] \to \mathbb{R}$. Green arrows are positively weighted edges, identified with the set $\{(s_1, s_2) \in [4] \times [4] : \epsilon(s_1, s_2) > 0\}$. Red arrows are negatively weighted edges, identified with the set $\{(s_1, s_2) \in [4] \times [4] : \epsilon(s_1, s_2) < 0\}$.

8.3.2 The Degree Matrix

Note: the material in this section is only used later on in section 8.4.4.

The degree matrix $\mathbf{D}$ of a network $\mathcal{G}$ with adjacency matrix $\mathbf{A}$ is defined as

$$[\mathbf{D}]_{s_1 s_2} := \begin{cases} \sum_{\sigma_2=1}^{S} \mathbf{A}_{s_1 \sigma_2} & s_1 = s_2 \\ 0 & \text{else} \end{cases} . \tag{8.3.2}$$

Intuitively speaking, the s'th entry on the diagonal of $\mathbf{D}$ encodes the "net influence" flowing out from node s. In the case of an unsigned and unweighted network, i.e. an adjacency

matrix consisting entirely of 1's and 0's, this is quantified by the number of the number of edges originating from node s.

8.3.3 Skeletons

Unsigned skeletons (section 8.3.3.1) preserve the least amount of structure of the original network. Signed skeletons (8.3.3.2) and magnitude skeletons (8.3.3.3) preserve more structure. By definition all notions of "graph skeleton" lead to networks with edge sets $\mathcal{E}$ that are the same as that of the original network.

8.3.3.1 (Unsigned) Skeletons

Given a network $\mathcal{G}$ with adjacency matrix $\mathbf{A}$, the **unsigned skeleton** $\mathcal{S}(\mathcal{G})$ is defined to be the network with adjacency matrix $\mathbf{B}$ such that

$$[\mathbf{B}]_{ij} := \begin{cases} 1 & [\mathbf{A}]_{ij} \neq 0 \,, \\ 0 & [\mathbf{A}]_{ij} = 0 \,. \end{cases} \qquad (8.3.3)$$

The unsigned skeleton $\mathcal{S}(\mathcal{G})$ preserves only the underlying connectivity/topology of $\mathcal{G}$, while "deleting" all information about edge weights or signs.

8.3.3.2 Signed Skeletons

Given a network $\mathcal{G}$ with adjacency matrix $\mathbf{A}$, the **signed skeleton** $\mathcal{S}^{\pm}(\mathcal{G})$ is defined to be the network with adjacency matrix $\mathbf{B}$ such that

$$[\mathbf{B}]_{ij} := \operatorname{sign}([\mathbf{A}]_{ij}) = \begin{cases} 1 & [\mathbf{A}]_{ij} > 0 \,, \\ -1 & [\mathbf{A}]_{ij} < 0 \,, \\ 0 & [\mathbf{A}]_{ij} = 0 \,. \end{cases} \qquad (8.3.4)$$

In addition to the underlying connectivity of the original network $\mathcal{G}$, the signed skeleton $\mathcal{S}^{\pm}(\mathcal{G})$ also preserves the signs of the edge weights. The signed skeleton $\mathcal{S}^{\pm}(\mathcal{G})$ still "deletes" all information about edge magnitudes.

Note that for a network with non-negative edge weights, the unsigned and signed skeletons coincide. Thus the signed skeleton is only a distinct and new notion for the more general setting of networks with mixed-sign edge weights.

8.3.3.3 Magnitude Skeletons

Given a network $\mathcal{G}$ with adjacency matrix $\mathbf{A}$, the **magnitude skeleton** $\mathcal{M}(\mathcal{G})$ is defined to be the network with adjacency matrix $\mathbf{B}$ such that

$$[\mathbf{B}]_{ij} := \operatorname{abs}([\mathbf{A}]_{ij}) := |[\mathbf{A}]_{ij}| = \begin{cases} [\mathbf{A}]_{ij} & [\mathbf{A}]_{ij} > 0 \,, \\ -[\mathbf{A}]_{ij} & [\mathbf{A}]_{ij} < 0 \,, \\ 0 & [\mathbf{A}]_{ij} = 0 \,. \end{cases} \qquad (8.3.5)$$

In addition to the underlying connectivity of the original network $\mathcal{G}$, the magnitude skeleton $\mathcal{M}(\mathcal{G})$ also preserves the magnitudes of the edge weights. The magnitude skeleton $\mathcal{M}(\mathcal{G})$ still "deletes" all information about edge signs.

Note that for a network with non-negative edge weights, the magnitude skeleton coincides with the original network. Thus the magnitude skeleton is only a distinct and new notion for the more general setting of networks with mixed-sign edge weights. Relatedly, observe how $\mathcal{S}(\mathcal{G}) = \mathcal{M}(\mathcal{S}^{\pm}(\mathcal{G}))$.

8.3.4 Positive and Negative Parts

Any adjacency matrix $\mathbf{A}$ admits the decomposition:

$$\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-\,, \quad [\mathbf{A}^+]_{s_1 s_2} := \max\{A_{s_1 s_2}, 0\}\,, \quad [\mathbf{A}^-]_{s_1 s_2} := \max\{-A_{s_1 s_2}, 0\}\,,$$

which is equivalent to decomposing the edge function ϵ in the standard way:

$$\epsilon = \epsilon^+ - \epsilon^-\,, \quad \epsilon^+ := \max\{\epsilon, \mathbf{0}\}\,, \quad \epsilon^- := \max\{-\epsilon, \mathbf{0}\}\,.$$

$\mathbf{A}^+/\epsilon^+$ corresponds to its own network, denoted $\mathcal{G}^+$, the **positive subnetwork**. Likewise, $\mathbf{A}^-/\epsilon^-$ also corresponds to its own network, denoted $\mathcal{G}^-$, the **negative subnetwork**. Cf. figures 8.4 and 8.5.

The edge weights of both $\mathcal{G}^+$ and $\mathcal{G}^-$ are by definition all positive, so standard dissimilarity measures for graphs can be applied to each subnetwork separately. This facilitates analysis and corresponds to the point of view that positive and negative edge weights correspond to qualitatively distinct phenomena.
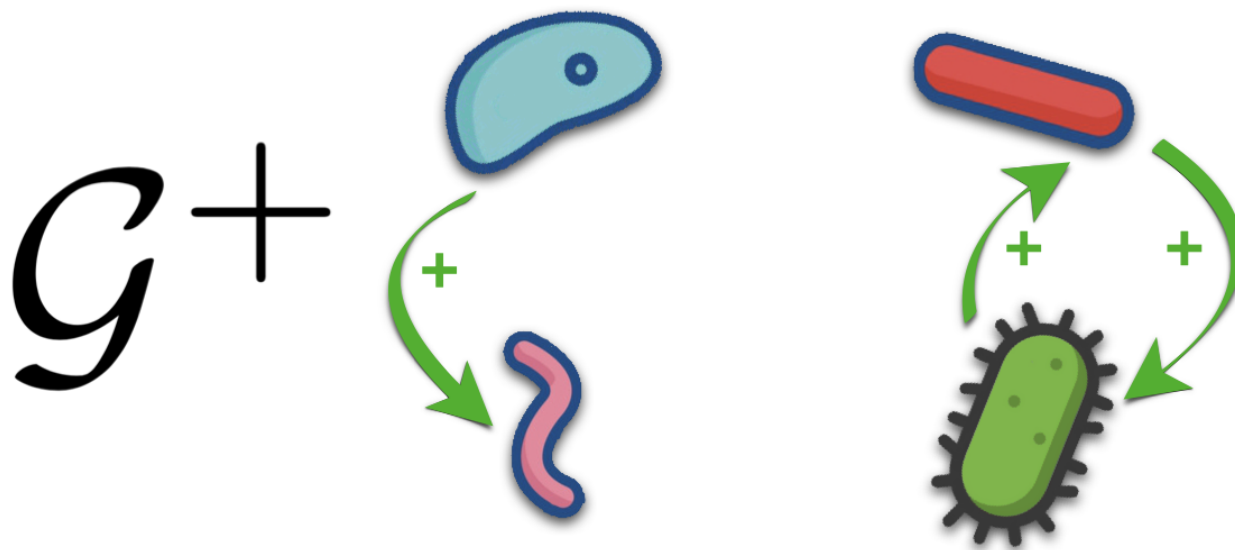
Figure 8.4: Positive Subnetwork of the Network from Figure 8.3.
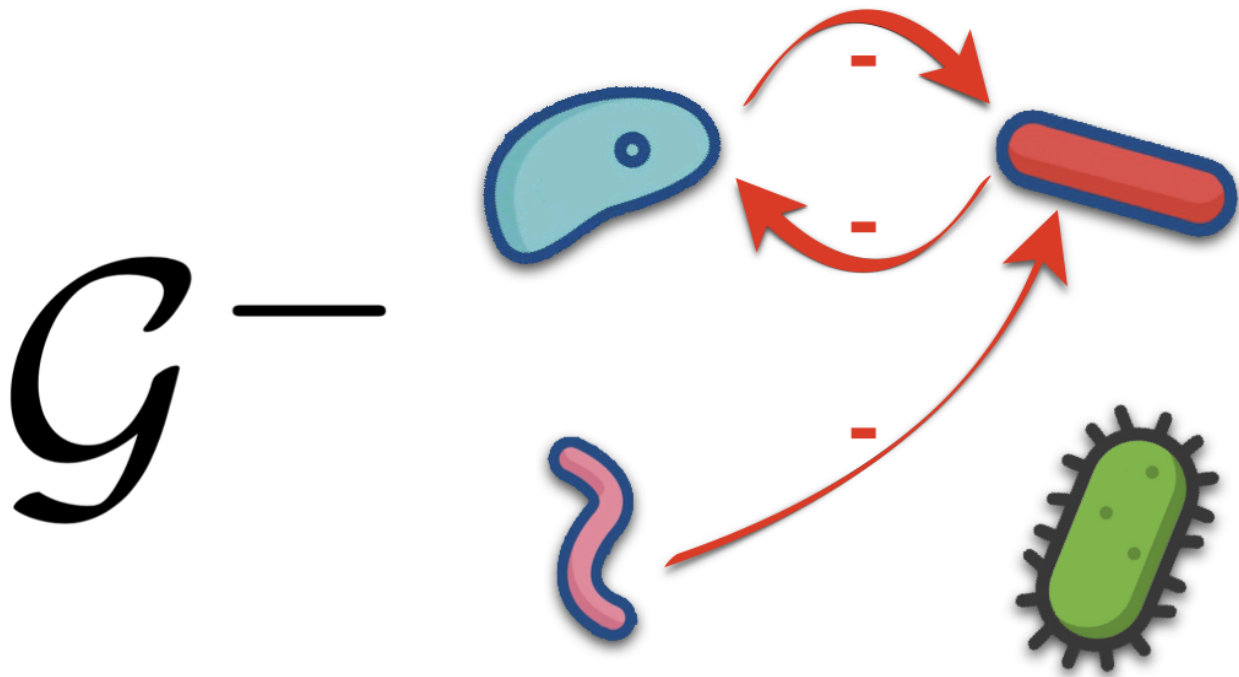
Figure 8.5: Negative Subnetwork of the Network from Figure 8.3.

8.3.5 Disjoint Union

Given two graphs, $\mathcal{G}_1$ with adjacency matrix $\mathbf{A}_1$ and $\mathcal{G}_2$ with adjacency matrix $\mathbf{A}_2$, define their "**disjoint union**" to be the graph $\mathcal{G}_1 \oplus \mathcal{G}_2$ corresponding to the adjacency matrix:

$$\mathbf{A}_1 \oplus \mathbf{A_2} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix},$$

i.e. the so-called direct sum of the matrices $\mathbf{A}_1$ and $\mathbf{A}_2$. The node set of this graph $\mathcal{G}_1 \oplus \mathcal{G}_2$ is the disjoint union of the node sets of $\mathcal{G}_1$ and $\mathcal{G}_2$, and likewise the edge set is the disjoint union of their edge sets $\mathcal{E}_1 \sqcup \mathcal{E}_2$. Cf. figure 8.6.
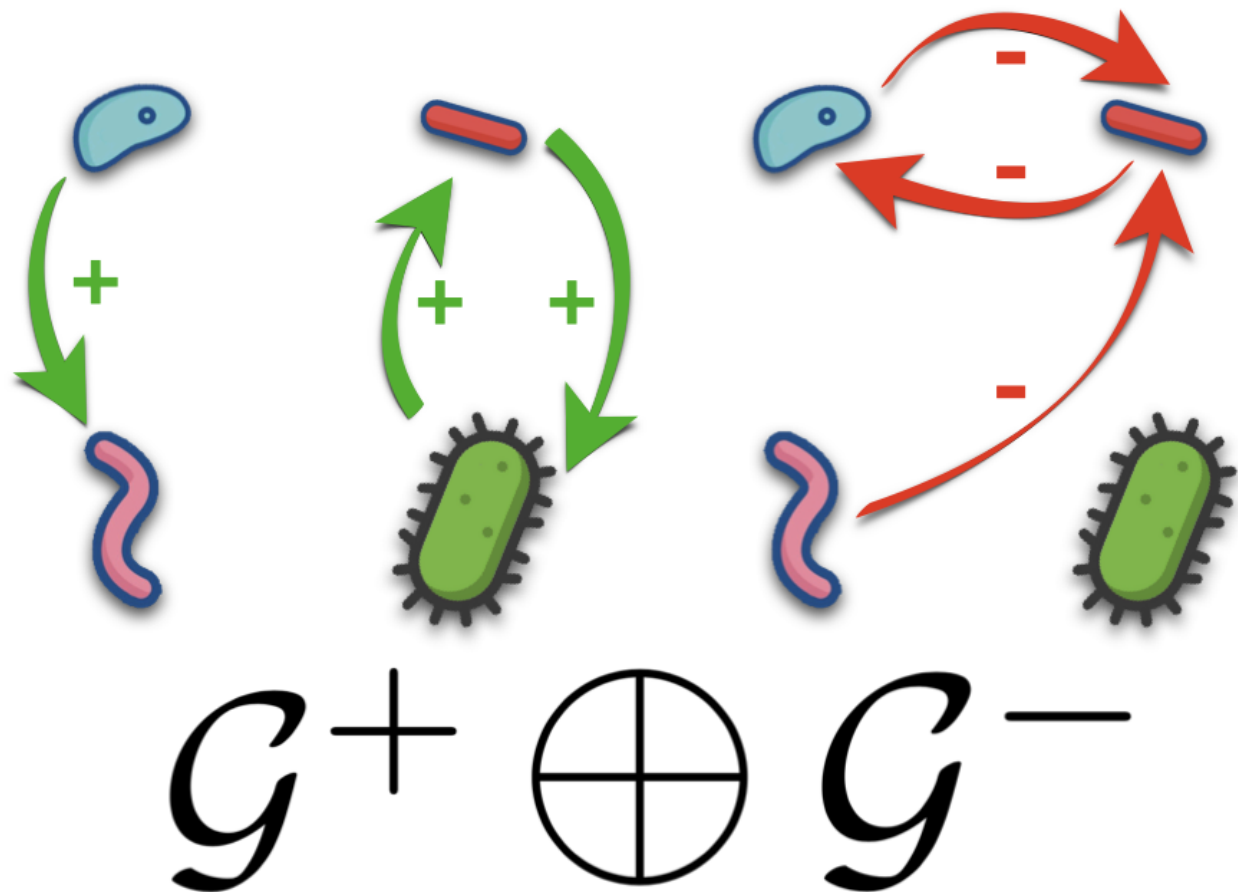
Figure 8.6: Direct Sum of the Positive and Negative Subnetworks of the Network from Figure 8.3. Notice how the nodes are duplicated compared to Figure 8.3.

8.4   (Dis)similarity Measures

Below I define and discuss comparison methods applicable to networks with mixed-sign edge weights which span the full range of qualitative and quantitative comparisons explained in section 8.2.8. When relevant, I explain how to modify methods for comparing unsigned networks to create new methods that give sensible results for networks with mixed-sign edge weights. I also explain the extent to which these proposed methods fall in line with the criteria discussed previously in sections 8.2.5, 8.2.6, and 8.2.7.

Section 8.4.1 provides mostly "quantitative" comparisons of networks, section 8.4.2 provides mostly "qualitative" comparisons of relative orderings of edge weights, section 8.4.3 provides "qualitative" and "quantitative" comparisons of edge presence/absence, and finally section 8.4.4 provides "qualitative" and "quantitative" comparisons of path correspondence.

8.4.1 (Entrywise L_1) Relative Error

Relative error is a "quantitative" comparison of networks. If we normalize the edge weights of both networks before computing their relative error, the comparison becomes somewhat more "qualitative" in the sense that the input edge-weight-normalized networks better capture the relative sizes of the edge weights. Herein I show that "out of the box" the entrywise L_1 relative error as defined below automatically considers both positive and negative edges separately and then sensibly recombines the results.

8.4.1.1 Entrywise L_1 Norm

The entrywise L_1 norm of an $S \times S$ matrix $\mathbf{A}$ will herein be defined as

$$\|\mathbf{A}\|_{[1]} := \sum_{\sigma_1=1}^{S} \sum_{\sigma_2=1}^{S} \left| [\mathbf{A}]_{\sigma_1 \sigma_2} \right| . \tag{8.4.1}$$

In other words, this is the sum of the absolute values of the entries of $\mathbf{A}$, or the L_1 norm of a vector resulting from vectorizing $\mathbf{A}$. It is the "L_1 analogue" of the Frobenius norm.

It is *not* to be confused with L_1 *operator* norm of $\mathbf{A}$, which is typically denoted $\|\mathbf{A}\|_1$. The notation $\|\mathbf{A}\|_{[1]}$ serves not only to emphasize this distinction, but also to remind of the notation for the entries of a matrix.

An operator norm would be inappropriate to use here, given that the adjacency matrix $\mathbf{A}$ here is not meant to represent a linear function[8] (it is "entirely contravariant"). The Frobenius norm would also be inappropriate to use here. The Frobenius norm incorrectly "inflates" large-magnitude entries and "deflates" small-magnitude entries.

8.4.1.2 Relative Error Definition

Given a "truth" matrix $\mathbf{A}_*$ and an "estimate" matrix $\hat{\mathbf{A}}$, the error of $\hat{\mathbf{A}}$ relative to $\mathbf{A}_*$, or just "relative error" if clear from context, is

$$\frac{\left\| \hat{\mathbf{A}} - \mathbf{A}_* \right\|_{[1]}}{\|\mathbf{A}_*\|_{[1]}} . \tag{8.4.2}$$

The fact that relative error normalizes by the norm of one of the networks arguably helps to avoid considering two networks highly similar merely because they are both sparse. Cf. section 8.2.7. Normalization could help to ensure that a small number of differences is considered important if, due to sparsity, there are only a small number of edges that might be different to begin with.

---

[8]Arguably except during the formation of the Neumann series in the definition of DeltaCon, cf. section 8.4.4.1.

8.4.1.3 Convex Combination Decomposition Property

One can directly verify the identity:

$$\left\|\hat{\mathbf{A}} - \mathbf{A}_*\right\|_{[1]} = \left\|\hat{\mathbf{A}}^+ - \mathbf{A}_*^+\right\|_{[1]} + \left\|\hat{\mathbf{A}}^- - \mathbf{A}_*^-\right\|_{[1]}, \tag{8.4.3}$$

which also happens to be a special case (for a counting measure) of the measure-theoretic identity $\int |f - g|\mathrm{d}\mu = \int |f^+ - g^+|\mathrm{d}\mu + \int |f^- - g^-|\mathrm{d}\mu$.

Therefore the overall relative error can be written as a convex combination/weighted average of the "positive relative error" and the "negative relative error". More specifically, the following is true:

$$\frac{\left\|\hat{\mathbf{A}} - \mathbf{A}_*\right\|_{[1]}}{\|\mathbf{A}_*\|_{[1]}} = c_1 \left(\frac{\left\|\hat{\mathbf{A}}^+ - \mathbf{A}_*^+\right\|_{[1]}}{\left\|\mathbf{A}_*^+\right\|_{[1]}}\right) + c_2 \left(\frac{\left\|\hat{\mathbf{A}}^- - \mathbf{A}_*^-\right\|_{[1]}}{\left\|\mathbf{A}_*^-\right\|_{[1]}}\right), \tag{8.4.4}$$

where the coefficients above are defined

$$c_1 := \frac{\left\|\mathbf{A}_*^+\right\|_{[1]}}{\|\mathbf{A}_*\|_{[1]}}, \quad c_2 := \frac{\left\|\mathbf{A}_*^-\right\|_{[1]}}{\|\mathbf{A}_*\|_{[1]}}. \tag{8.4.5}$$

This also ensures that any positive interactions estimated as negative, or vice versa, will appear twice in the above expression. We can see how the convex combination property implies the double penalization principle is satisfied.

8.4.1.4 Normalizing to get More Qualitative Comparisons

If a priori we don't believe that the edge weights of the two networks are on the same "length scale", e.g. they are measured in different units, but we still want to compare them in some qualitative, "scale-free" way, then we can normalize both networks' adjacency matrices before computing their relative error:

$$\frac{\left\|\dfrac{\hat{\mathbf{A}}}{\left\|\hat{\mathbf{A}}\right\|_{[1]}} - \dfrac{\mathbf{A}_*}{\|\mathbf{A}_*\|_{[1]}}\right\|_{[1]}}{\left\|\dfrac{\mathbf{A}_*}{\|\mathbf{A}_*\|_{[1]}}\right\|_{[1]}} = \frac{\left\|\dfrac{\hat{\mathbf{A}}}{\left\|\hat{\mathbf{A}}\right\|_{[1]}} - \dfrac{\mathbf{A}_*}{\|\mathbf{A}_*\|_{[1]}}\right\|_{[1]}}{1} = \left\|\frac{\hat{\mathbf{A}}}{\left\|\hat{\mathbf{A}}\right\|_{[1]}} - \frac{\mathbf{A}_*}{\|\mathbf{A}_*\|_{[1]}}\right\|_{[1]}. \tag{8.4.6}$$

If we wanted to compare whether the relatively largest magnitude edge weights in both networks tended to have the same sign, or belong to edges connecting the same nodes, this could be useful. If an edge has the same sign, but three times the magnitude, in $\mathcal{G}_*$ compared to $\hat{\mathcal{G}}$, for example, that won't be penalized if the sum of the magnitudes of the edge weights is also three times as large in $\mathcal{G}_*$ compared to $\hat{\mathcal{G}}$ (hence why the comparison is "scale-free").

Note that the resulting values are necessarily always in the interval $[0, 2]$. Also, unlike the relative error in general (8.4.2), the resulting values are symmetric when interchanging $\hat{\mathbf{A}}$ and $\mathbf{A}_*$ in the formula (8.4.6).

Normalizing both matrices first before computing their relative error arguably also helps to avoid considering two networks highly similar merely because they are both sparse. Cf. again section 8.2.7.

8.4.2 Mixed-Sign Spearman Correlation

These methods based on Spearman correlation are designed to provide a qualitative comparison of two networks, by comparing the relative ordering of their edge weights. The formation of rank vectors destroys any direct quantitative information about exact edge weight values. The mixed-sign Spearman correlation defined below provides an advantage over "raw" Spearman correlation by separately comparing the relative ordering of positive and negative edges and then sensibly recombining the results. By considering this extra information contained in the signs of the edge weights, it is not fooled by drastic changes to the network structure to which the "raw" Spearman correlation or Spearman correlation of the weights' magnitudes are completely oblivious.

8.4.2.1 Sparsity Adjustment

Given two networks $\mathcal{G}_1$, $\mathcal{G}_2$ with adjacency matrices $\mathbf{A}_1$, $\mathbf{A}_2$ respectively, consider the subset of index pairs for which the edge corresponding to each index pair is present in at least one of the two networks:

$$\mathcal{I}_{\mathcal{G}_1,\mathcal{G}_2} := \{(s_1, s_2) \in [S] \times [S] : [\mathbf{A}_1]_{s_1 s_2} \neq 0 \text{ or } [\mathbf{A}_2]_{s_1 s_2} \neq 0\} . \tag{8.4.7}$$

Then define a modified vectorization operator $\text{vect}_{\mathcal{G}_1,\mathcal{G}_2}$ which, given $[S] \times [S]$ matrices, returns vectors of length $|\mathcal{I}_{\mathcal{G}_1,\mathcal{G}_2}|$ whose indices correspond to the index pairs belonging to $\mathcal{I}_{\mathcal{G}_1,\mathcal{G}_2}$. This is a "sparsity adjustment" because it removes values corresponding to index pairs for which no edge exists in either $\mathcal{G}_1$ or $\mathcal{G}_2$.

8.4.2.2 Rank Vectors

Let $\mathcal{R}$ denote the "rank operator" which replaces each value of a vector with its relative ordering (e.g. the smallest element is assigned 1, the largest is assigned the length of the vector). Assume $\mathcal{R}$ is such that ties are replaced with the mean value of the tied positions (as is done by default e.g. in the 'rankdata' function of the SciPy stats library[VGO+20]).

Given a network $\mathcal{G}$ with adjacency matrix $\mathbf{A}$, define the following shorthand:

$$\mathcal{R}(\mathcal{G}) := \mathcal{R}\left(\text{vect}_{\mathcal{G}_1,\mathcal{G}_2}(\mathbf{A})\right) , \tag{8.4.8}$$
$$\scriptstyle \mathcal{G}_1\mathcal{G}_2$$

where $\text{vect}_{\mathcal{G}_1,\mathcal{G}_2}$ is the same sparsity-adjusted vectorization operator that was defined in section 8.4.2.1.

8.4.2.3 Positive and Negative Spearman Correlation

The "**positive Spearman correlation**" of $\mathcal{G}_1$ and $\mathcal{G}_2$ is defined to be the (sparsity-adjusted) Spearman correlation[9] of $\mathcal{G}_1^+$ and $\mathcal{G}_2^+$:

$$\rho_S^+(\mathcal{G}_1, \mathcal{G}_2) := \frac{\hat{\mathrm{cov}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^+), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+)\right]}{\sqrt{\hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^+)\right]} \cdot \sqrt{\hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+)\right]}} . \tag{8.4.9}$$

Notice that the sparsity adjustment is in terms of $\mathcal{G}_1$ and $\mathcal{G}_2$, not $\mathcal{G}_1^+$ and $\mathcal{G}_2^+$. This is noteworthy to the extent that the union of the edge sets of $\mathcal{G}_1^+$ and $\mathcal{G}_2^+$ could, in principle, be strictly smaller than the union of the edge sets of $\mathcal{G}_1$ and $\mathcal{G}_2$. Therefore one might sensibly argue that the strictest possible sparsity adjustment, that in terms of $\mathcal{G}_1^+$ and $\mathcal{G}_2^+$, should be applied when defining a notion of "positive Spearman correlation". Anything else might not be compatible with the principle outlined in section 8.2.7. On the other hand, one could also argue that a correlation which "preserves the context of the original networks" admits a more useful interpretation, at least when trying to compare or combine[10] the values of the positive and negative Spearman correlations.

Completely analogously, the "**negative Spearman correlation**" of $\mathcal{G}_1$ and $\mathcal{G}_2$ is defined to be the (sparsity-adjusted) Spearman correlation of $\mathcal{G}_1^-$ and $\mathcal{G}_2^-$:

$$\rho_S^-(\mathcal{G}_1, \mathcal{G}_2) := \frac{\hat{\mathrm{cov}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^-), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)\right]}{\sqrt{\hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)\right]}} \tag{8.4.10}$$

Analogous comments about sparsity adjustments apply here too, of course.

8.4.2.4 Vector Concatenation

The concatenation or "direct sum" $\vec{\mathbf{v}}_1 \oplus \vec{\mathbf{v}}_2 \in \mathbb{R}^{I_1+I_2}$ of two vectors $\vec{\mathbf{v}}_1 \in \mathbb{R}^{I_1}$ and $\vec{\mathbf{v}}_2 \in \mathbb{R}^{I_2}$ is defined such that

$$(\vec{\mathbf{v}}_1 \oplus \vec{\mathbf{v}}_2)^{(i)} := \begin{cases} v_1^{(i)} & 1 \le i \le I_1 \\ v_2^{(i-I_1)} & I_1 + 1 \le i \le I_1 + I_2 . \end{cases} \tag{8.4.11}$$

---

[9]The "S" in the notation "ρ_S" can be interpreted to refer to either "Spearman", "sparsity-adjusted", or both. "ρ" of course is commonly-used notation for a correlation coefficient.

[10] Admittedly my main motivation for defining the sparsity adjustments for the positive and negative Spearman correlations using the edge sets of the original networks was to facilitate the interpretation of the mixed-sign Spearman correlation. If the sparsity adjustment for the positive Spearman correlation was in terms of $\mathcal{G}_1^+$ and $\mathcal{G}_2^+$, and that for the negative Spearman correlation in terms of $\mathcal{G}_1^-$ and $\mathcal{G}_2^-$, then as far as I can tell there would in general be no straightforward way to write the mixed-sign Spearman correlation as a weighted sum of the positive and negative Spearman correlations. I came up with the notion of mixed-sign Spearman correlation before observing the need for sparsity adjustments, so it perhaps is not altogether surprising that the two notions do not seem to be easily compatible.

This has the intuitive interpretation of first listing the entries of $\vec{\mathbf{v}}_1$ and then appending to that list the entries of $\vec{\mathbf{v}}_2$.

Some properties are obvious consequences of the definition. For example:

$$(\vec{\mathbf{v}}_1 + \vec{\mathbf{w}}_1) \oplus (\vec{\mathbf{v}}_2 + \vec{\mathbf{w}}_2) = (\vec{\mathbf{v}}_1 \oplus \vec{\mathbf{v}}_2) + (\vec{\mathbf{w}}_1 \oplus \vec{\mathbf{w}}_2). \tag{8.4.12}$$

Here is a second example using the standard inner product, $\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle := \sum_{i=1}^{I} x^{(i)} y^{(i)}$:

$$\langle \vec{\mathbf{v}}_1 \oplus \vec{\mathbf{v}}_2, \vec{\mathbf{w}}_1 \oplus \vec{\mathbf{w}}_2 \rangle = \langle \vec{\mathbf{v}}_1, \vec{\mathbf{w}}_1 \rangle + \langle \vec{\mathbf{v}}_2, \vec{\mathbf{w}}_2 \rangle. \tag{8.4.13}$$

As seen above, for many operations on concatenated vectors in $\mathbb{R}^{I_1 + I_2}$ we can treat what happens in $\mathbb{R}^{I_1}$ separately of what happens in $\mathbb{R}^{I_2}$.

8.4.2.5   Mixed-Sign Spearman Definition

Without explaining or motivating anything, the definition may be written as

$$\rho_S^{\pm}(\mathcal{G}_1, \mathcal{G}_2) := \frac{\widehat{\mathrm{cov}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^+) \oplus \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^-), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+) \oplus \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)\right]}{\sqrt{\widehat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^+) \oplus \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\widehat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+) \oplus \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)\right]}}. \tag{8.4.14}$$

Using the sparsity-adjusted rank operator, instead of the typical rank operator $\mathcal{R}$, prevents two networks from scoring high values merely because both are highly sparse. This is because the sparsity-adjusted rank operator only looks at the subset of edges which are nonzero in at least one of the networks, and does not consider edges missing from both networks. The use of the sparsity-adjusted rank operator follows from the principle outlined in section 8.2.7.

8.4.2.6   Subconvex Combination Decomposition Property

The sense in which expression (8.4.14) considers the positive and negative subnetworks separately and then combines the results is made precise below.

Because of Lemma 8.2, the hypotheses of Lemma 8.3 are satisfied. Applying Lemma 8.3

to expressions in both the numerator and denominator of (8.4.14):

$$\rho_S^{\pm}(\mathcal{G}_1, \mathcal{G}_2) = \frac{\hat{\text{cov}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+)\right] + \hat{\text{cov}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)\right]}{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+)\right] + \hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^+)\right] + \hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^-)\right]}}$$

$$= c_+ \frac{\hat{\text{cov}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+)\right]}{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^+)\right]}} + c_- \frac{\hat{\text{cov}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-), \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)\right]}{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^-)\right]}}$$

$$= c_+ \cdot \rho_S^+(\mathcal{G}_1, \mathcal{G}_2) + c_- \cdot \rho_S^-(\mathcal{G}_1, \mathcal{G}_2),$$

(8.4.15)

where the coefficients c_+, c_- equal

$$c_+ = \frac{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^+)\right]}}{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+)\right] + \hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^+)\right] + \hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^-)\right]}},$$

$$c_- = \frac{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^-)\right]}}{\sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^+)\right] + \hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_1^-)\right]} \cdot \sqrt{\hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^+)\right] + \hat{\text{var}}_{\mathcal{G}_1\mathcal{G}_2}\left[\mathcal{R}(\mathcal{G}_2^-)\right]}}.$$

(8.4.16)

Although in equation (8.4.16) we have $c_+, c_- \geq 0$, in general we only have that $c_+ + c_- \leq 1$, cf. Lemma 8.4. In other words, in general the mixed-sign Spearman correlation is only guaranteed to be a **subconvex combination** of the positive Spearman correlation and the negative Spearman correlation. In special cases it may still also be a convex combination sensu stricto. Because this subconvex property can be derived using the Cauchy-Schwarz inequality, one way to interpret any deficit $1 - (c_+ + c_-) > 0$ is as a measurement of how much[11] $\mathcal{G}_1$ and $\mathcal{G}_2$ "fail to align in almost the same direction".

Although the subconvex property does still correspond to a weighted sum, and does have some merit for possessing a useful interpretation, it would nevertheless still be fair to argue that the mixed-sign Spearman correlation does not fully live up to the promised easy

---

[11] More precisely, it measures how much $\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^+) \oplus \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_1^-)$ and $\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^+) \oplus \underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}(\mathcal{G}_2^-)$ "fail to almost align onto the same line". The Cauchy-Schwarz inequality is an equality only when two vectors are linearly dependent, so in the same (or exactly opposite) directions.

interpretation from section 8.2.6. The mixed-sign Spearman correlation does still satisfy the double penalization principle. Despite in general being a subconvex combination but not necessarily a convex combination, in practice the mixed-sign Spearman correlation also does not seem to necessarily give overly "conservative" results "biased" towards 0.

Perhaps one could argue that any such "shrinkage" might even be desirable, given that the sparsity adjustments defining the positive and negative Spearman correlations are not as aggressive as possible (cf. section 8.4.2.3). By possibly failing to adhere to the principle of section 8.2.7, one might be concerned that the positive and negative Spearman correlations artifactually overstate similarities. While this argument might be true to an extent, I would also caution that "two wrongs don't make a right". A priori we have no reason to believe that hypothetical "corrections" from the subconvex property would (or even could) exactly counterbalance hypothetical artifacts from the chosen sparsity adjustments for the positive and negative Spearman correlations.

8.4.3 Mixed-Sign Jaccard Similarity

Unweighted Jaccard similarity provides a qualitative comparison of networks by comparing the presence/absence of edges. Weighted Jaccard similarity provides a comparison that is a little quantitative and a little qualitative, by increasing the importance of comparisons for edge pairs where one of the edges has a large edge weight. Mixed-sign versions of both comparison methods provide an advantage over unsigned versions by separately considering presence/absence for positive and negative edges and then sensibly recombining the results. By considering this extra information contained in the signs of the edge weights, the proposed mixed-sign versions are not fooled by drastic changes to the network structure to which the unsigned versions are completely oblivious.

8.4.3.1 Binary Classification and Mixed Signs

Considering positive and negative edges to correspond to distinct phenomena means we effectively have a ternary classification problem (present and positive, present and negative, absent). Our goal is to recast this ternary classification problem as two inter-related binary classification problems, one for the positive edges and one for the negative edges.

To avoid confusion with the edge signs, below I use the following alternative terminology to refer to the four possible binary classification outcomes:

- "**True Hit**" (TH) := "True Positive",

- "**False Hit**" (FH) := "False Positive",

- "**True Miss**" (TM) := "True Negative",

- "**False Miss**" (FM) := "False Negative".

Positive Edge Binary Classification Problem   The positive edge binary classification problem has

- hits = $\{+\}$,

- misses = $\{0, -\}$.

So for the positive edge binary classification problem:

- a "**true hit**" occurs when both the truth and the estimate are positive,

- a "**false hit**" occurs when the truth is non-positive (0 or $-$) but the estimate is positive,

- a "**true miss**" occurs when the truth is non-positive and the estimate is non-positive, and

- a "**false miss**" occurs when the truth is positive but the estimate is non-positive.

Negative Edge Binary Classification Problem   Analogously the negative edge binary classification problem has

- hits = $\{-\}$,

- misses = $\{0, +\}$.

So for the negative edge binary classification problem:

- a "**true hit**" occurs when both the truth and the estimate are negative,

- a "**false hit**" occurs when the truth is non-negative (0 or $+$) but the estimate is negative,

- a "**true miss**" occurs when the truth is non-negative and the estimate is non-negative, and

- a "**false miss**" occurs when the truth is negative but the estimate is non-negative.

Origin of Double Penalization   From this perspective of having two inter-related binary classification problems, confusing a positive interaction with a negative interaction, or vice versa, is arguably both

- a **false hit**, inferring a type of interaction which does not exist, and

- a **false miss**, failing to infer a type of interaction which does exist.

Thus an ideal performance metric should "doubly penalize" such errors. This was the original insight which led to the more general double penalization principle from section 8.2.5.

8.4.3.2 (Unsigned) Unweighted Definition

The (unweighted) Jaccard similarity of two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ (with non-negative edge weights) is the Jaccard similarity of their edge sets $\mathcal{E}_1$ and $\mathcal{E}_2$:

$$\mathcal{S}_J(\mathcal{G}_1, \mathcal{G}_2) := \frac{|\mathcal{E}_1 \cap \mathcal{E}_2|}{|\mathcal{E}_1 \cup \mathcal{E}_2|} = \frac{|\mathcal{E}_1 \cap \mathcal{E}_2|}{|\mathcal{E}_1 \cap \mathcal{E}_2| + |\mathcal{E}_1 \setminus \mathcal{E}_2| + |\mathcal{E}_2 \setminus \mathcal{E}_1|} . \tag{8.4.17}$$

Recall that

$$\mathcal{E}_1 := \mathrm{supp}(\epsilon_1) \,, \mathcal{E}_2 := \mathrm{supp}(\epsilon_2)$$

where $\epsilon_1 \geq 0$ is the edge function corresponding to $\mathcal{G}_1$, and likewise for $\epsilon_2 \geq 0$.

Notice that if $\mathcal{G}_1$ is an "estimate" $\hat{\mathcal{G}}$, and $\mathcal{G}_2$ is the "ground truth" $\mathcal{G}_*$ then their (unweighted) Jaccard similarity corresponds to

$$\mathcal{S}_J(\hat{\mathcal{G}}, \mathcal{G}_*) = \frac{TH}{TH + FH + FM} , \tag{8.4.18}$$

where TH is the number of true hits, FH is the number of false hits, and FM is the number of false misses. Thus Jaccard similarity is no less conservative than the minimum of precision and recall, since it has the same numerator and one additional non-negative term in the denominator.

It is beneficial that Jaccard similarity does not reward true misses[12], because not doing so helps to prevent considering two highly different networks as similar merely because both are highly sparse. Cf. again section 8.2.7.

As an aside, in meteorology Jaccard similarity (as used for binary classification) is often called the "critical success index" or "threat score"[HFJS10]. Many binary classification metrics considered improvements over Jaccard similarity (e.g. the "equitable threat score") are studied in meteorology[HFJS10]. Future work examining whether these other binary classification metrics also admit extensions to the comparison of networks of mixed-sign edge weights, and whether such extensions would be useful, would be interesting.

8.4.3.3 (Unsigned) Weighted Definition

For two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ with non-negative edge weights and edge functions ϵ_1 and ϵ_2 as above, their (unsigned) weighted Jaccard similarity is

$$\mathcal{S}_J^W(\mathcal{G}_1, \mathcal{G}_2) := \frac{\displaystyle\sum_{i,j} \min\{\epsilon_1(i,j), \epsilon_2(i,j)\}}{\displaystyle\sum_{i,j} \max\{\epsilon_1(i,j), \epsilon_2(i,j)\}} . \tag{8.4.19}$$

---

[12]TM would denote the number of true misses. Notably this is absent from the numerator of equation (8.4.18). Thus higher values of TM do not lead to higher values of Jaccard similarity.

By considering the graphs' "skeletons", which have edge functions $(\text{sign} \circ \epsilon_1)$ and $(\text{sign} \circ \epsilon_2)$, one recovers the regular (unweighted) Jaccard similarity. Specifically

$$\mathcal{S}_J^W(\mathcal{S}(\mathcal{G}_1), \mathcal{S}(\mathcal{G}_2)) = \mathcal{S}_J(\mathcal{G}_1, \mathcal{G}_2). \tag{8.4.20}$$

Recall the definitions of graph (unsigned) skeletons from section 8.3.3.1.

8.4.3.4 Positive and Negative Jaccard Similarities

From here on let $\mathcal{G}_1$, $\mathcal{G}_2$ denote two general networks with *mixed-sign* edge weights. (In the previous sections 8.4.3.2 and 8.4.3.3 it was assumed for the sake of simplicity that the networks $\mathcal{G}_1$, $\mathcal{G}_2$ had non-negative edge weights.)

One can define the "positive (unweighted) Jaccard similarity" of two general networks as:

$$\mathcal{S}_J^+(\mathcal{G}_1, \mathcal{G}_2) := \mathcal{S}_J(\mathcal{G}_1^+, \mathcal{G}_2^+), \tag{8.4.21}$$

as well as their "positive weighted Jaccard similarity",

$$(\mathcal{S}_J^W)^+(\mathcal{G}_1, \mathcal{G}_2) := \mathcal{S}_J^W(\mathcal{G}_1^+, \mathcal{G}_2^+). \tag{8.4.22}$$

Completely analogously, the "negative (unweighted) Jaccard similarity" of two general networks is:

$$\mathcal{S}_J^-(\mathcal{G}_1, \mathcal{G}_2) := \mathcal{S}_J(\mathcal{G}_1^-, \mathcal{G}_2^-), \tag{8.4.23}$$

and their "negative weighted Jaccard similarity",

$$(\mathcal{S}_J^W)^-(\mathcal{G}_1, \mathcal{G}_2) := \mathcal{S}_J^W(\mathcal{G}_1^-, \mathcal{G}_2^-). \tag{8.4.24}$$

The above definitions make sense because the positive $\mathcal{G}^+$ and negative subnetworks $\mathcal{G}^-$ have non-negative edge weights by definition. Therefore "standard" network dissimilarity methods may be applied to them.

Observe how the weighted versions $(\mathcal{S}_J^W)^+$, $(\mathcal{S}_J^W)^-$ reduce to their respective unweighted counterparts $\mathcal{S}_J^+$, $\mathcal{S}_J^-$ when applied to (signed) skeletons:

$$(\mathcal{S}_J^W)^+(\mathcal{S}^\pm(\mathcal{G}_1), \mathcal{S}^\pm(\mathcal{G}_2)) = \mathcal{S}_J^+(\mathcal{G}_1, \mathcal{G}_2), \quad (\mathcal{S}_J^W)^-(\mathcal{S}^\pm(\mathcal{G}_1), \mathcal{S}^\pm(\mathcal{G}_2)) = \mathcal{S}_J^-(\mathcal{G}_1, \mathcal{G}_2). \tag{8.4.25}$$

8.4.3.5 Mixed-Sign Unweighted Jaccard Similarity

The "mixed sign (unweighted) Jaccard similarity" of two networks $\mathcal{G}_1$, $\mathcal{G}_2$ with mixed-sign edge weights is defined to equal

$$\mathcal{S}_J^\pm(\mathcal{G}_1, \mathcal{G}_2) := \mathcal{S}_J(\mathcal{G}_1^+ \oplus \mathcal{G}_1^-, \mathcal{G}_2^+ \oplus \mathcal{G}_2^-) = \frac{|\mathcal{E}_1^+ \cap \mathcal{E}_2^+| + |\mathcal{E}_1^- \cap \mathcal{E}_2^-|}{|\mathcal{E}_1^+ \cup \mathcal{E}_2^+| + |\mathcal{E}_1^- \cup \mathcal{E}_2^-|}. \tag{8.4.26}$$

($\mathcal{E}^+$ is the edge set of a graph's positive part, analogously for $\mathcal{E}^-$.)

Notice that if a positive edge is estimated as negative, it will show up once both as a false hit for the negative edges (contributing once to $|\mathcal{E}_1^- \cup \mathcal{E}_2^-|$) as well as a false miss for the positive edges (contributing once to $|\mathcal{E}_1^+ \cup \mathcal{E}_2^+|$). The analogous conclusions hold when a negative edge is estimated as positive.

Therefore, *mixed-sign (unweighted) Jaccard similarity satisfies the desired double penalization property for incorrectly estimated signs.* In fact:

$$\mathcal{S}_J^{\pm}(\hat{\mathcal{G}}, \mathcal{G}_*) = \frac{TH^+ + TH^-}{TH^+ + TH^- + FH^+ + FH^- + FM^+ + FM^-} \, . \qquad (8.4.27)$$

The analogy between (8.4.18) and (8.4.27) strongly suggests that mixed-sign Jaccard similarity is "the correct" generalization of (unsigned) Jaccard similarity.

Notice how this definition again does not reward true misses[13], which could occur merely due to either the positive and/or negative subnetworks being highly sparse. This helps mixed-sign (unweighted) Jaccard similarity to avoid considering two highly different networks as similar merely because both are highly sparse, and/or both have few positive/negative edges. Cf. again section 8.2.7.

Moreover the mixed-sign (unweighted) Jaccard similarity can be described simply as a convex combination/weighted average of the positive (unweighted) Jaccard similarity and the negative (unweighted) Jaccard similarity:

$$\mathcal{S}_J^{\pm}(\mathcal{G}_1, \mathcal{G}_2) = c_+ \mathcal{S}_J^+(\mathcal{G}_1, \mathcal{G}_2) + c_- \mathcal{S}_J^-(\mathcal{G}_1, \mathcal{G}_2) \, , \qquad (8.4.28)$$

with the coefficients of the convex combination above being

$$c_+ = \frac{|\mathcal{E}_1^+ \cup \mathcal{E}_2^+|}{|\mathcal{E}_1^+ \cup \mathcal{E}_2^+| + |\mathcal{E}_1^- \cup \mathcal{E}_2^-|} \, , \quad c_- = \frac{|\mathcal{E}_1^- \cup \mathcal{E}_2^-|}{|\mathcal{E}_1^+ \cup \mathcal{E}_2^+| + |\mathcal{E}_1^- \cup \mathcal{E}_2^-|} \, . \qquad (8.4.29)$$

8.4.3.6 Mixed-Sign Weighted Jaccard Similarity

Completely analogously the "mixed sign weighted Jaccard similarity" of two networks $\mathcal{G}_1$, $\mathcal{G}_2$ with mixed-sign edge weights is defined to equal

$$
\begin{aligned}
(\mathcal{S}_J^W)^{\pm}(\mathcal{G}_1, \mathcal{G}_2) :=\, & \mathcal{S}_J^W(\mathcal{G}_1^+ \oplus \mathcal{G}_1^-, \mathcal{G}_2^+ \oplus \mathcal{G}_2^-) \\
=\, & \frac{\displaystyle\sum_{i,j} \min\{\epsilon_1^+(i,j), \epsilon_2^+(i,j)\} + \sum_{i,j} \min\{\epsilon_1^-(i,j), \epsilon_2^-(i,j)\}}{\displaystyle\sum_{i,j} \max\{\epsilon_1^+(i,j), \epsilon_2^+(i,j)\} + \sum_{i,j} \max\{\epsilon_1^-(i,j), \epsilon_2^-(i,j)\}} \, .
\end{aligned}
\qquad (8.4.30)
$$

---

[13]TM^+ would denote the number of true misses for the positive subnetworks, TM^- would denote the number of true misses for the negative subnetworks. Notably these are absent from the numerator of equation (8.4.27). Thus neither higher values of TM^+, nor higher values of TM^-, lead to higher values of mixed-sign Jaccard similarity.

Like the unsigned version (cf. section 8.4.3.3), the mixed-sign weighted Jaccard similarity reduces to its unweighted counterpart when applied to skeletons:

$$(\mathcal{S}_J^W)^{\pm}(\mathcal{S}^{\pm}(\mathcal{G}_1), \mathcal{S}^{\pm}(\mathcal{G}_2)) = \mathcal{S}_J^{\pm}(\mathcal{G}_1, \mathcal{G}_2). \tag{8.4.31}$$

Like its unweighted counterpart, the mixed-sign weighted Jaccard similarity also enjoys the convex combination decomposition property:

$$(\mathcal{S}_J^W)^{\pm}(\mathcal{G}_1, \mathcal{G}_2) = c_+ (\mathcal{S}_J^W)^+(\mathcal{G}_1, \mathcal{G}_2) + c_- (\mathcal{S}_J^W)^-(\mathcal{G}_1, \mathcal{G}_2), \tag{8.4.32}$$

with the coefficients c_+, c_- of the convex combination being

$$\begin{aligned} c_+ &:= \frac{\sum_{i,j} \max\{\epsilon_1^+(i,j), \epsilon_2^+(i,j)\}}{\sum_{i,j} \max\{\epsilon_1^+(i,j), \epsilon_2^+(i,j)\} + \sum_{i,j} \max\{\epsilon_1^-(i,j), \epsilon_2^-(i,j)\}}, \\[2em] c_- &:= \frac{\sum_{i,j} \max\{\epsilon_1^-(i,j), \epsilon_2^-(i,j)\}}{\sum_{i,j} \max\{\epsilon_1^+(i,j), \epsilon_2^+(i,j)\} + \sum_{i,j} \max\{\epsilon_1^-(i,j), \epsilon_2^-(i,j)\}}. \end{aligned} \tag{8.4.33}$$

Therefore, like its unweighted counterpart, mixed-sign weighted Jaccard similarity also satisfies the double penalization principle.

8.4.3.7 Mixed-Sign False Discovery Rate and Mixed-Sign False Miss Rate

Herein assume again that $\mathcal{G}_1$ is an "estimate" $\hat{\mathcal{G}}$, and $\mathcal{G}_2$ is the "ground truth", and both have non-negative edge weights. Similar to the "measure of goodness" defined by (8.4.18), we can also define two corresponding "measures of badness":

- the **false discovery rate**:

$$\text{FDR}(\hat{\mathcal{G}}, \mathcal{G}_*) := \frac{FH}{TH + FH}, \tag{8.4.34}$$

- and the **false miss rate**[14]:

$$\text{FMR}(\hat{\mathcal{G}}, \mathcal{G}_*) := \frac{FM}{TH + FM}. \tag{8.4.35}$$

Note that, unlike the definition of Jaccard similarity, these definitions are not symmetric with respect to the roles of $\hat{\mathcal{G}}$ and $\mathcal{G}_*$.

Now allow $\hat{\mathcal{G}}$ and $\mathcal{G}_*$ to again have mixed-sign edge weights. Analogous to the relationship between definitions (8.4.18) and (8.4.27), we can also define mixed-sign versions of the false discovery rate (8.4.34) and the false miss rate (8.4.35):

---

[14]More typically called "false negative rate", but that would be confusing in this context.

- the **mixed-sign false discovery rate**:

$$\mathrm{FDR}^{\pm}(\hat{\mathcal{G}}, \mathcal{G}_*) := \frac{FH^+ + FH^-}{TH^+ + TH^- + FH^+ + FH^-}, \tag{8.4.36}$$

- and the **mixed-sign false miss rate**:

$$\mathrm{FMR}^{\pm}(\hat{\mathcal{G}}, \mathcal{G}_*) := \frac{FM^+ + FM^-}{TH^+ + TH^- + FM^+ + FM^-}. \tag{8.4.37}$$

Just like the mixed-sign Jaccard similarity, these can also be written straightforwardly as weighted averages of the corresponding single-sign performance metrics applied to the positive $\hat{\mathcal{G}}^+, \mathcal{G}_*^+$ and negative $\hat{\mathcal{G}}^-, \mathcal{G}_*^-$ subnetworks.

8.4.4   Mixed-Sign DeltaCon Distance

Unweighted DeltaCon provides a qualitative comparison of networks by quantifying the correspondence of their paths. Weighted DeltaCon provides a comparison that is a little quantitative and a little qualitative, by increasing the importance of comparisons for paths where some edges have large edge weights. Herein mixed-sign versions of both unweighted and weighted DeltaCon are proposed which are not fooled by drastic changes to the network structure to which the unsigned versions from previous work are completely oblivious.

8.4.4.1   Paths and Powers of Adjacency Matrices

Given an adjacency matrix (cf. section 8.3.1), for any $n \geq 1$, the matrix $\mathbf{A}^n$ gives information about the paths of length n in the network. (The case where $n = 1$ corresponds to paths of length 1, i.e. the edges.) Specifically, given a pair (s_1, s_2), the (s_1, s_2)'th entry $[\mathbf{A}^n]_{s_1, s_2}$ of the nth power $\mathbf{A}^n$ of the adjacency matrix $\mathbf{A}$ equals the sum over all paths from s_1 to s_2 with length exactly n of the product of all of the weights of all of the edges in each path.

For example, if there were two paths of length exactly 3 from s_1 to s_2, and the weights of the edges in the first were 1, -4, 2, and the weights of the edges in the second were -2, 6, -5, then the (s_1, s_2)'th entry $[\mathbf{A}^3]_{s_1, s_2}$ of $\mathbf{A}^3$ would be $(1)(-4)(2) + (-2)(6)(-5) = -8 + 60 = 52$.

Usefulness of this Formalism   The usefulness of encoding information about paths this way is uncontroversial in the case of an adjacency matrix with only 1's and 0's. Then the entries count the number of paths of a given length in between any two nodes.

In the case of an unsigned weighted graph the usefulness of this notion is perhaps more controversial. One might wonder if it makes more sense to take the sum of the weights along each path rather than the product. It is probably even more controversial in the case of signed and weighted graphs, since taking products over weights with different signs leads to difficult to predict "fluctuations" in the sign of the contribution from each path.

In particular, one might question whether the effect on signs given by multiplying the edges in the path is scientifically meaningful in a given context, or whether (instead of paths potentially cancelling one another out) edges within a path that have opposite signs should be allowed to cancel each other out.

Usefulness for Microbial Ecology I believe there is unlikely to be a universal answer. Instead, whether powers of adjacency matrices encode useful information about a mixed-sign network's paths likely depends on the context established by the specific scientific questions of interest. In this context I believe that powers of adjacency matrices do make sense.

For example, consider a situation where strain[15] s_1 strongly promotes the growth of strain s_2, and s_2 produces compounds toxic to strain s_3. Cf. figure 8.3, where s_1 could be the blue blob microbe, s_2 could be the pink spiral microbe, and s_3 could be the red rod microbe. I think it makes more sense for the resulting path from s_1 to s_3 to be given a strong negative weight, which corresponds only to the product of the weights of the individual edges.

If one used the sum of the edge weights instead, one would instead "cancel" the effects of s_1 on s_2 and of s_2 on s_3. That makes no sense, because large numbers of s_1 would promote large numbers of s_2, which would in turn produce large amounts of the compound toxic to s_3. It seems clear that the net indirect effect of s_1 on s_3 is strong and negative.

Neumann Series One might ask which length of path n should be considered. A possible answer to this is "all of them", by using the Neumann series of the adjacency matrix $\mathbf{A}$. For a general matrix $\mathbf{M}$, its Neumann series is

$$\mathbf{M} + \mathbf{M}^2 + \mathbf{M}^3 + \cdots + \mathbf{M}^n + \cdots = (\mathbf{I} - \mathbf{M})^{-1}. \tag{8.4.38}$$

This converges[16] if and only if the spectral radius of $\mathbf{M}$ is less than 1. Therefore it makes sense to think of the Neumann series as the matrix analogue of the geometric series for (real or complex) numbers.

The spectral radius lower bounds the Frobenius norm and any operator norm[17], so a sufficient (but not necessary) condition for the Neumann series of $\mathbf{M}$ to converge is if the value of one of these norms is less than 1 for $\mathbf{M}$. Therefore in general convergence of the corresponding Neumann series can be accomplished by rescaling $\mathbf{M}$ by a positive constant.

8.4.4.2 DeltaCon Distance Overview

For both the unsigned and mixed-sign versions of DeltaCon distance, the general idea can be divided into two steps:

---

[15]Herein I use "strains" to refer equally to strains belonging to the same species(/genus/family/etc.) as well as to strains belonging to different species(/genera/families/etc.), because the distinction is irrelevant for setting up the abstract problem. It may matter for the implementation of a specific experiment.

[16]This fact is not obvious, but appears to be a "folk theorem" such that it is difficult to find a "canonical" reference for the proof.

[17]See footnote 16, because the same comments apply here too.

1. For each of the two networks, define a matrix which encodes the "path structure" of the network.

2. Define a distance between the two networks by evaluating a distance between the two matrices produced in the first step.

For the first step we form the Neumann series $\mathbf{S}$ of a suitable matrix $\mathbf{W}$. For the second step, we evaluate (a modified version of) the Matusita ("root Euclidean") distance of the two matrices $\mathbf{S}_1$, $\mathbf{S}_2$ constructed in the first step.

8.4.4.3 DeltaCon Distance Definition

The definition of DeltaCon does not use the Neumann series of (a rescaled version of) the adjacency matrix $\mathbf{A}$, but instead uses the Neumann series of a matrix $\mathbf{W}$ based on $\mathbf{A}$. The formula (8.4.39) for $\mathbf{W}$ in terms of $\mathbf{A}$ is derived from a linearized version of the belief propagation algorithm called FaBP[KKK$^+$11]. The authors of [KSV$^+$16] state that nevertheless the intuition for using the Neumann series of the matrix $\mathbf{W}$ is still the same as the intuition described above for using the Neumann series of the adjacency matrix $\mathbf{A}$. The scaling constant ε used below in the definition (8.4.39) of this matrix $\mathbf{W}$ has the interpretation of encoding "the influence between neighboring nodes"[KSV$^+$16].

Combining the exact definition[18] from [KKK$^+$11] with the notation from [KSV$^+$16], I use the following for the definition of $\mathbf{W}$:

$$\mathbf{W} := \frac{1}{1-\varepsilon^2}(\varepsilon\mathbf{A} - \varepsilon^2\mathbf{D}). \tag{8.4.39}$$

(The degree matrix $\mathbf{D}$ is a general notion defined in section 8.3.2.) The Neumann series $\mathbf{S}$ of $\mathbf{W}$, which encodes the "path structure" of the network, is called the "similarity matrix" in [KSV$^+$16] because each entry can be interpreted as measuring the "affinity" of two nodes in terms of the paths between them.

$$\mathbf{S} := (\mathbf{I} - \mathbf{W})^{-1} = \mathbf{I} + \mathbf{W} + \mathbf{W}^2 + \dots. \tag{8.4.40}$$

To ensure the definition of DeltaCon distance works we need to ensure the convergence of the Neumann series defining the similarity matrix $\mathbf{S}$. To guarantee the convergence of the Neumann series defining $\mathbf{S}$, we choose, for any given input $\mathbf{A}$, a corresponding value for ε in (8.4.39) that guarantees that the spectral radius of the resulting $\mathbf{W}$ defined in terms of $\mathbf{A}$ and ε will be less than 1.

The normalization implicit in the definition of $\mathbf{W}$ (8.4.39) and of the similarity matrix (8.4.40) may help DeltaCon distance to avoid judging two highly different networks as being similar merely because both are sparse. Cf. again section 8.2.7, as well as the comments at the end of section 8.4.1.2. I am uncertain whether this is really true in practice, but it does seem plausible.

---

[18]The definition of $\mathbf{W}$ from [KSV$^+$16] slightly differs and is not the original definition.

Unsigned Version The definition given in [KSV⁺16] was intended to work for unsigned but weighted networks. Without being modified, it does not work for general networks with mixed-sign edge weights. I describe in detail the proposals made in [KSV⁺16] for (1) the value of ε in the definition (8.4.39) of $\mathbf{W}$, and (2) the distance to apply to the similarity matrices $\mathbf{S_1}$ and $\mathbf{S_2}$. Along with using the original definition of $\mathbf{W}$ from [KKK⁺11], these are the two aspects of the definition from [KSV⁺16] that I changed in order to guarantee the new definition would work for all networks with mixed-sign edge weights.

Step 1: The authors of [KSV⁺16] propose using the following for ε:

$$\varepsilon := \frac{1}{1 + \max_i [\mathbf{D}]_{ii}} . \tag{8.4.41}$$

However, they don't include an argument showing why this value (ostensibly) works. This is notable because the previous paper[KKK⁺11] (i) only discussed unweighted (and unsigned) networks, (ii) used a different version of $\mathbf{W}$ (corresponding to (8.4.39) above) than the definition used in [KSV⁺16], and (iii) proposed using a slightly different value for ε (for which a proof was provided).

Step 2: Given two adjacency matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, and defining their corresponding similarity matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ as in equation (8.4.40), [KSV⁺16] defines the (unsigned) DeltaCon distance to be the Matusita distance of $\mathbf{S}_1$ and $\mathbf{S}_2$:

$$d_{DC}(\mathcal{G}_1, \mathcal{G}_2) := \sqrt{\sum_{i,j} (\sqrt{[\mathbf{S}_1]_{ij}} - \sqrt{[\mathbf{S_2}]_{ij}})^2} . \tag{8.4.42}$$

The authors of [KSV⁺16] suggest several reasons for using Matusita distance instead of e.g. Frobenius distance, but one important reason they mention is "boosting" small entries in $[0, 1]$ (given that $\sqrt{x} \geq x$ for all $x \in [0, 1]$).

Mixed-Sign Version To modify the definition from [KSV⁺16] to apply to networks with mixed-sign edge weights, I made the following two changes.

Step 1: I showed that a sufficient condition for a value of ε to work with any network with mixed-sign edge weights is for $\varepsilon < \frac{1}{1+\|\mathbf{A}\|_\infty}$ (see Lemma 8.5 from section 8.A.2). I also showed that this bound is sharp (see Lemma 8.6 from section 8.A.2), in the sense that there exist adjacency matrices $\mathbf{A}$ of networks with mixed-sign edge weights for which choosing $\varepsilon := \frac{1}{1+\|\mathbf{A}\|_\infty}$ leads to a $\mathbf{W}$ with spectral radius exactly equal to 1. While such problematic adjacency matrices do not appear to be generic, for safety I chose to use instead:

$$\varepsilon := \frac{1}{2 + \|\mathbf{A}\|_\infty} . \tag{8.4.43}$$

In practice, even when $\varepsilon := \frac{1}{1+\|\mathbf{A}\|_\infty}$ technically works for a given network, if the spectral radius of the resulting $\mathbf{W}$ is still very close to 1, the DeltaCon distance appears to be poorly behaved, e.g. by producing unreasonably large values (data not shown). Using the more

conservative definition (8.4.43) of ε, the resulting $\mathbf{W}$ matrices had spectral radii bounded further away from 1, making the resulting values of the DeltaCon distance more stable than they were before.

Step 2: In equation (8.4.42), I replaced $\sqrt{[\mathbf{S}]_{ij}}$ with $\mathrm{sign}\left([\mathbf{S}]_{ij}\right) \cdot \sqrt{|[\mathbf{S}]_{ij}|}$:

$$d_{DC}^{\pm}(\mathcal{G}_1, \mathcal{G}_2) := \sqrt{\sum_{i,j}\left(\mathrm{sign}([\mathbf{S}_1]_{ij}) \cdot \sqrt{|[\mathbf{S}_1]_{ij}|} - \mathrm{sign}([\mathbf{S}_2]_{ij}) \cdot \sqrt{|[\mathbf{S_2}]_{ij}|}\right)^2}. \qquad (8.4.44)$$

The choice to use $\mathrm{sign}\left([\mathbf{S}]_{ij}\right) \cdot \sqrt{|[\mathbf{S}]_{ij}|}$ instead of $\sqrt{[\mathbf{S}]_{ij}}$ appears to be necessary for at least two reasons. First, it ensures that the distance is defined (by not assuming that the entries of $\mathbf{S}$ are necessarily non-negative in the more general case of mixed-sign edge weights). Second, it also guarantees that the double penalization principle (cf. section 8.2.5) is satisfied. At the same time, the new definition (8.4.44) also retains the same "boosting" property that helped motivate definition (8.4.42), because $\sqrt{|x|} \geq |x|$ for all $x \in [-1, 1]$. So nothing important appears to be lost because of the change, as reflected perhaps in how (8.4.44) reduces to (8.4.42) in the case that the entries of $\mathbf{S}_1$ and $\mathbf{S}_2$ are non-negative.

8.5   Methods

$1,000$ random networks were generated (section 8.5.1) and classified according to which sign had the larger number of edges (section 8.5.2). Three attacks were applied to each of the networks (section 8.5.3). The values of the mixed-sign network comparison methods between the original networks and their attacked versions were compared with the corresponding values for variants which do not use the entirety of the mixed-sign network structure or which do not obey the double penalization principle (section 8.5.4). Violin plots of the results were generated using Matplotlib [Hun07] version 3.4.1 and Seaborn [Was21] version 0.11.1. Complete implementation details can be found in the code at the relevant GitLab repository. See https://gitlab.com/krinsman/mixed-sign-networks.

8.5.1   Generating Random Networks

$1,000$ random networks were independently generated as follows:

- A hyperparameter $p \in (0, 1)$ was selected from the uniform distribution.

- 100^2 random entries of an adjacency matrix were independently generated according to a $\mathrm{Beta}(2p, 2(1-p))$ distribution. (Considered as a Dirichlet distribution, this has the same concentration parameter as the uniform distribution.)

- 0.5 was subtracted from all entries, and then all entries were multiplied by 2, with the effect that their support changed from $(0, 1)$ to $(-1, 1)$.

- Each of the entries in the adjacency matrix was set to zero independently and with probability $\frac{1}{2}$ (this means the random networks are Erdos-Renyi).

8.5.2 Predominant and Non-Predominant Sign Edges

For each network, if a majority of the nonzero entries were positive, + was the "predominant sign" and − the "non-predominant sign". If a majority of the nonzero entries were negative, − was the "predominant sign" and + the "non-predominant sign". Results looking at "predominant sign edges" consider only the subnetworks (using the definitions from section 8.3.4) corresponding to the "predominant sign" for any given networks. Similarly for results looking at "non-predominant sign edges". Results were split according to the "predominant sign" and "non-predominant sign" for each network, rather than always grouping the positive subnetworks together and the negative subnetworks together, because the corresponding distributions usually depended only on whether the subnetwork corresponded to the numerically predominant sign for that network, and not to what the particular sign was.

8.5.3 Definitions of the Attacks

Three types of attack were applied to each of the random networks. The corresponding values of similiarity/dissimilarity were then computed between each original random network and their attacked counterparts.

The three types were as follows.

8.5.3.1 Shift Attack

This attack affected both the signs and magnitudes of the edges of the random network. The effect on the signs was the same as the sign flip attack. First, the predominant sign of the network was identified. If the predominant sign was positive, then twice the maximum value of any edge in the network was subtracted from all nonzero edges, forcing all edges to be negative (the non-predominant sign). If the predominant sign was negative, then twice the absolute value of the minimum value of any edge in the network was added to all nonzero edges, forcing all edges to be positive (the non-predominant sign). In both cases the values of all nonzero edges were shifted by a constant (whence the name), leaving their relative ordering unaffected.

8.5.3.2 Magnitude Swap Attack

This attack affected the magnitudes of the edges of the random network, but not the signs. The relative ordering of all magnitudes of all nonzero edges was computed. Then the edge with the largest magnitude had its magnitude swapped with the magnitude of the edge with the smallest magnitude, the edge with the second largest magnitude had its magnitude swapped with the magnitude of the edge with the second smallest magnitude, and so on.

For all n (less than or equal to the number of nonzero edges), the edge with the nth largest magnitude had its magnitude swapped with the magnitude of the edge with the nth smallest magnitude.

8.5.3.3   Sign Flip Attack

This attack affected the signs of the edges of the random network, but not the magnitudes. First, the predominant sign of the network was identified. Then all of the edges corresponding to the predominant sign had their sign flipped (i.e. multiplied by -1), while all edges corresponding to the non-predominant sign were unchanged.

8.5.4   Definitions of Variants

Mixed-sign network comparison methods were compared against related comparison methods that only considered subsets of the features of the networks. Section 8.5.4.1 explains the definitions of the variants used for (entrywise L_1) relative error. Section 8.5.4.2 explains the definitions of the variants used for Spearman correlation. Section 8.5.4.3 explains the definitions of the variants used for Jaccard similarity. Section 8.5.4.4 explains the definitions of the variants used for DeltaCon distance.

8.5.4.1   Relative Error

First I compared the values of relative error with those values resulting from considering (i) only the magnitudes of the original networks or (ii) only the signs of the original networks. I then compared the distribution of relative error values for the whole network with the corresponding values for the subnetworks corresponding to (i) the predominant sign edges only or (ii) the non-predominant sign edges only.

   Whenever computing the relative error of two (sub)networks, the magnitudes of the edges were first normalized by the value of the entrywise L_1 norm of the (sub)networks. Cf. section 8.4.1.4. This served two purposes. First, it ensured that the distribution of values was confined to $[0, 2]$ and thus made the results easier to visualize. Second, it corresponds to what it was done in preliminary work, where only the relative but not absolute sizes of estimates are of interest. The convex combination relationship mentioned in section 8.4.1.3 is therefore no longer directly relevant. Thus the most accurate interpretation of results is somewhat obscured.

   "Magnitudes only" refers to the relative error between the "magnitude skeletons" $\mathcal{M}(\mathcal{G}_1)$, $\mathcal{M}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. "Signs only" refers to the relative error between the "signed skeletons" $\mathcal{S}^{\pm}(\mathcal{G}_1), \mathcal{S}^{\pm}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. (See section 8.3.3 for skeleton definitions.)

8.5.4.2 Spearman Correlation

First I compared the values of mixed-sign Spearman correlation with those values resulting from (i) using the raw Spearman correlation, considering (ii) only the magnitudes of the original networks, or (iii) only the signs of the original networks. I then compared the distribution of mixed-sign Spearman correlation values for the whole network with the corresponding (raw) Spearman correlation values for the subnetworks corresponding to (i) the predominant sign edges only or (ii) the non-predominant sign edges only.

"Raw" Spearman correlation refers to the sparsity-adjusted Spearman correlation of the (unaltered) original networks $\mathcal{G}_1, \mathcal{G}_2$. "Magnitudes only" refers to the sparsity-adjusted Spearman correlation of the "magnitude skeletons" $\mathcal{M}(\mathcal{G}_1), \mathcal{M}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. Similarly "signs only" refers to the sparsity-adjusted Spearman correlation of the "signed skeletons" $\mathcal{S}^{\pm}(\mathcal{G}_1), \mathcal{S}^{\pm}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. (See section 8.3.3 for skeleton definitions.)

As mentioned already in section 8.4.2.5, only looking at the subset of edges which are nonzero in at least one of the networks (i.e. not considering edges missing from both networks) avoids considering two very different networks as similar due only to them both being highly sparse.

Also note that when one of the compared vectors of edge values was constant (and thus its rank vector had 0 variance), the Spearman correlation herein is defined by convention to be 0 (since the ranks from one vector have zero predictive value for predicting the ranks of the other vector).

8.5.4.3 Jaccard Similarity

First I compared the values of mixed-sign unweighted Jaccard similarity and mixed-sign weighted Jaccard similarity with the values resulting from considering only (i) the presence/absence of edges or (ii) the magnitudes of edges. I then compared the mixed-sign unweighted Jaccard similarity values with the Jaccard similarity values for the subnetworks corresponding to (i) the predominant sign edges only or (ii) the non-predominant sign edges only, and then analogously also for the mixed-sign weighted Jaccard similarity.

"Presence/Absence Only" refers to the weighted Jaccard similarity of the "unsigned skeletons" $\mathcal{S}(\mathcal{G}_1), \mathcal{S}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$ (or equivalently the unweighted Jaccard similarity of the "magnitude skeletons" $\mathcal{M}(\mathcal{G}_1), \mathcal{M}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$). "Magnitudes only" refers to the weighted Jaccard similarity of the "magnitude skeletons" $\mathcal{M}(\mathcal{G}_1), \mathcal{M}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. (See section 8.3.3 for skeleton definitions.)

Variants called "Signs Only" for relative error or Spearman correlation effectively correspond to the mixed-sign unweighted Jaccard similarity (which again is equivalent to the mixed-sign weighted Jaccard similarity of the "signed skeletons" of the original networks). More variants are considered than for relative error or Spearman correlation because the unweighted versions of those (their "Signs only" variants) are not as interesting as they are for Jaccard similarity.

8.5.4.4   DeltaCon Distance

First I compared the values of mixed-sign unweighted DeltaCon distance and mixed-sign weighted DeltaCon distance with the values resulting from considering only (i) the presence/absence of edges or (ii) the magnitudes of edges. I then compared the mixed-sign unweighted DeltaCon distance values with the DeltaCon distance values for the subnetworks corresponding to (i) the predominant sign edges only or (ii) the non-predominant sign edges only, and then analogously also for the mixed-sign weighted DeltaCon distance.

"Presence/Absence Only" refers to the DeltaCon distance of the "unsigned skeletons" $\mathcal{S}(\mathcal{G}_1), \mathcal{S}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$, analogous to Jaccard similarity. "Magnitudes Only" refers to the DeltaCon distance of the "magnitude skeletons" $\mathcal{M}(\mathcal{G}_1), \mathcal{M}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. Variants called "Signs Only" elsewhere effectively correspond to the mixed-sign *unweighted* DeltaCon distance, which by definition equals the mixed-sign weighted DeltaCon distance of the "signed skeletons" $\mathcal{S}^{\pm}(\mathcal{G}_1), \mathcal{S}^{\pm}(\mathcal{G}_2)$ of the original networks $\mathcal{G}_1, \mathcal{G}_2$. (See section 8.3.3 for skeleton definitions.)

8.6   Results

Section 8.6.1 discusses the behavior of variants of relative error for all three attacks. Section 8.6.2 discusses the behavior of variants of Spearman correlation for all three attacks. Section 8.6.3 discusses the behavior of variants of Jaccard similarity for all three attacks. Section 8.6.4 discusses the behavior of variants of DeltaCon distance for all three attacks.

8.6.1   Relative Error

Section 8.6.1.1 overviews the results, explaining which variants were sensitive to which attacks. Section 8.6.1.2 explains the insight gleaned from comparing with the signs only and magnitudes only variants. Section 8.6.1.3 explains why and when the signs only variant may or may not be useful. Section 8.6.1.4 explains the indirect evidence for the convex combination property seen in the results. (Note that nevertheless the convex combination property is proven and doesn't actually require empirical evidence to be substantiated.)

8.6.1.1   Sensitivity to Attacks

For all three attacks shown in figures 8.7b, 8.7a, 8.7c, we see that the relative error is always sensitive to the major changes that occur. In contrast, the relative error of the signed skeletons of the networks is completely oblivious to the magnitude swap attack in figure 8.7a, while the relative errors of the magnitude skeletons of the networks is completely oblivious to the sign flip attack in figure 8.7c.

Moreover, even for the sign flip attack in figure 8.7c when only signs are changed, the wider flare at the top of the distribution for relative error than for "signs only" seems to indicate greater sensitivity to the attack. Similarly for the shift attack in figure 8.7b, the wider flare
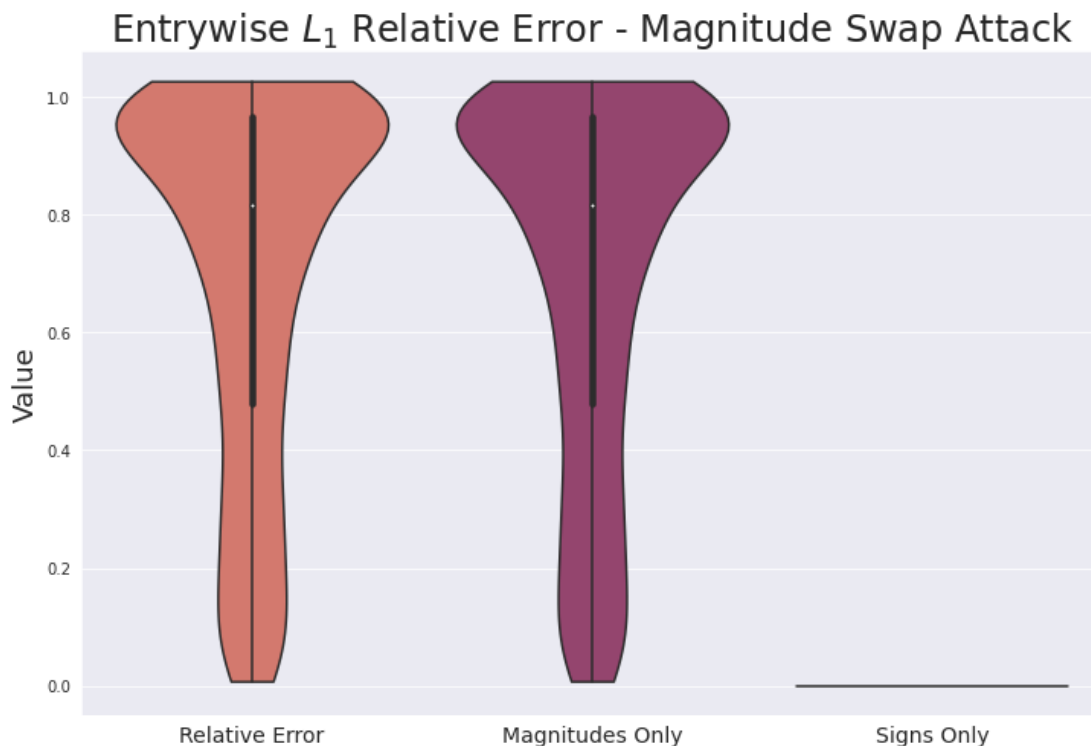
(a)

Figure 8.7: Responses of relative error variants to adversarial attacks.

at the top of the distribution for relative error compared to signs only suggests that relative error is able to combine both sign and magnitude information to be more sensitive to attacks than considering either magnitudes only or signs only. Despite the changes in magnitude the accompany the shift attack, in figure 8.7b we see that considering magnitudes only still leads to substantially less sensitivity.

8.6.1.2 Sign Information and Magnitude Information

Observe also how in figures 8.7b, 8.7a, 8.7c, the results for the relative error are not even weakly intermediate between the relative errors when considering either the magnitudes only or the signs only. In other words, the "whole" of using the relative error and considering both the sign and magnitude information in a mixed-sign network is "greater than the sum of the parts" of using either the sign information or the magnitude information alone.

8.6.1.3 Potential Usefulness for Signs Only Variant

Note that, as long as the attack does affect the signs (the shift attack in figure 8.7b and the sign flip attack in 8.7c), the signs only variant is actually fairly qualitatively similar to the
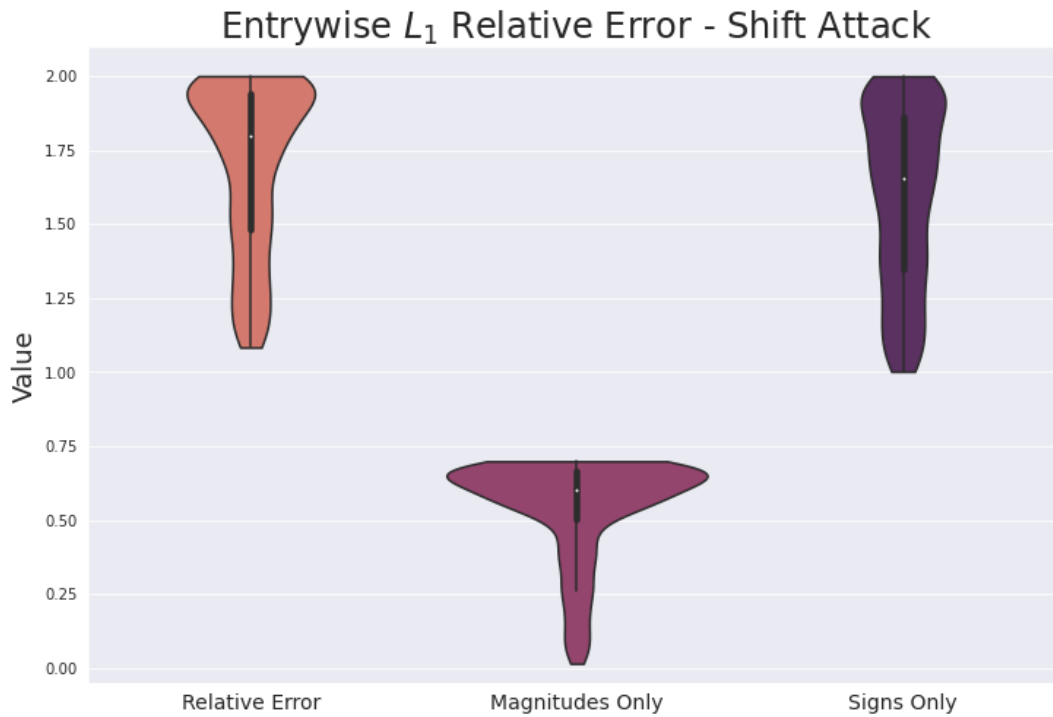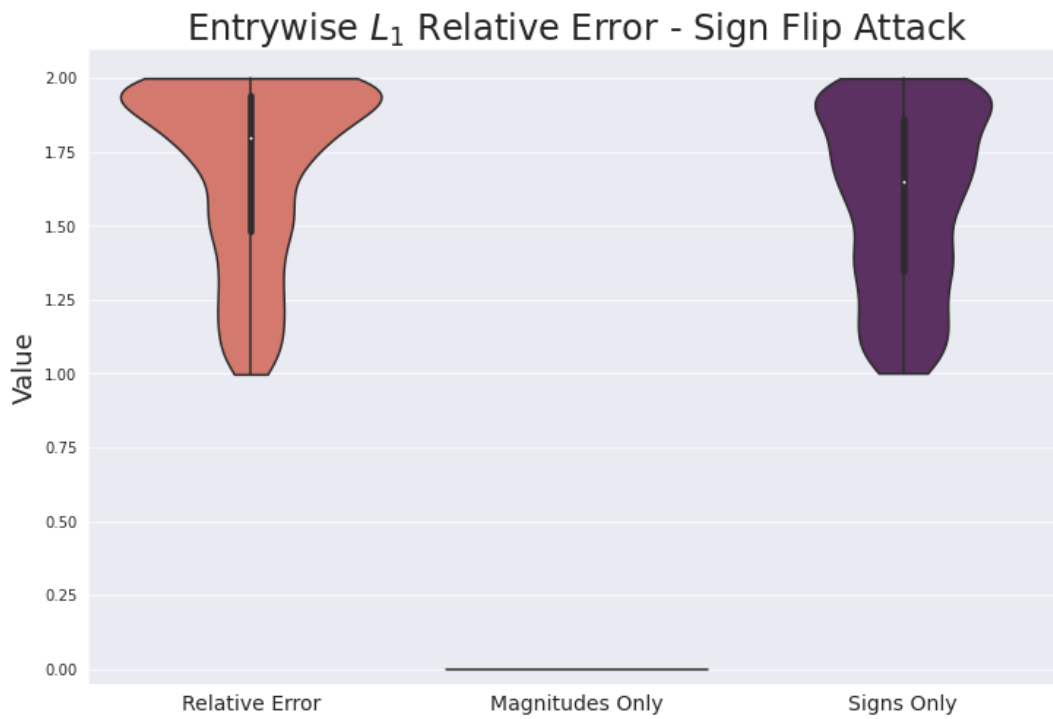
Entrywise L_1 Relative Error - Shift Attack

(b)

Entrywise L_1 Relative Error - Sign Flip Attack

(c)

Figure 8.7: Responses of relative error variants to adversarial attacks.
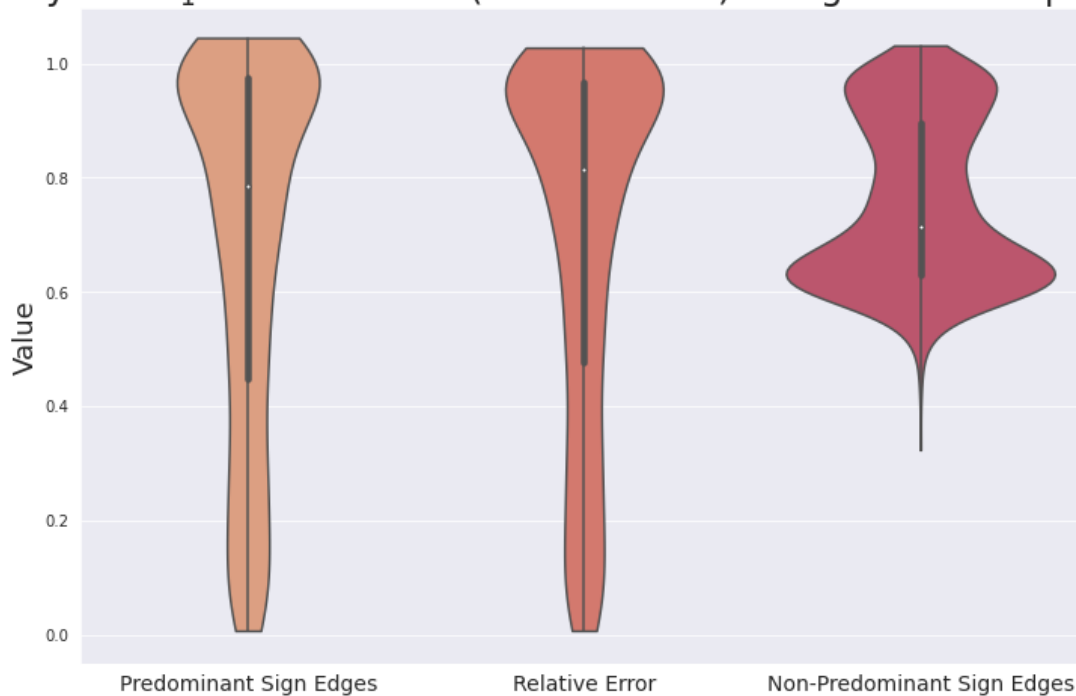
(a)

Figure 8.8: Relative error combines values for positive and negative parts.

relative error. This is perhaps to be expected at least inasmuch as the signs only variant does technically satisfy the double penalization principle as well. However this may also be largely an artifact of the magnitudes being bounded to be no greater than 1 (while the "magnitudes" of the signed skeletons, on which the signs only variant acts, are 1).

8.6.1.4 Effects of Convex Combination Decomposition Property

Subnetworks were normalized before computing the relative error (cf. sections 8.5.4.1 and 8.4.1.4). This means the interpretation as a convex combination as in equation (8.4.4) does not apply. Nevertheless, we still do see in figures 8.8b, 8.8a, and 8.8c results for the relative error which are weakly intermediate between the relative errors for the predominant sign edges and the non-predominant sign edges. Note that for the shift attack in figure 8.8b and the sign flip attack in figure 8.8c, the fact that the relative errors are all exactly 1 is because the corresponding subnetwork has 0 edges in the attacked versions of the networks.
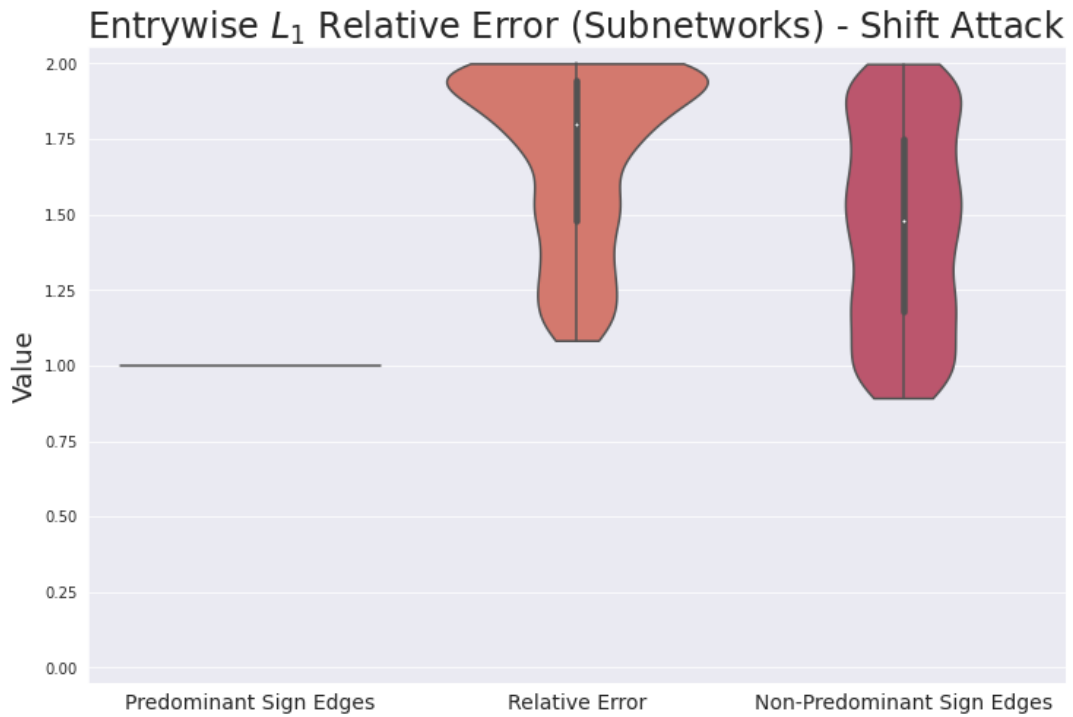
(b)
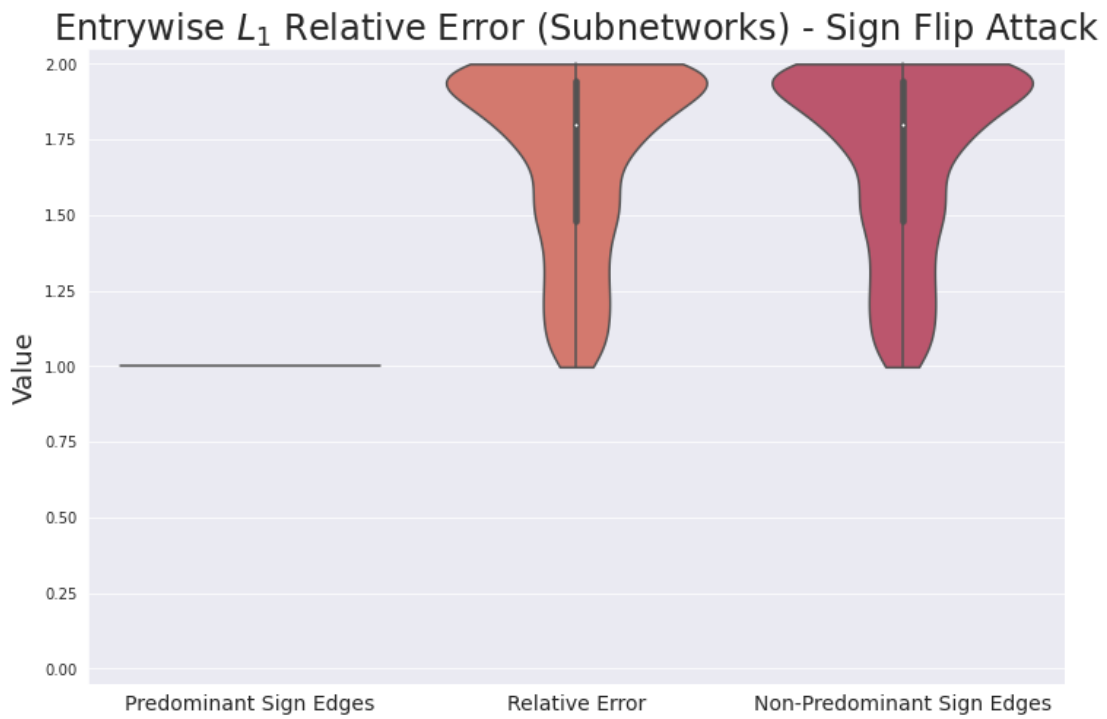
(c)

Figure 8.8: Relative error combines values for positive and negative parts.
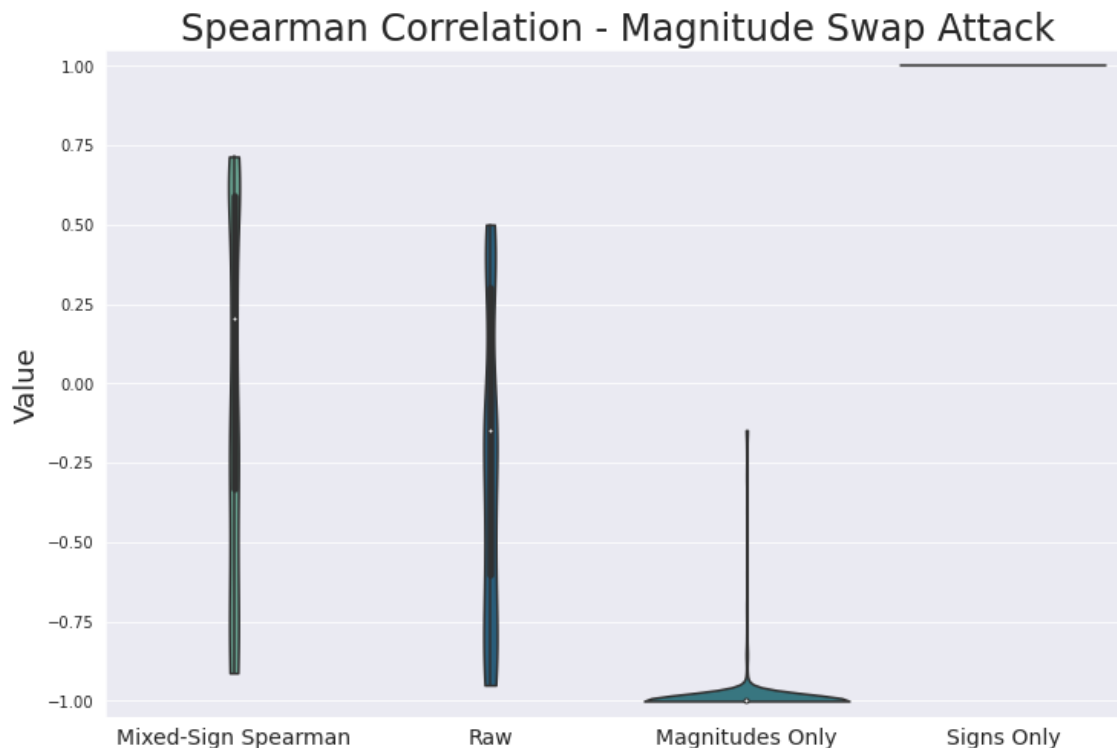
(a)

Figure 8.9: Responses of Spearman correlation variants to adversarial attacks.

8.6.2 Spearman Correlation

Section 8.6.2.1 overviews which variants were sensitive to which attacks. Section 8.6.2.2 explains the evidence from the results that the mixed-sign Spearman not only directly combines the information from the positive and negative edges, but also achieves performance superior to that from directly combining the sign and magnitude information.

8.6.2.1 Sensitivity to Attacks

For all three attacks shown in figures 8.9b, 8.9a, 8.9c, we see that the mixed-sign Spearman correlation is always sensitive to the major changes that occur.

For the shift attack, we see from figure 8.9b that the raw Spearman correlation is completely oblivious to the attack, while the signs only and magnitudes only variants are unable to detect any similarity to the original underlying network. The mixed-sign Spearman correlation is the only method that is both sensitive to the attack and able to ascertain some similarity to the original underlying network.

For the magnitude swap attack, we see from figure 8.9a that the signs only variant is (unsurprisingly) completely oblivious, while the magnitudes only variant is (unsurprisingly)
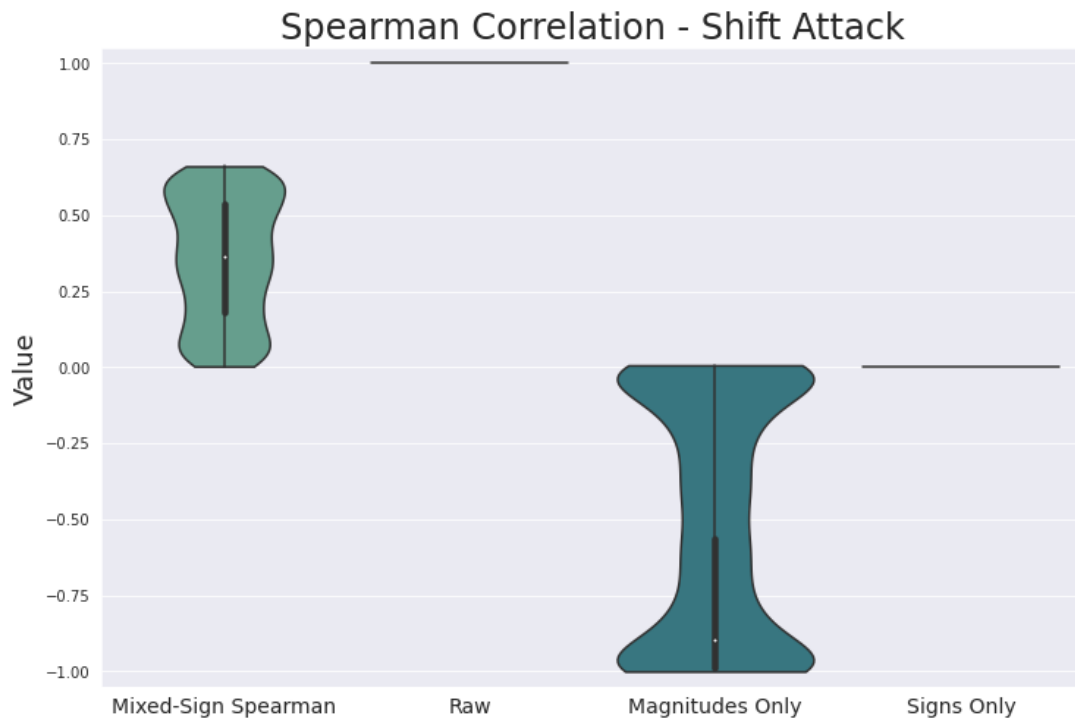
(b)
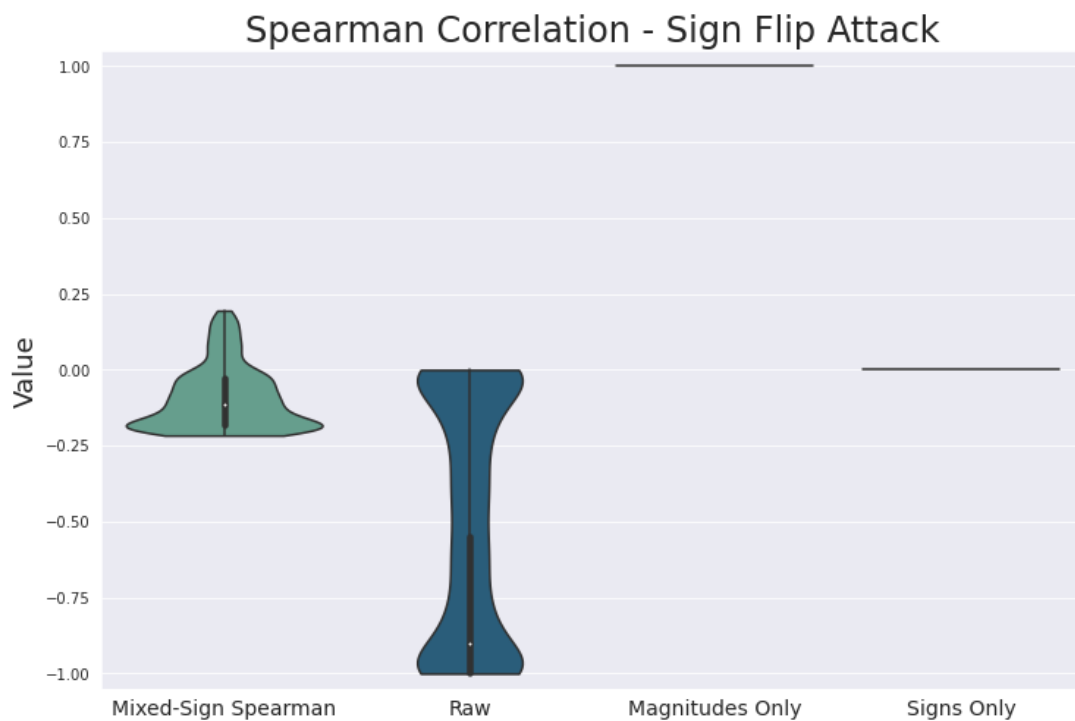
(c)

Figure 8.9: Responses of Spearman correlation variants to adversarial attacks.

completely unable to detect any similarity to the original underlying network. Both the mixed-sign and raw Spearman correlations report an amount of similarity depending on the particular structure of the original network, although the mixed-sign Spearman seems better able to detect similarities and has a wider spread.

For the sign flip attack, we see from figure 8.9c that the magnitudes only variant is (unsurprisingly) completely oblivious, while the signs only is unable to provide any useful information (because the sign vector for the attacked network is constant). Meanwhile the raw Spearman correlation is never able to detect any similarities to the original underlying network (even though all of the magnitudes and non-predominant sign edges were left unaffected), but the mixed-sign Spearman correlation reports values depending on the particular structure of the original network, correctly reporting some amount of similarity in cases when the non-predominant sign edges constitute a large minority.

8.6.2.2 Combining Information

The mixed-sign Spearman correlation is the only variant that is not oblivious to any of the attacks, and which always reports values depending on the particular structure of the underlying network. We also see from figures 8.10b, 8.10a, and 8.10c that the mixed-sign Spearman reports values intermediate between those for the predominant sign and non-predominant sign subnetworks, corresponding to the subconvex combination property mentioned in section 8.4.2.6 and demonstrating double penalization. The mixed-sign Spearman directly combines information from the positive and negative edges. In contrast, the mixed-sign Spearman can be seen to *not* be intermediate between the magnitudes only and signs only variants in figures 8.9b, 8.9a, 8.9c, demonstrating that the "whole" of the magnitude and sign information is greater than the "sum of the parts" of the magnitudes only and signs only information.

8.6.3 Jaccard Similarity

Section 8.6.3.1 explains how the standard notion of Jaccard similarity is completely useless for all three attacks. Section 8.6.3.2 explains how the results appear to indicate greater importance for sensitivity to sign information than for sensitivity to magnitude information. Section 8.6.3.3 explains how the mixed-sign weighted Jaccard similarity is more sensitive than would follow from directly combining the sign and magnitude information. Section 8.6.3.4 explains how the convex combination property satisfied by both the mixed-sign unweighted and mixed-sign weighted Jaccard similarity manifests itself in the results.

8.6.3.1 Presence/Absence Only is Always Completely Oblivious

For all three attacks shown in figures 8.11b, 8.11a, and 8.11c, we see that the presence/absence only variant is completely oblivious (and thus completely useless). This is important, since using Jaccard similarity of presence/absence of edges might have been the default choice of some researchers. We see from this that failing to respect the richer structure of networks
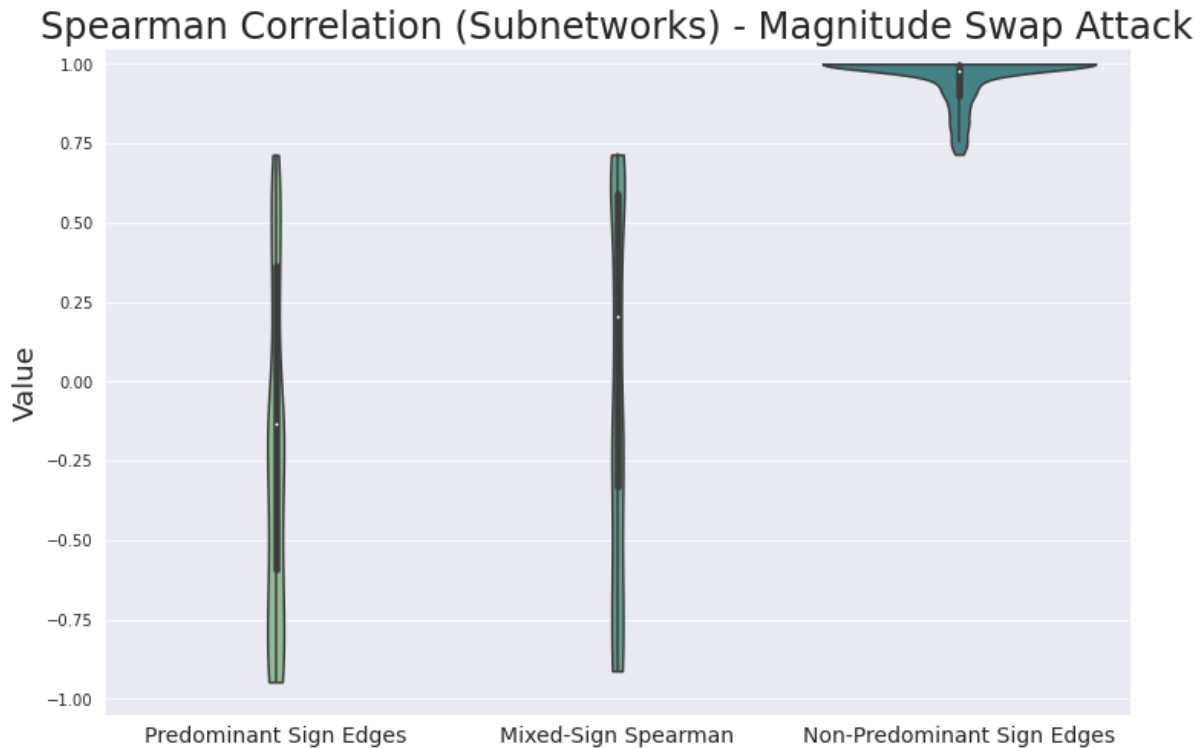
(a)

Figure 8.10: Spearman correlation combines values for positive and negative parts.

with mixed-sign edge weights, and attempting to blindly apply standard methods applicable for less structured networks, can lead to drastically misleading conclusions.

8.6.3.2 Sign Information Appears More Important than Magnitude Information

From figure 8.11c, we see that the magnitudes only variant of Jaccard similarity is completely oblivious to the sign flip attack, as expected since the attack does not change the sign of any edges. For the two attacks shown in figures 8.11b and 8.11a which do affect the magnitudes of the edges, we do see some sensitivity on the part of the magnitudes only variant. Moreover, since the magnitudes are the only aspect of the network that changed for the magnitude swap attack, it makes sense that we see in figure 8.11a that the magnitudes only variant and the mixed-sign weighted Jaccard similarity give the same results, while the mixed-sign unweighted Jaccard similarity is completely oblivious.

However, the magnitudes only variant is surprisingly insensitive to the shift attack as shown in figure 8.11b. While it is not necessarily surprising that the mixed-sign weighted Jaccard similarity is more sensitive, it is even the case that the lower range of values of the magnitudes only variant overlaps with the upper range of values of the mixed-sign un-
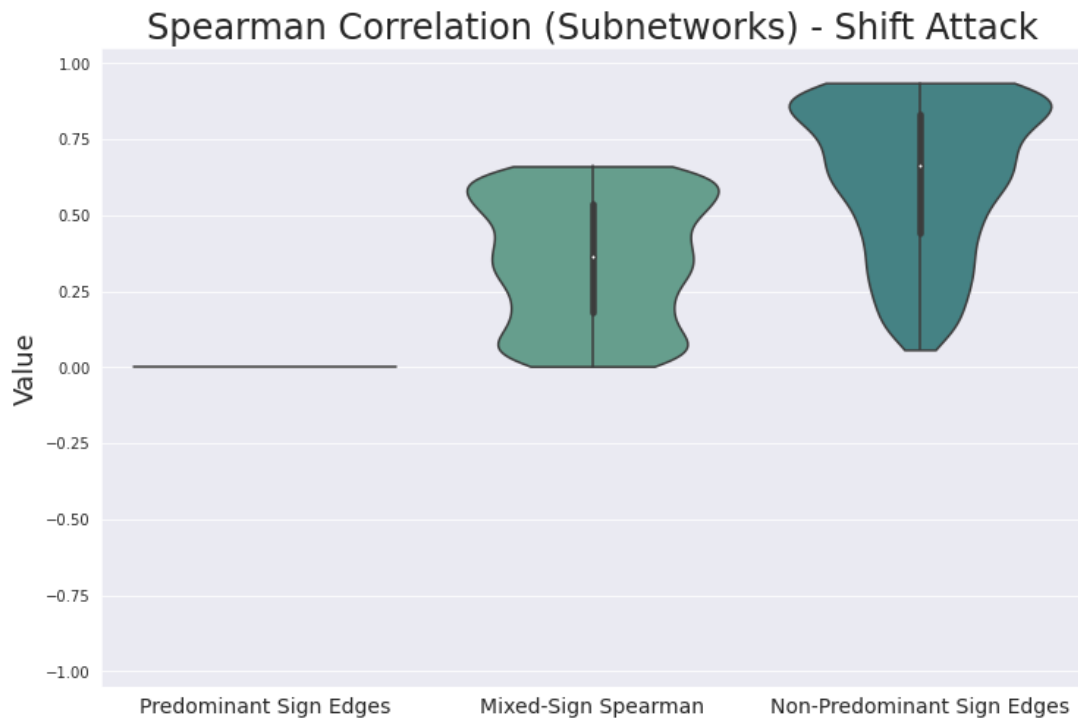
(b)
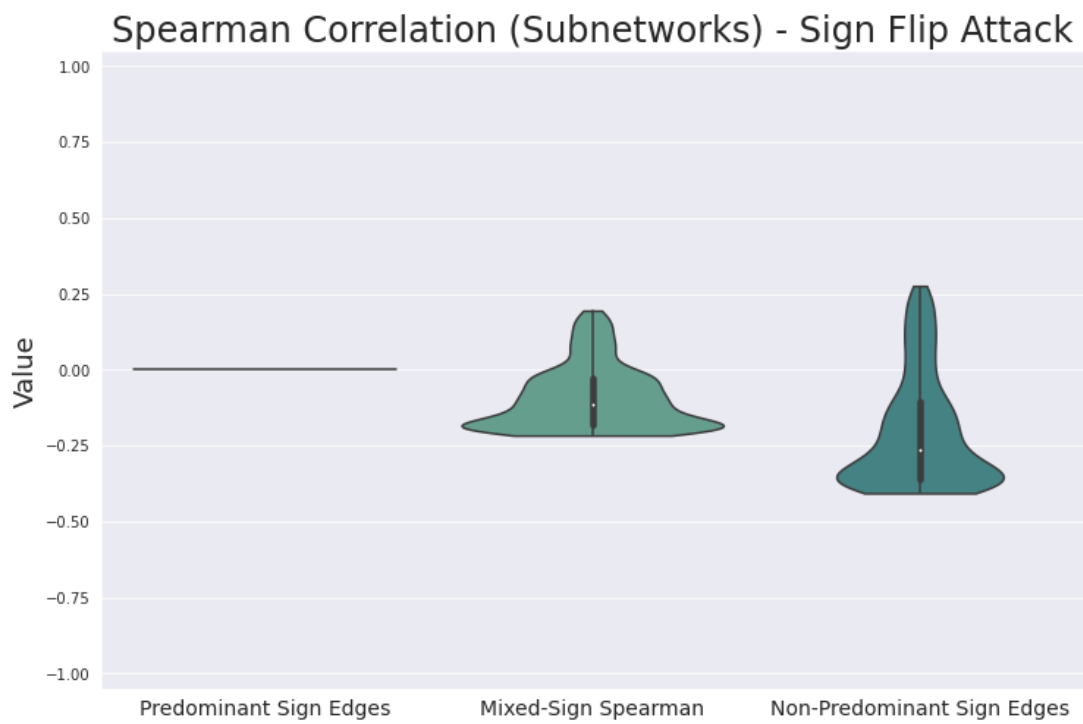
(c)

Figure 8.10: Spearman correlation combines values for positive and negative parts.
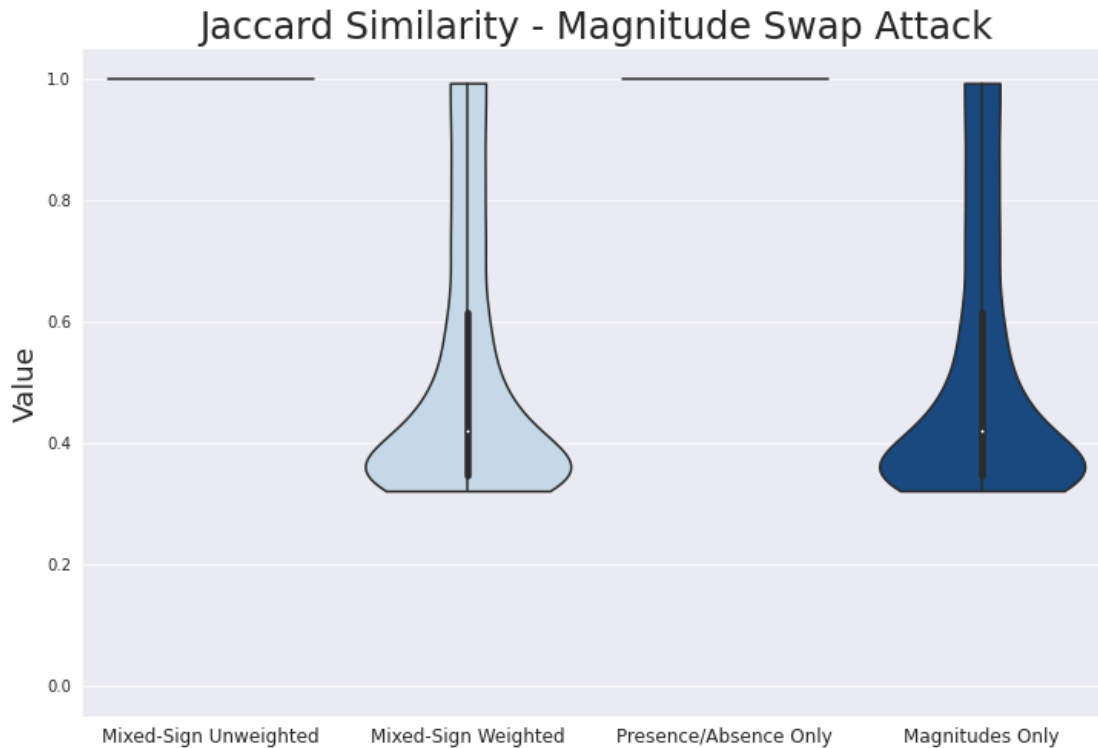
(a)

Figure 8.11: Responses of Jaccard similarity variants to adversarial attacks.

weighted variant. This is probably because, even though the mixed-sign unweighted Jaccard similarity technically only looks at the signs of the edges, it nevertheless does obey the double penalization principle in doing so. In contrast the magnitudes only variant does not obey the double penalization principle and largely has "the wool pulled over its eyes" when it comes to the shift attack, even though it is not completely oblivious.

This seems to suggest that, although signs only variants will never be able to detect attacks that only affect magnitudes, they can still be substantially useful in general as long as they obey the double penalization principle. Since the double penalization principle requires sensitivity to signs, but not necessarily to magnitudes, this seems to suggest that sensitivity to signs is more important to sensitivity to magnitudes, *as long as the double penalization principle is satisfied*. This is important to note because blind application of default methods for unsigned networks will generally[19] be completely insensitive to signs, much less obey the double penalization principle.

---

[19]Entrywise L_1 relative error is of course a notable exception.
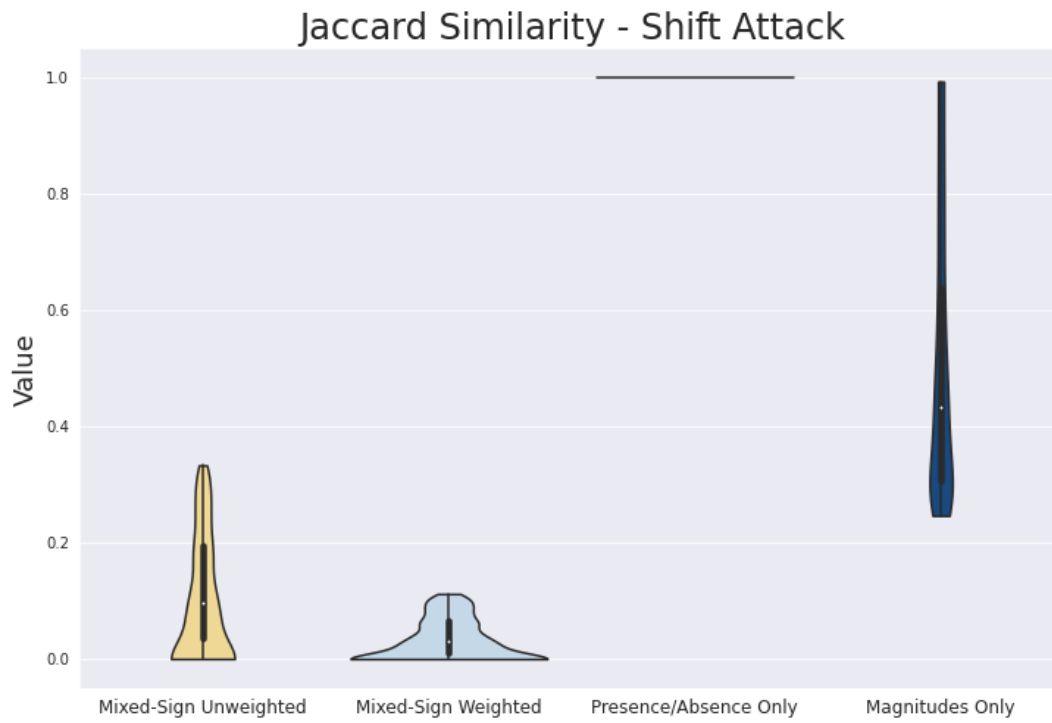
(b)
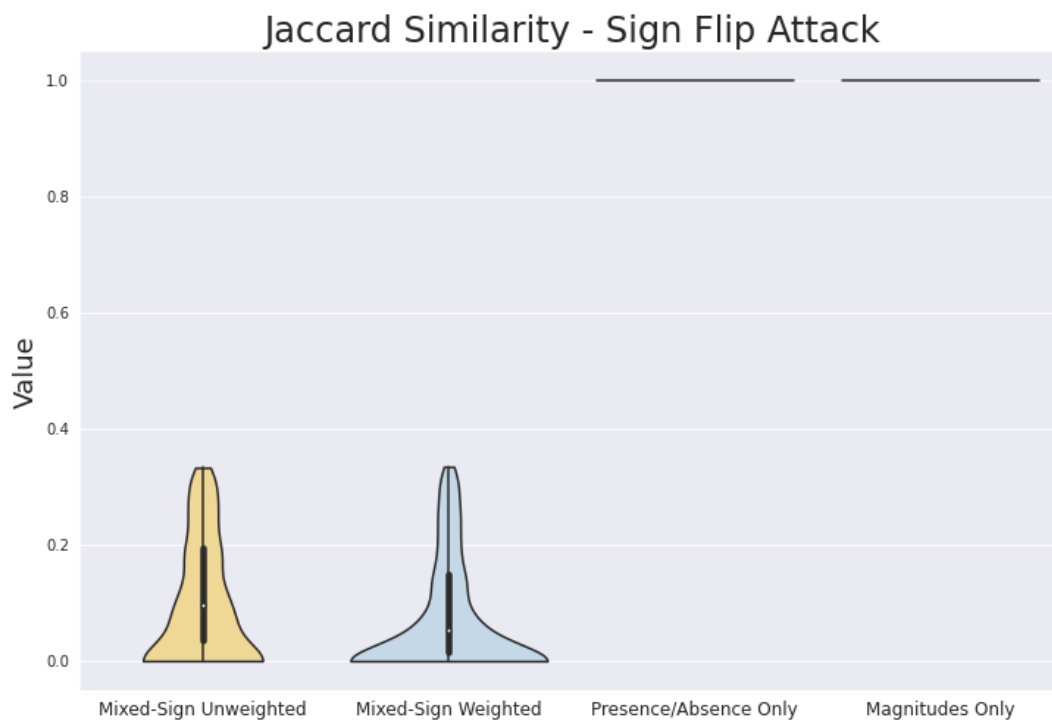
(c)

Figure 8.11: Responses of Jaccard similarity variants to adversarial attacks.

8.6.3.3 Combining Sign and Magnitude Information

We see in figure 8.11c that the distribution of values for the mixed-sign unweighted and mixed-sign weighted Jaccard similarities are most similar for the sign flip attack. This makes sense intuitively because that is the only attack that leaves the magnitudes unaffected. However the wider flare at the bottom of the distribution of the mixed-sign weighted Jaccard similarities in figure 8.11c suggests that it is even more sensitive to this attack than the mixed-sign unweighted Jaccard similarity, in spite of the attack leaving magnitudes unaffected. This would seem to follow because the distribution of magnitudes is not constant for any of the networks. Thus again we see an instance where the "whole" of being sensitive to both sign and magnitude information is greater than the "sum of the parts" of the magnitude and sign information, even in instances where seemingly only one of those two types of information would appear to be directly relevant.

8.6.3.4 Effects of Convex Combination Decomposition Property

Figures 8.12b, 8.12a, and 8.12c correctly suggest how the mixed-sign unweighted Jaccard similarity is a convex combination of the unweighted Jaccard similarities for the positive and negative subnetworks.

Similarly, figures 8.13b, 8.13a, and 8.13c also correctly suggest how the mixed-sign weighted Jaccard similarity is a convex combination of the weighted Jaccard similarities for the positive and negative subnetworks.

Thus both mixed-sign versions of Jaccard similarity incorporate information from both subnetworks in a reasonable fashion.

8.6.4 DeltaCon Distance

Only the mixed-sign weighted DeltaCon distance is sensitive to all three attacks, with mixed-sign unweighted and magnitudes only both completely oblivious to at least one attack, and presence/absence only completely oblivious to all three attacks. Mixed-sign weighted is always more sensitive than magnitudes only. Mixed-sign weighted appears slightly less sensitive than mixed-sign unweighted for these attacks, but that may be an artifact.

Section 8.6.4.1 explains the importance of presence/absence only being completely oblivious. Section 8.6.4.2 explains an important caveat when comparing the results for mixed-sign unweighted and mixed-sign weighted DeltaCon distance. Section 8.6.4.3 explains how mixed-sign unweighted may actually still be useful sometimes, despite disregarding information about edge magnitudes. Sections 8.6.4.4, 8.6.4.5, and 8.6.4.6 give "sanity checks" that demonstrate the results are sensible and in line with expectations. Section 8.6.4.7 explains a subtle observation that argues strongly for the importance of the double penalization principle. Section 8.6.4.8 generalizes the observations from section 8.6.4.7.
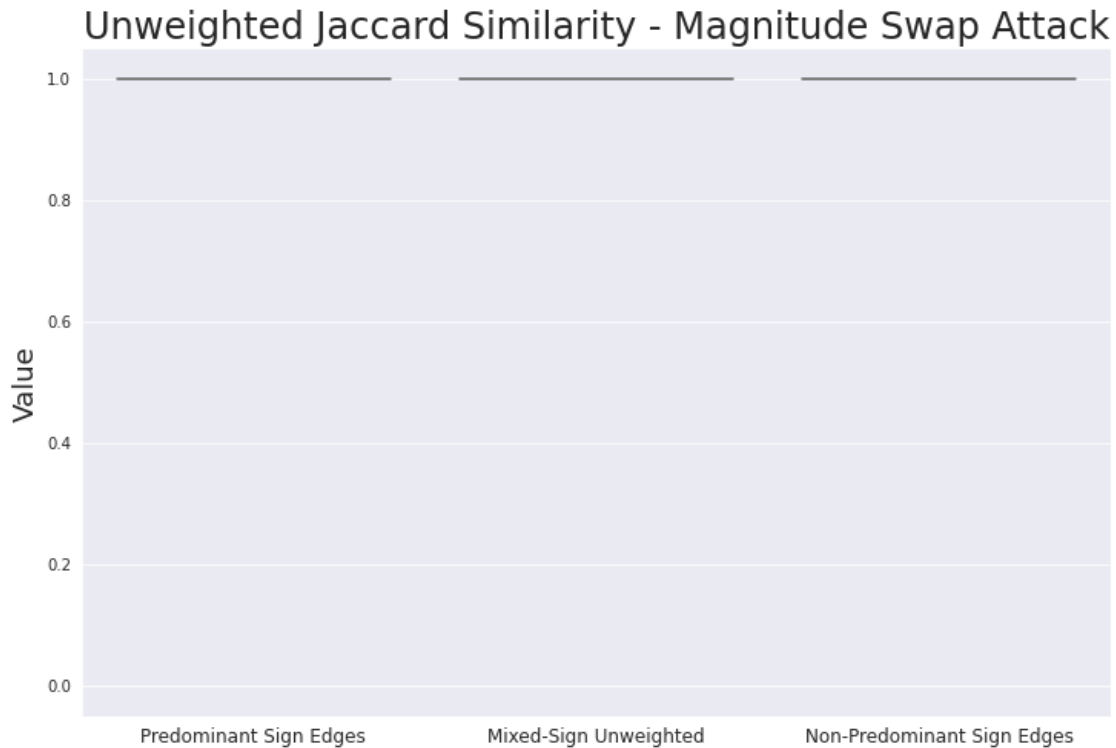
(a)

Figure 8.12: Unweighted Jaccard similarity combines values for positive and negative parts.

8.6.4.1 Presence/Absence Only is Always Completely Oblivious

Just as was the case for Jaccard similarity, the presence/absence only variant of DeltaCon distance can be seen from figures 8.14b, 8.14a, and 8.14c to be completely oblivious to all three attacks. This is again important because it shows that blindly seeking to directly transfer similar methods for less general networks to this context can lead to misleading conclusions. For example the proofs in the FaBP paper [KKK+11] seem to assume an unweighted network, as do some code implementations for DeltaCon that can be found.

8.6.4.2 Mixed-Sign Weighted May Actually Not Be Less Sensitive than Mixed-Sign Unweighted

That the values of the mixed-sign unweighted DeltaCon distance trend slightly larger than those for the mixed-sign weighted DeltaCon distance in figures 8.14b and 8.14c appears to be an artifact of the magnitudes all being less than or equal to 1. (Thus the "magnitudes" of the signed skeleton are in general greater, and mixed-sign unweighted DeltaCon is equivalent to mixed-sign weighted DeltaCon applied to the signed skeletons of the original networks.)

(b)
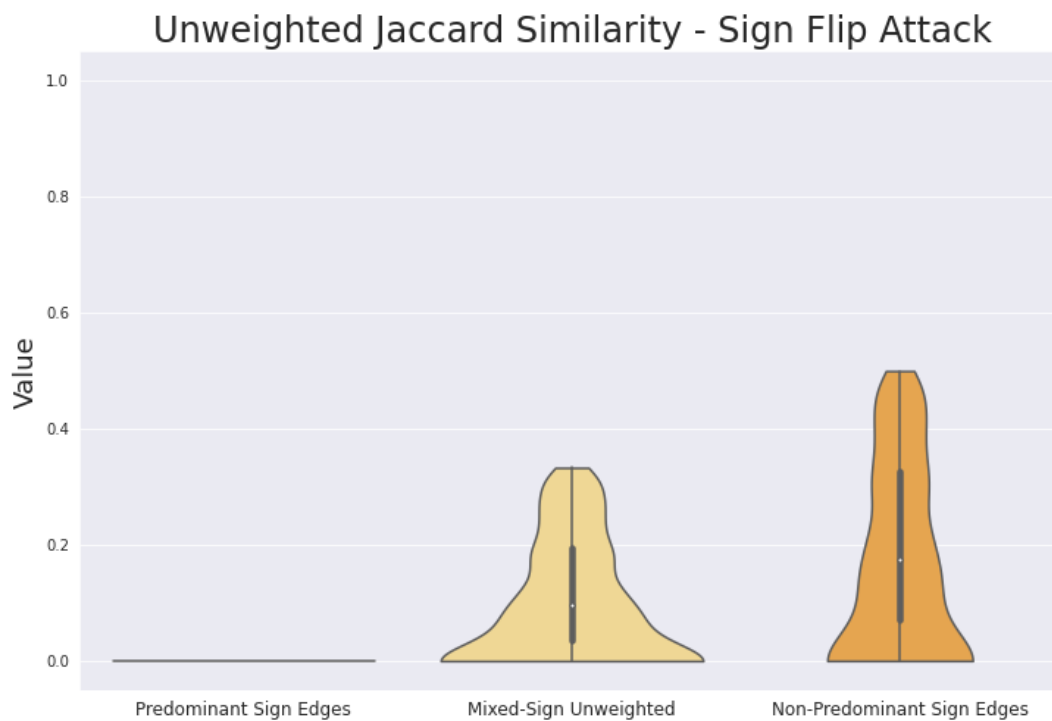
(c)

Figure 8.12: Unweighted Jaccard similarity combines values for positive and negative parts.
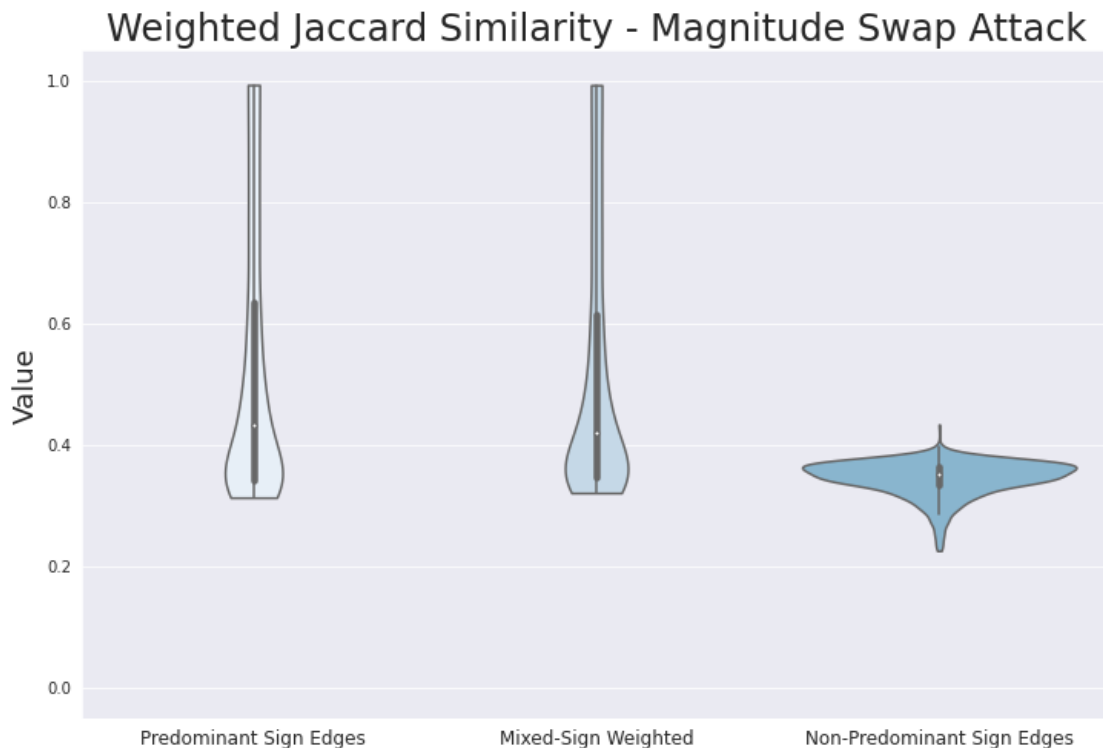
(a)

Figure 8.13: Weighted Jaccard similarity combines values for positive and negative parts.

8.6.4.3 Mixed-Sign Unweighted Performs Better than Magnitudes Only

We again see from figure 8.14b that, as long as the double penalization principle is satisfied by the "signs only" variant, sensitivity to signs (mixed-sign unweighted DeltaCon) may be more important than sensitivity to magnitudes for attacks which affect both. The mixed-sign unweighted DeltaCon distance is more sensitive to the shift attack than the magnitudes only DeltaCon distance. One could argue that this is an artifact of the magnitudes being less than or equal to 1, but that seems unlikely because the distribution of values for the mixed-sign weighted DeltaCon distance for the shift attack is much more similar to that of the mixed-sign unweighted DeltaCon distance than of the magnitudes only variant.

8.6.4.4 Sensitivity to Changes in Sign Only

Unsurprisingly only the variants that are sensitive to sign, namely the mixed-sign unweighted and mixed-sign weighted DeltaCon distances, can detect that sign flip attack from figure 8.14c at all. The distribution of values of the mixed-sign weighted DeltaCon distance may be more similar to that of the mixed-sign unweighted DeltaCon distance for the sign flip attack (figure 8.14c) than for the shift attack (figure 8.14b) since the latter also involves a
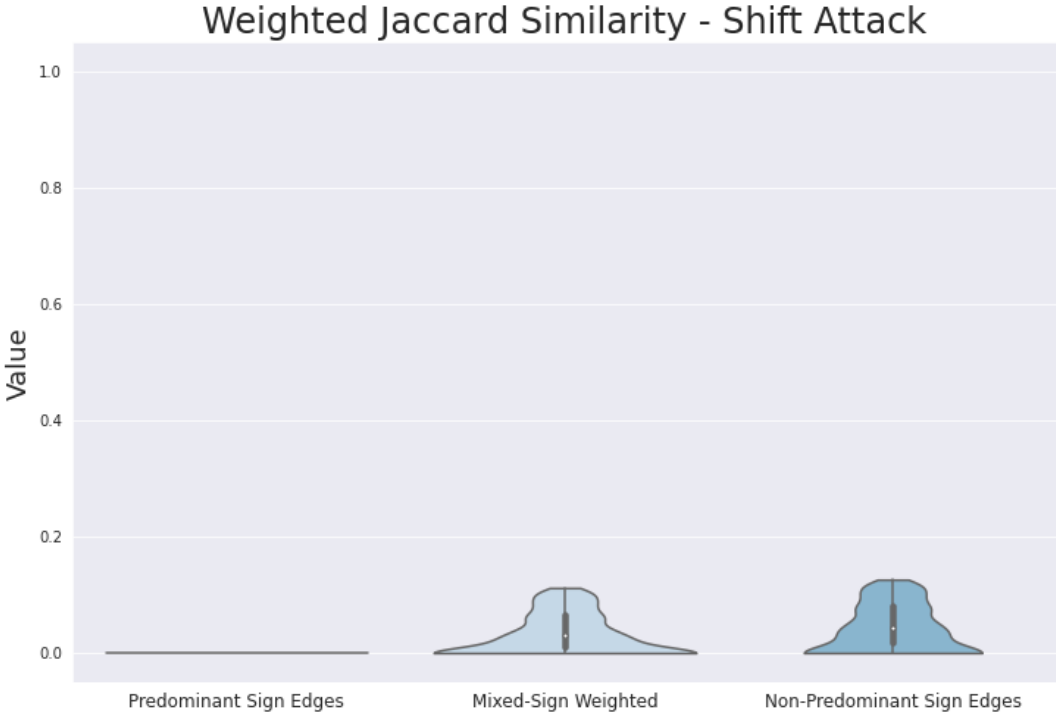
(b)
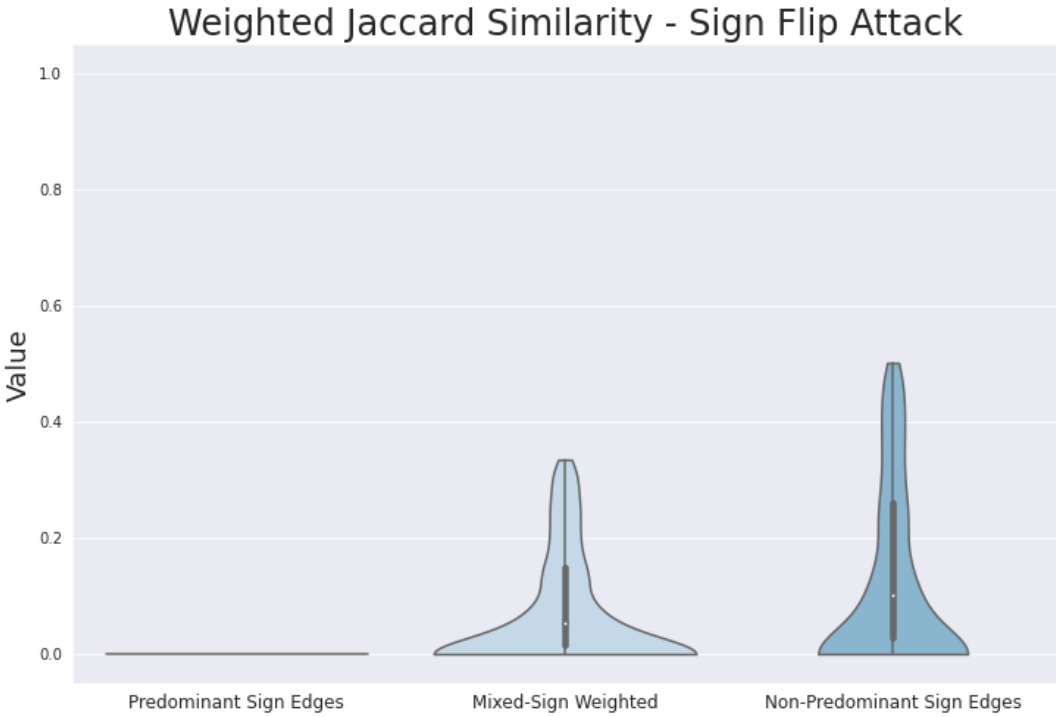
(c)

Figure 8.13: Weighted Jaccard similarity combines values for positive and negative parts.
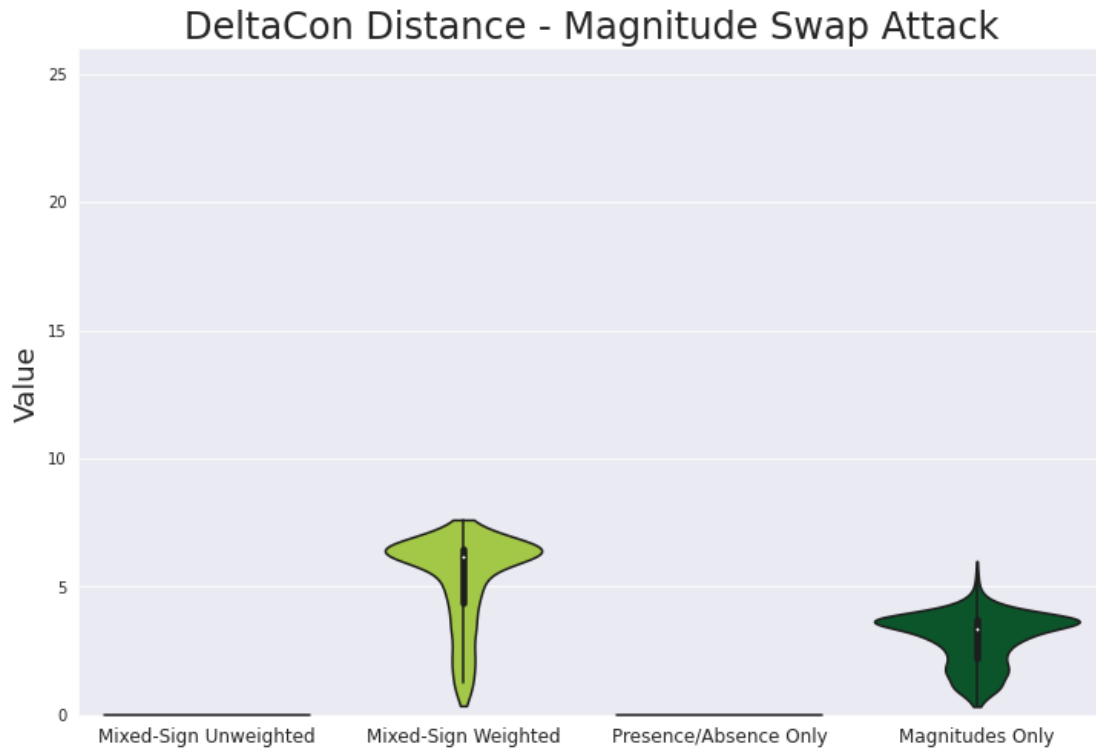
(a)

Figure 8.14: Responses of DeltaCon variants to adversarial attacks.

change of magnitudes.

8.6.4.5 Sensitivity to Changes in Magnitude Only

Unsurprisingly, only the mixed-sign weighted DeltaCon distance and the magnitudes only variant are sensitive to the magnitude swap attack from figure 8.14a (which only affects magnitude). Not only presence/absence only, but also mixed-sign unweighted DeltaCon, are completely oblivious.

8.6.4.6 Mixed-Sign Weighted Picks up Changes in Magnitude

Note that the distributions of values look the same between figures 8.15b and 8.15c, reflecting the fact that the shift and sign flip attacks have the same effect on the signs of the edges, and that the mixed-sign unweighted DeltaCon is only sensitive to changes in sign.

In contrast, the distributions of values look different between figures 8.16b and 8.16c, reflecting that the shift attack affects magnitudes (and the sign flip attack does not), and also demonstrating that the mixed-sign weighted version of DeltaCon is truly sensitive to changes in both sign and magnitude.
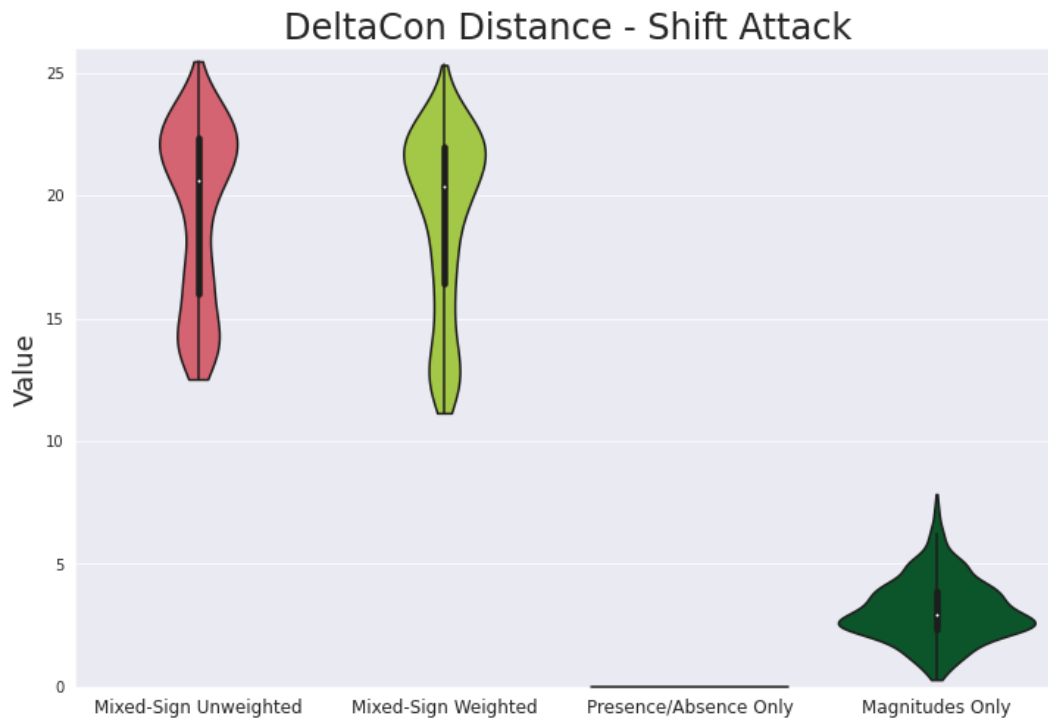
(b)
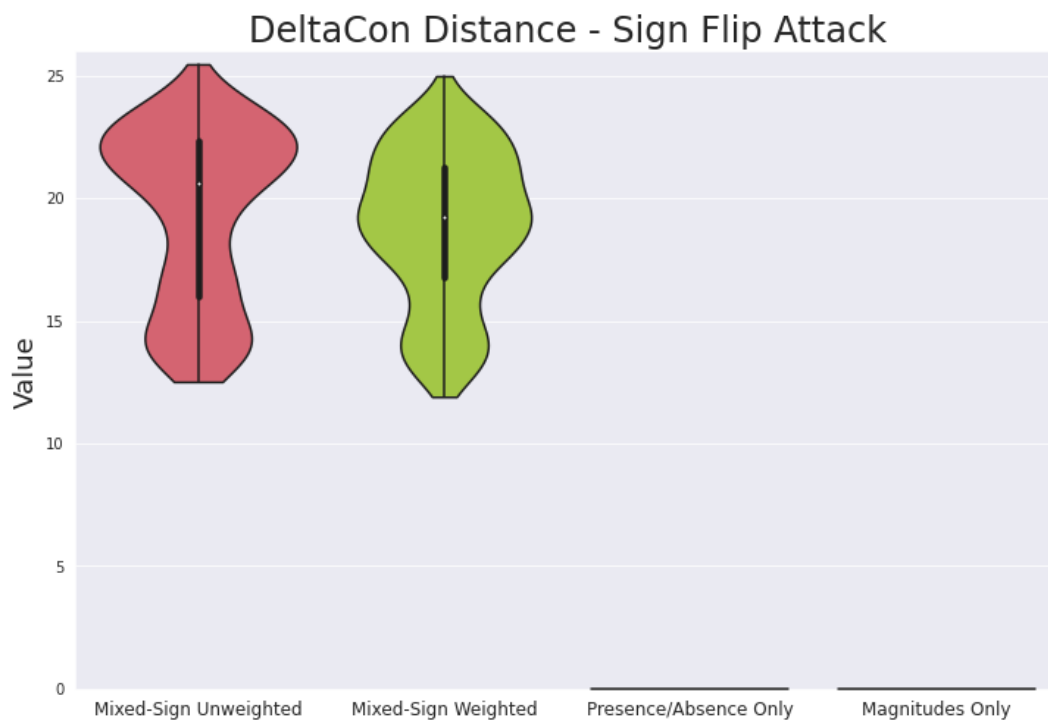
(c)

Figure 8.14: Responses of DeltaCon variants to adversarial attacks.

(a)

Figure 8.15: Unweighted DeltaCon exceeds values for positive and negative parts.

8.6.4.7 Mixed-Sign Weighted is More Sensitive than Magnitudes Only even when Only Magnitudes Have Changed

In spite of the fact that the attack only directly affects magnitudes, the mixed-sign DeltaCon (which also takes signs into account) appears to be more sensitive to the magnitude swap attack than the magnitudes only version. This may actually be a consequence of the double penalization principle being satisfied, since even though the signs agree for all edges when comparing the original network and its attacked counterpart, the signs may not necessarily agree when comparing the edges of the networks which correspond to interpreting their similarity matrices (8.4.40) as adjacency matrices of a mixed-sign network. In any case, it is another instance of the "whole" of the magnitude and sign information being greater than the "sum of the parts" of the magnitude or sign information considered separately. That this occurs is nevertheless still relatively surprising when one notes how the analogous phenomenon does not occur for entrywise L_1 relative error (figure 8.7a) nor for Jaccard similarity (figure 8.11a). That discrepancy is probably further evidence for the explanation that the phenomenon is a result of DeltaCon's consideration of "higher-order connectivities" (i.e. graph motifs consisting of more than one edge).
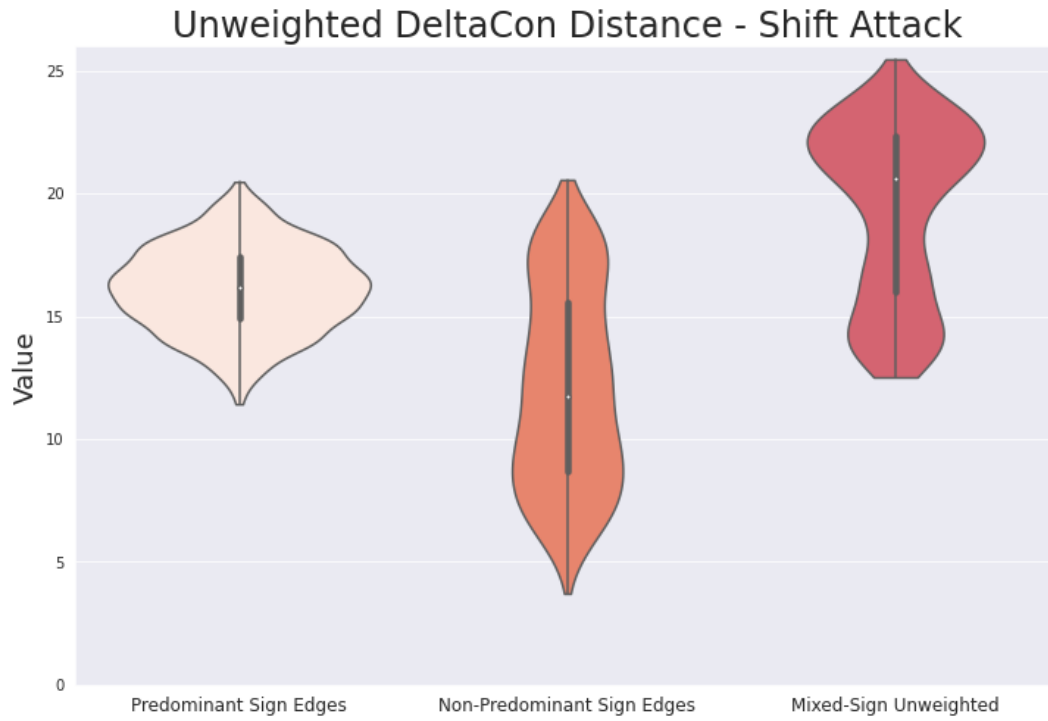
(b)
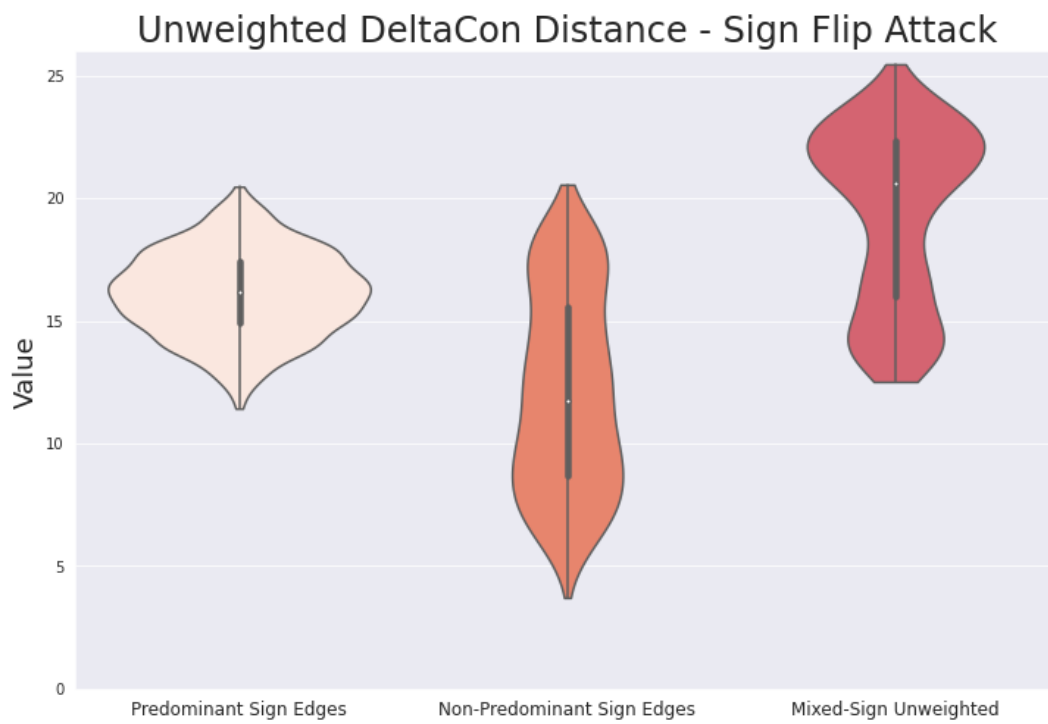
(c)

Figure 8.15: Unweighted DeltaCon exceeds values for positive and negative parts.
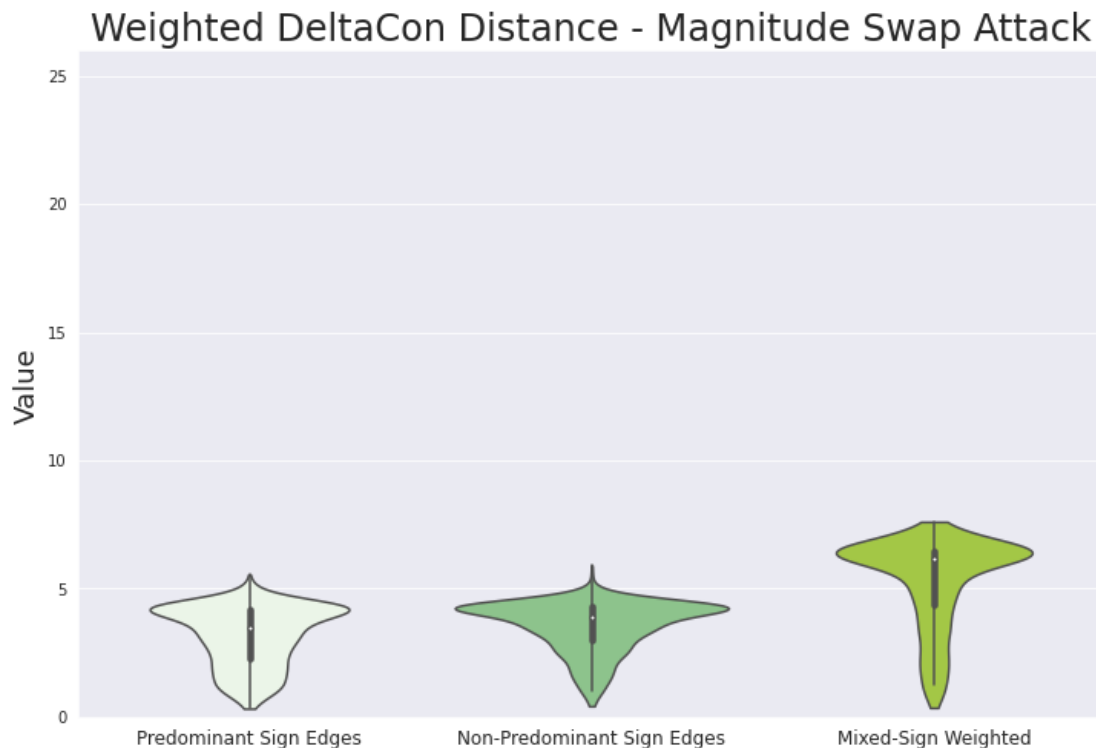
(a)

Figure 8.16: Weighted DeltaCon exceeds values for positive and negative parts.

8.6.4.8 Mixed-Sign Variants Always More Sensitive

We see from figures 8.14b, 8.14a, 8.14c that the mixed-sign variants, weighted or unweighted, are more sensitive to all three attacks than considering presence/absence only or magnitudes only. (With the exception of the mixed-sign unweighted variant for the magnitude attack.) This seems to be evidence for double penalization occurring for DeltaCon as well. Figures 8.15b and 8.15c show that the mixed-sign unweighted version is more sensitive than the changes to either the predominant or non-predominant sign subnetworks alone when the attack changes sign, which is further evidence that double penalization is occurring. Figures 8.16b, 8.16a, 8.16c also show that the mixed-sign weighted DeltaCon is more sensitive to all of the attacks than either the predominant or non-predominant sign subnetworks. This is not only because the support of the distributions is higher – their distributions are also skewed towards higher values, whereas the distributions of values for either subnetwork are more "evenly" distributed. This apparent effectiveness of the double penalization principle even for DeltaCon, which is affected by the problem that $(A^+ + A^-)^n \neq (A^+)^n + (A^-)^n$ for "higher-order connectivities" $(n \geq 2)$, suggests that the double penalization principle is still valuable even when it doesn't make sense to define a mixed-sign (dis)similarity as a convex combination of (dis)similarities for the positive and negative parts of the networks.
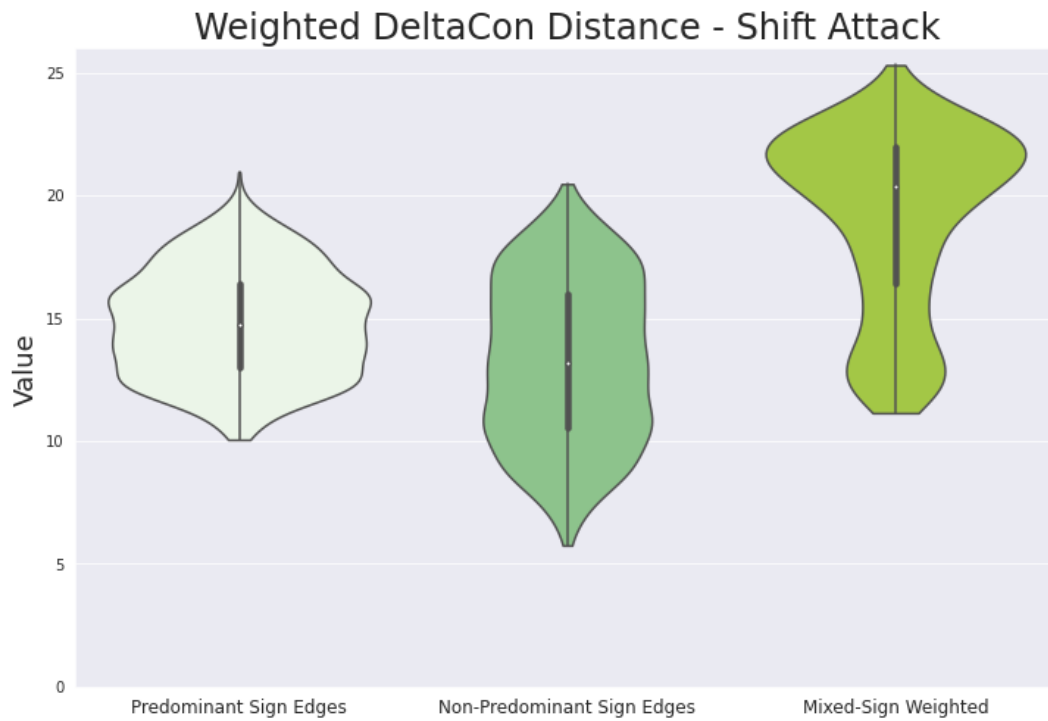
(b)
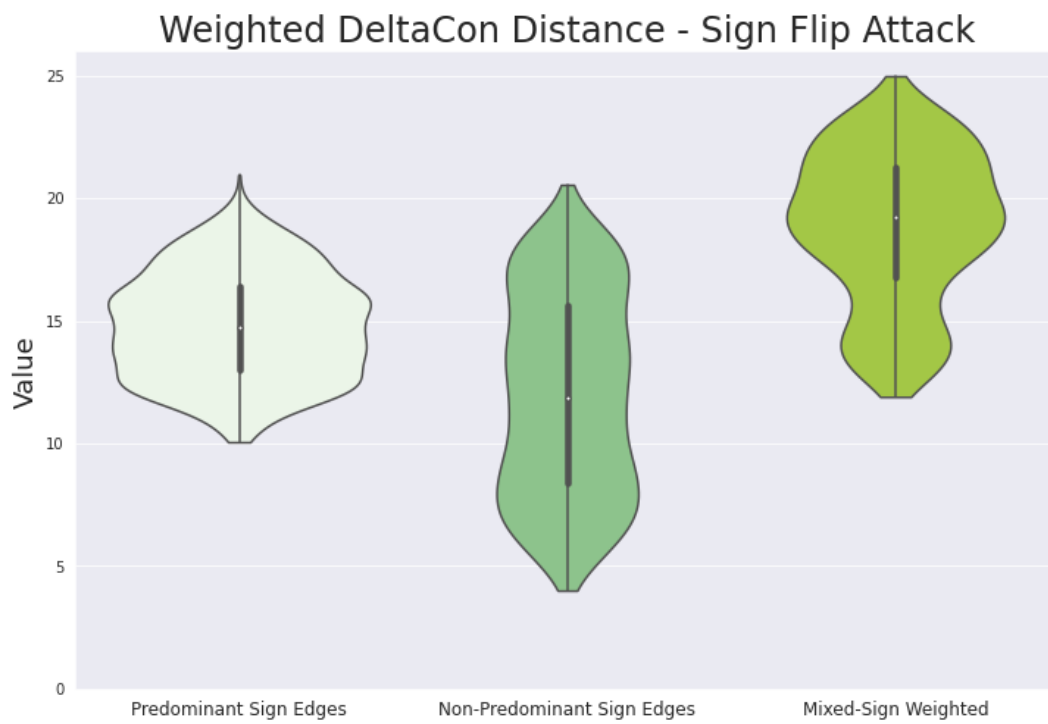
(c)

Figure 8.16: Weighted DeltaCon exceeds values for positive and negative parts.

8.7 Discussion

There are potential limitations to consider when interpreting the general applicability of the results.

Sections 8.7.1, 8.7.2, and 8.7.3 discuss the potential limitations of the chosen simulations. Section 8.7.1 discusses the number and variety of attacks for which these methods were tested. Section 8.7.2 discusses the statistical properties of the random networks on which these methods were tested. Section 8.7.3 discusses the alternatives that these methods were tested against.

Sections 8.7.4, 8.7.5, and 8.7.6 discuss the general limitations of the scope of this paper. Section 8.7.4 discusses the potential limited applicability of the convex combination decomposition property. Section 8.7.5 discusses which type of network comparison problems are addressed by this work. Section 8.7.6 discusses the limited applicability of this work for extending network comparison methods not discussed in this paper.

8.7.1 The Number of Investigated Attacks Was Small

These results only considered a limited number of attacks. The goal of this initial work was to consider attacks which did not affect the overall connectivity of the network and which affected limited aspects of the structure in obvious ways. One can think of many more changes to structure of mixed-sign networks for which we would want (dis)similarity measures to behave well. Future work can investigate whether double penalization leads to good responses for other kinds of attacks.

8.7.2 Distribution of Random Networks

The goal of this initial work was to consider a large number of networks with varying structure, such that their composite structure might be enough to encompass all "typical" networks. For example, always using a uniform distribution for the edge weights, rather than a Beta distribution with a varying hyperparameter, led to results (data not shown) that more reflected the particular nature of the uniform distribution rather than say anything about the network comparison methods themselves. For example, the spread of (dis)similarity values tended to be much smaller, and the effects of the magnitude swap attack were too predictably constrained. The chosen distribution led to mixed-sign networks with a wide range of properties.

Moreover the networks all had an expected sparsity of 50%. (Each possible edge was present or absent independently with probability 50%.) Future work might examine distributions with less uniform distributions of sparsity. Furthermore future work might want to choose magnitudes and signs more independently, and it might be worthwhile to look at magnitudes sampled from distributions which are not bounded (e.g. the exponential distribution). The Beta distributions have support limited to $(0, 1)$ and we made an edge negative if its original magnitude fell in $(0, \frac{1}{2})$ and positive otherwise (see section 8.5.1 for reference).

Also one could argue that I had mostly generated random sparse matrices rather than true "random networks". While this concern seems unfounded given the one-to-one correspondence between networks with mixed-sign edge weights and their adjacency matrices (which could correspond to any square matrix), it is nevertheless true that networks with "realistic" connectivity have adjacency matrices whose sparsity pattern is non-generic. (In other words, matrices with no 0 entries are generic under typical random matrix distributions, which correspond to "complete" networks with every possible edge present.) I did use a distribution for which networks with many missing edges were generic, but still it could be argued that this did not correspond to "state of the art" or "typical" models for defining random network topologies.

Having each edge be present or absent with probability 50% does seem to correspond to an Erdos-Renyi model with parameter $p = \frac{1}{2}$, but that still leaves open the criticism of not having considered Erdos-Renyi for other values of p (or with p as a random hyperparameter), or other models of random network topology. The case where $p = \frac{1}{2}$ corresponds to choosing all possible network topologies (unsigned skeletons) with equal probability, but nevertheless "most" such topologies may still be considered unrealistic compared to "typical real-world networks". The (dis)similarity methods chosen all adjust implicitly for sparsity somewhat, so it seems unlikely that any bias in the network topologies affects the validity of the overall conclusions. Nevertheless it would still be a valuable issue for future work to investigate.

8.7.3   Fairness of Comparisons

One could argue that the shown results might make the proposed (dis)similarity measures look better unfairly because they were compared against "strawmen" alternatives. Nevertheless, with the exception of the raw spearman and the relative error, it seems most standard (dis)similarity measures that one would use for networks with less structure cannot be applied to networks with mixed-sign edge weights. Therefore to attempt to use such standard methods, it seems conceivable that many people in practice might disregard some of extra structure of networks with mixed-sign edge weights (e.g. by considering only the magnitude skeleton or the unsigned skeleton). So I believe it is fair to compare against such methods. Even to the extent such alternatives are unrealistic, comparing with them illustrates the understanding is lost when specific features of networks with mixed-sign edge weights are ignored, and thus can be justified from a "pedagogical" perspective. Future work might try to envision and make comparisons with more "realistic" or "fair" alternatives that don't obey the double penalization principle.

8.7.4   Convex Combinations and Higher-Order Connectivity

A lot of emphasis was given to (dis)similarity measures that can be decomposed into a convex combination (weighted average) of their values for the positive edge subnetwork and the negative edge subnetwork. This emphasis may not be realistic, because such an approach does not seem to be generally applicable. (Dis)similarity measures which consider

graph motifs involving two or more edges most likely require some reference to "higher-order adjacency matrices" $\mathbf{A}^n$ describing the number of paths between two nodes with n edges. Although it is true that $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$, in general the "freshman's dream" is false, namely $\mathbf{A}^n = (\mathbf{A}^+ - \mathbf{A}^-)^n \neq (\mathbf{A}^+)^n - (\mathbf{A}^-)^n$ for $n \geq 2$. Because the analogue of the "freshman's dream" for $n = 1$ is most often used to show the convex combination decomposition property (when it exists/is valid), I do not expect analogous properties to be applicable or available for general (dis)similarity measures considering graph motifs involving two or more edges. One could attempt to define reasonable convex combinations in terms of all 2^n terms of $(\mathbf{A}^+ - \mathbf{A}^-)^n$ when expanded, but such methods are unlikely to be computationally feasible (unless their computation reduces to something with far fewer terms).

Nevertheless, I do not believe that the failure of the convex combination paradigm to generalize to arbitrary (dis)similarity measures spells trouble for the Double Penalization Principle. Methods with the convex combination decomposition property are only a special case of methods satisfying the double penalization principle. The results for both unweighted and weighted DeltaCon appear to suggest that the double penalization principle remains useful even for (dis)similarity measures that consider "higher-order connectivities". Future work considering extensions to networks with mixed-sign edge weights of other (dis)similarity measures considering graph motifs with two or more edges would be worthwhile. For example, this work only considers extending the exact version of DeltaCon, but not also the approximate (and much more scalable) version of DeltaCon defined in [KSV⁺16].

8.7.5 Node Correspondence

Using the dichotomy between "known node correspondence" and "unknown node correspondence" network comparison methods as defined in [TITP19], as mentioned before this work only considers the known node correspondence problem for networks with mixed-sign edge weights. The extent to which the double penalization principle or other ideas from this work are applicable as well to the unknown node correspondence problem for networks with mixed-sign edge weights is unclear. Future work investigating whether we need to start "from scratch" for the unknown node correspondence problem when it comes to principles for extending methods to networks with mixed-sign edge weights would be highly valuable. It seems conceivable that the ideas applicable for the known node correspondence problem might still be applicable at least indirectly, but this still needs to be investigated.

8.7.6 General Methods

This work provides no "algorithm" or systematic procedure for extending arbitrary (dis)similarity measures to the case of networks with mixed-sign edge weights. The double penalization principle and the convex composition decomposition properties are merely "desired properties" or "design specifications" (the former "essential", the latter "optional"). I provide concrete implementations of these principles for only a handful of network comparison methods, and perhaps not even necessarily the most important ones. It seems unlikely that such

a general systematic procedure could exist (or at least not any reasonable one). For example, even amongst the three methods with the convex combination decomposition property, the formulas for the convex coefficients were different each time, following from "just so" arguments specific to each method. The absence of such a general systematic procedure can be considered to limit the usefulness of this work.

8.8 Conclusion

Findings and Contributions I discussed how comparisons of ecological interactions can be recast into the statistical framework of loss functions for signed networks. I explained how avoiding unexpected behavior requires loss functions for signed networks to satisfy what I call "the double penalization principle". Starting from loss functions of unsigned networks, I derived several specific and concrete examples of loss functions for signed networks that satisfy this property. I also identified a simple property (sufficient but not necessary), the convex combination decomposition property, which implies the double penalization principle.

Practical Implications The results show that naive attempts to use pre-existing methods for comparing unsigned networks can lead to useless or even actively misleading results when applied to signed networks. The extra structure of signed networks means that they need to be compared differently. Considering changes in sign alone, or changes in magnitude alone, are usually insufficient. Methods satisfying the double-penalization principle can avoid such misleading results. This work allows us to directly address the difficulties inherent in making meaningful comparisons between signed networks, rather than ignore those difficulties.

Next Steps and Open Questions Future work should further clarify the double penalization principle. For example, what are the most general contexts in which it makes sense? In which ways can it be generalized? How does it relate to other criteria we might want a loss function to satisfy? Also worth investigating would be to more directly investigate the statistical consequences of loss functions for signed networks either satisfying or not satisfying certain criteria. For example, if we say that a random sequence of signed networks asymptotically converges to a certain limit network if the value of the loss with respect to that network approaches 0 in probability, then what properties must such a random sequence of signed networks have? The answer clearly depends on the properties of the chosen loss function. Finally, it should be desirable to connect these ideas to other potentially related areas. For example, to the extent that we can consider a random signed network as a random (sparse) matrix via the adjacency matrix representation, there should be meaningful connections to the already substantially developed field of random matrix theory. Likewise, thinking of networks as functions (e.g. as is done in sections 8.3.1 or 8.B), at least "morally" one should expect to be able to apply or modify ideas from functional data analysis [WCM16] [RS05] [HE15] [KR21] about what are "reasonable" choices of loss functions.

Chapter 8.   Appendices

Appendix 8.A   Lemmas

The symbol $\mathbf{1}_I$ denotes the all ones vector $(1,\ldots,1)$ in $\mathbb{R}^I$, the notation $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}]$ denotes the arithmetic mean (empirical expectation) of $\vec{\mathbf{x}} \in \mathbb{R}^I$, $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] := \dfrac{1}{I}\sum_{i=1}^{I} x^{(i)}$, and $\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle$ denotes the standard inner product[20] $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{R}^I$, $\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle := \sum_{i=1}^{I} x^{(i)} y^{(i)}$.

8.A.1   Mixed-Sign Spearman Correlation

Lemma 8.1. *Projection Lemma: when taking the (standard) inner product, "mean-centering one vector is as good as mean-centering both vectors":*

$$\begin{aligned}
\langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle &= \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle \\
&= \langle \vec{\mathbf{x}}, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle .
\end{aligned} \tag{8.A.1}$$

Here we assume $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{R}^I$, and recall that $\mathbf{1}_I \in \mathbb{R}^I$ denotes the all 1's vector.

   Proof: Because inner products are symmetric,

$$\langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle = \langle \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I, \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I \rangle . \tag{8.A.2}$$

Therefore it suffices to show only the first of the two claimed equalities, namely

$$\langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle = \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle ,$$

with the second equality following by applying symmetry (8.A.2), repeating the steps of the proof of the first equality, and then applying symmetry (8.A.3):

$$\langle \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I, \vec{\mathbf{x}} \rangle = \langle \vec{\mathbf{x}}, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle . \tag{8.A.3}$$

Anyway the conclusion follows readily from the definitions and simple algebra, so providing a proof to this level of detail is probably an over-explanation making the result seem more

---

[20]Also known as the "dot product".

confusing than it actually is. Nevertheless, at least for the sake of being thorough, below is a computation proving the first equality:

$$
\begin{aligned}
\langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle =& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \mathbf{1}_I \rangle \\
=& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot (\langle \vec{\mathbf{x}}, \mathbf{1}_I \rangle - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \langle \mathbf{1}_I, \mathbf{1}_I \rangle) \\
=& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \left(\sum_{i=1}^{I} x^{(i)} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot I \right) \\
=& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \left(\sum_{i=1}^{I} x^{(i)} - \sum_{i=1}^{I} x^{(i)} \right) \\
=& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle - 0 . \quad \square
\end{aligned}
\tag{8.A.4}
$$

For geometric intuition about why we should expect this lemma to be true, consider how $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I$ is the projection[21] of $\vec{\mathbf{x}}$ onto the span of $\mathbf{1}_I$. Thus mean-centering $\vec{\mathbf{x}}$ is the same as projecting[22] $\vec{\mathbf{x}}$ onto the orthogonal complement of (the span of) $\mathbf{1}_I$. Hence the projection of $\vec{\mathbf{y}}$ onto the span of $\mathbf{1}_I$, namely $\hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I$, is orthogonal to the mean-centered version of $\vec{\mathbf{x}}$. So in the expression $\langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle$, the contribution from $\hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I$ cancels out:

$$
\begin{aligned}
\langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} \rangle =& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, (\vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I) + \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle \\
=& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle + \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle \\
=& \langle \vec{\mathbf{x}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}] \cdot \mathbf{1}_I, \vec{\mathbf{y}} - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}] \cdot \mathbf{1}_I \rangle + 0 .
\end{aligned}
\tag{8.A.5}
$$

In general, if π denotes an operator projecting onto some subspace, then for analogous reasons it is always true that $\langle \pi(\vec{\mathbf{x}}), \pi(\vec{\mathbf{y}}) \rangle = \langle \pi(\vec{\mathbf{x}}), \vec{\mathbf{y}} \rangle = \langle \vec{\mathbf{x}}, \pi(\vec{\mathbf{y}}) \rangle$.

Lemma 8.2. *As long as tied values are replaced with the mean of the tied values[23], then any two rank vectors of the same length have the same mean.*

Proof: In the case that there are no ties, then it follows easily that all rank vectors of length L have the same mean as the vector $(1, \ldots, L)$. Because addition is commutative, if two vectors are the same up to a permutation of their entries, then they must have the same mean. In the absence of ties, all rank vectors by definition have entries which are permuted from the vector $(1, \ldots, L)$.

In the case of tied values, we can use the general fact that replacing any subset $\mathcal{S} \subseteq [I]$ of the entries of a vector with the mean value of the entries in the subset does not change the mean of the vector. In the computation below, let $\vec{\mathbf{v}}$ denote the original vector, and let $\vec{\mathbf{v}}'$ denote the vector created by replacing the entries of $\vec{\mathbf{v}}$ corresponding to the subset $\mathcal{S}$ with

---

[21] Herein, whenever I say "projection", I mean specifically "orthogonal projection".

[22] Likewise, herein "projecting" always refers to "orthogonally projecting".

[23] This is the way tied values are replaced by default in SciPy[VGO+20]. "The average of the ranks that would have been assigned to all the tied values is assigned to each value".

their mean value:

$$
\begin{aligned}
\hat{\boldsymbol{\mu}}[\vec{\mathbf{v}}'] :=& \frac{1}{I} \sum_{i=1}^{I} (v')^{(i)} \\
=& \frac{1}{I} \left(\sum_{i \in \mathcal{S}} (v')^{(i)} + \sum_{i \notin \mathcal{S}} (v')^{(i)} \right) \\
=& \frac{1}{I} \left(\sum_{i \in \mathcal{S}} \left(\frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} v^{(i)} \right) + \sum_{i \notin \mathcal{S}} v^{(i)} \right) \\
=& \frac{1}{I} \left(|\mathcal{S}| \cdot \left(\frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} v^{(i)} \right) + \sum_{i \notin \mathcal{S}} v^{(i)} \right) \\
=& \frac{1}{I} \left(\sum_{i \in \mathcal{S}} v^{(i)} + \sum_{i \notin \mathcal{S}} v^{(i)} \right) \\
=& \frac{1}{I} \sum_{i=1}^{I} v^{(i)} \\
=& \hat{\boldsymbol{\mu}}[\vec{\mathbf{v}}].
\end{aligned}
\tag{8.A.6}
$$

Applying this argument for each of the groups of tied elements, we get that the mean of a rank vector with tied values is the same as that of a rank vector without any ties, which in turn is the same as the mean of $(1, \ldots, L)$. $\square$

Lemma 8.3. ***Conditions for additivity of covariance for concatenated vectors:*** *Let $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2 \in \mathbb{R}^I$, and $\vec{\mathbf{y}}_1, \vec{\mathbf{y}}_2 \in \mathbb{R}^J$. If at least one of*

- *$\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1]$, or*

- *$\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_2] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_2]$*

is true, then covariance is additive for $\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1$ and $\vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2$, i.e.

$$
\hat{\mathrm{cov}}[\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1, \vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2] = \hat{\mathrm{cov}}[\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2] + \hat{\mathrm{cov}}[\vec{\mathbf{y}}_1, \vec{\mathbf{y}}_2].
\tag{8.A.7}
$$

Proof: Because covariance is symmetric, following reasoning analogous to that from the proof of Lemma 8.1, it suffices to give a proof only for the case that $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1]$. Note

that $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1]$ implies $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1]$.

$$
\begin{aligned}
\hat{\mathrm{cov}}[\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1, \vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2] &= \langle \vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1] \cdot \mathbf{1}_{I+J} \ , \ \vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2] \cdot \mathbf{1}_{I+J} \rangle \\
&= \langle \vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1] \cdot \mathbf{1}_{I+J} \ , \ \vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2 \rangle \\
&= \langle \vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1 - (\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] \cdot \mathbf{1}_I) \oplus (\hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1] \cdot \mathbf{1}_J) \ , \ \vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2 \rangle \\
&= \langle (\vec{\mathbf{x}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] \cdot \mathbf{1}_I) \oplus (\vec{\mathbf{y}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1] \cdot \mathbf{1}_J) \ , \ \vec{\mathbf{x}}_2 \oplus \vec{\mathbf{y}}_2 \rangle \\
&= \langle \vec{\mathbf{x}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] \cdot \mathbf{1}_I \ , \ \vec{\mathbf{x}}_2 \rangle + \langle \vec{\mathbf{y}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1] \cdot \mathbf{1}_J \ , \ \vec{\mathbf{y}}_2 \rangle \\
&= \langle \vec{\mathbf{x}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] \cdot \mathbf{1}_I \ , \ \vec{\mathbf{x}}_2 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_2] \cdot \mathbf{1}_I \rangle + \langle \vec{\mathbf{y}}_1 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1] \cdot \mathbf{1}_J \ , \ \vec{\mathbf{y}}_2 - \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_2] \cdot \mathbf{1}_J \rangle \\
&= \hat{\mathrm{cov}}[\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2] + \hat{\mathrm{cov}}[\vec{\mathbf{y}}_1, \vec{\mathbf{y}}_2].
\end{aligned}
$$
(8.A.8)

The second and second-to-last equalities of (8.A.8) use the Projection Lemma, Lemma 8.1. The third equality of (8.A.8) uses $\hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1 \oplus \vec{\mathbf{y}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{x}}_1] = \hat{\boldsymbol{\mu}}[\vec{\mathbf{y}}_1]$. The fourth equality of (8.A.8) uses property (8.4.12) from section 8.4.2.4. The fifth equality of (8.A.8) uses property (8.4.13) from section 8.4.2.4. The first and last equalities of (8.A.8) are just the definition of (empirical) covariance, of course. $\square$

Lemma 8.4. *Subconvex Combination Decomposition Property:* *For the coefficients* c_+, c_- *from (8.4.16) with the property that*

$$
\rho_S^{\pm}(\mathcal{G}_1, \mathcal{G}_2) = c_+ \rho_S^+(\mathcal{G}_1, \mathcal{G}_2) + c_- \rho_S^-(\mathcal{G}_1, \mathcal{G}_2),
$$

we always have that $c_+ + c_- \le 1$. *(Trivially,* $c_+, c_- \ge 0$.*)*

Proof: For convenience, define the following variables:

$$
\begin{aligned}
a &:= \hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}\left(\mathcal{G}_1^+\right)\right] \\
b &:= \hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}\left(\mathcal{G}_1^-\right)\right] \\
c &:= \hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}\left(\mathcal{G}_2^+\right)\right] \\
d &:= \hat{\mathrm{var}}\left[\underset{\mathcal{G}_1\mathcal{G}_2}{\mathcal{R}}\left(\mathcal{G}_2^-\right)\right].
\end{aligned}
$$
(8.A.9)

It follows then that the coefficients c_+, c_- from (8.4.16) equal

$$
c_+ = \frac{\sqrt{ac}}{\sqrt{a+b} \cdot \sqrt{c+d}}, \quad c_- = \frac{\sqrt{bd}}{\sqrt{a+b} \cdot \sqrt{c+d}}.
$$
(8.A.10)

Thus the inequality $c_+ + c_- \le 1$ is equivalent to

$$
\frac{\sqrt{ac} + \sqrt{bd}}{\sqrt{a+b} \cdot \sqrt{c+d}} \le 1 \quad \Longleftrightarrow \quad \sqrt{ac} + \sqrt{bd} \le \sqrt{a+b} \cdot \sqrt{c+d}.
$$
(8.A.11)

However this follows immediately from the Cauchy-Schwarz inequality, namely

$$|\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle| \leq \|\vec{\mathbf{u}}\|_2 \|\vec{\mathbf{v}}\|_2 \,,$$

if we define the vectors $\vec{\mathbf{u}} := (\sqrt{a}, \sqrt{b})$ and $\vec{\mathbf{v}} := (\sqrt{c}, \sqrt{d})$.

Another, equivalent way to show that the inequality (8.A.11) is true is by using the AM-GM inequality applied to the vector (ad, bc):

$$
\begin{aligned}
\frac{ad + bc}{2} &\geq \sqrt{abcd} \\
\Longleftrightarrow\ ad + bc + (ac + bd) &\geq 2\sqrt{abcd} + (ac + bd) \\
\Longleftrightarrow\ (a + b)(c + d) &\geq (\sqrt{ac} + \sqrt{bd})^2 \\
\Longleftrightarrow\ \sqrt{a + b} \cdot \sqrt{c + d} &\geq \sqrt{ac} + \sqrt{bd} \,.
\end{aligned}
\tag{8.A.12}
$$

The analysis is somewhat simplified given that we can assume $a, b, c, d \geq 0$. $\square$

8.A.2   DeltaCon Distance

Lemma 8.5. *Any choice of $\varepsilon < \frac{1}{1+\|\mathbf{A}\|_\infty}$ is sufficient to guarantee that the Neumann series $(\mathbf{I} - \mathbf{W})^{-1}$ converges.*

Proof: To review, the ∞-operator (or "induced") matrix norm is the "maximum absolute row sum" of a matrix $\mathbf{M}$:

$$\|\mathbf{M}\|_\infty := \max_i \left[\sum_{j=1}^J |[\mathbf{M}]_{ij}| \right] . \tag{8.A.13}$$

This operator norm is arguably one of the easiest to use to get a bound on the spectral radius because the definition of the degree matrix $\mathbf{D}$ from equation (8.3.2) of section 8.3.2 involves the row sums of $\mathbf{A}$.

Recall from section 8.4.4 that the matrix $\mathbf{W}$ used to create the Neumann series $(\mathbf{I} - \mathbf{W})^{-1}$ is defined as

$$\mathbf{W} := \frac{\varepsilon}{1 - \varepsilon^2} \mathbf{A} - \frac{\varepsilon^2}{1 - \varepsilon^2} \mathbf{D} \,. \tag{8.A.14}$$

Reviewing from section 8.4.4.1, to ensure that the Neumann series $(\mathbf{I} - \mathbf{W})^{-1}$ exists, our goal is to ensure that the spectral radius $\mathrm{sr}(\mathbf{W})$ of $\mathbf{W}$ is less than 1, $\mathrm{sr}(\mathbf{W}) < 1$, by ensuring the sufficient condition that $\|\mathbf{W}\|_\infty < 1$.

As long as $1 - \varepsilon^2 > 0$, tedious algebra with the definition of $\mathbf{D}$ shows that

$$\|W\|_\infty = \frac{\varepsilon}{1 - \varepsilon^2} \max_i \left[\sum_{j \neq i} |[\mathbf{A}]_{ij}| + \left| [\mathbf{A}]_{ii} - \varepsilon \sum_{j=1}^J [\mathbf{A}]_{ij} \right| \right] . \tag{8.A.15}$$

Therefore the requirement that $\|\mathbf{W}\|_\infty < 1$ is equivalent to the condition:

$$\frac{1-\varepsilon^2}{\varepsilon}\|\mathbf{W}\|_\infty = \max_i \left[\sum_{j\neq i}|[\mathbf{A}]_{ij}| + \left|[\mathbf{A}]_{ii} - \varepsilon\sum_{j=1}^J[\mathbf{A}]_{ij}\right|\right] < \frac{1-\varepsilon^2}{\varepsilon}. \qquad (8.A.16)$$

Using the triangle inequality, we can bound $\frac{1-\varepsilon^2}{\varepsilon}\|\mathbf{W}\|_\infty$ by $(1+\varepsilon)\|\mathbf{A}\|_\infty$:

$$\begin{aligned}\frac{1-\varepsilon^2}{\varepsilon}\|\mathbf{W}\|_\infty &= \max_i\left[\sum_{j\neq i}|[\mathbf{A}]_{ij}| + \left|[\mathbf{A}]_{ii} - \varepsilon\sum_{j=1}^J[\mathbf{A}]_{ij}\right|\right] \\ &\leq \max_i\left[\sum_{j\neq i}|[\mathbf{A}]_{ij}| + |[\mathbf{A}]_{ii}| + \varepsilon\sum_{j=1}^J|[\mathbf{A}]_{ij}|\right] = (1+\varepsilon)\|\mathbf{A}\|_\infty.\end{aligned} \qquad (8.A.17)$$

Therefore it suffices to show that $(1+\varepsilon)\|\mathbf{A}\|_\infty < \frac{1-\varepsilon^2}{\varepsilon} = \frac{(1-\varepsilon)(1+\varepsilon)}{\varepsilon}$, which is equivalent to $\|\mathbf{A}\|_\infty < \frac{1-\varepsilon}{\varepsilon}$, which in turn (as can be shown via tedious algebra) is equivalent to the inequality $\varepsilon < \frac{1}{1+\|\mathbf{A}\|_\infty}$.

This satisfies $1-\varepsilon^2 > 0$ as long as $\|\mathbf{A}\|_\infty > 0$, i.e. $\mathbf{A}\neq\mathbf{0}$, where $\mathbf{0}$ denotes the all-zeros matrix (correponding to an empty network with no edges). However we can assume that $\mathbf{A}\neq\mathbf{0}$ without loss of generality, because when $\mathbf{A}=\mathbf{0}$, in turn $\mathbf{D}=\mathbf{0}$, implying that $\mathbf{W}=\mathbf{0}$, so that $(\mathbf{I}-\mathbf{W})^{-1}=\mathbf{I}^{-1}=\mathbf{I}$, so the Neumann series exists regardless of our choice of ε.

Therefore we have shown that $\varepsilon < \frac{1}{1+\|\mathbf{A}\|_\infty}$ is sufficient to guarantee that $\|\mathbf{W}\|_\infty < 1$, which in turn is sufficient to guarantee that $\mathrm{sr}(\mathbf{W}) < 1$, guaranteeing convergence of the Neumann series $(\mathbf{I}-\mathbf{W})^{-1}$. $\square$

By seeking to upper bound the Frobenius norm by 1, and using very naive bounding techniques (mostly triangle inequality), it is also possible to prove that satisfying the much weaker bound:

$$\varepsilon < \frac{1}{1+\sqrt{S}\|\mathbf{A}\|_\infty} \qquad (8.A.18)$$

is sufficient to guarantee convergence of the Neumann series. It is probably possible to refine that argument to get a much better bound from the Frobenius norm, analogous to what was done in a much less general context in [KKK$^+$11].

It follows from the proof of Lemma 8.5 that if *at least one* of the two **strict** inequalities $\mathrm{sr}(\mathbf{W}) < \|\mathbf{W}\|_\infty$ or $\frac{1-\varepsilon^2}{\varepsilon}\|\mathbf{W}\|_\infty < (1+\varepsilon)\|\mathbf{A}\|_\infty$ is true, then we can use $\varepsilon = \frac{1}{1+\|\mathbf{A}\|_\infty}$ without any problems.

However, there do exist matrices $\mathbf{A}$ such that both $\mathrm{sr}(\mathbf{W}) = \|\mathbf{W}\|_\infty$ *and* $\frac{1-\varepsilon^2}{\varepsilon}\|\mathbf{W}\|_\infty = (1+\varepsilon)\|\mathbf{A}\|_\infty$. These should be the only matrices for which $\varepsilon = \frac{1}{1+\|\mathbf{A}\|_\infty}$ is too large. The general conditions under which $\mathrm{sr}(\mathbf{W}) = \|\mathbf{W}\|_\infty$ is true are obscure to me, but I do know that the second requirement that needs to be satisfied for a counterexample, $\frac{1-\varepsilon^2}{\varepsilon}\|\mathbf{W}\|_\infty = (1+\varepsilon)\|\mathbf{A}\|_\infty$, should be true if and only if all entries in the row of $\mathbf{A}$ with the largest sum of absolute values have the same sign (because then the absence of terms with

opposite signs means there are no cancellations and thus that the triangle inequality holds with equality). Through happenstance I have found $\mathbf{A}$ that also satisfy the first requirement, that $\mathrm{sr}(\mathbf{W}) = \|\mathbf{W}\|_\infty$ when $\varepsilon = \frac{1}{1+\|\mathbf{A}\|_\infty}$.

Lemma 8.6. *There exist matrices $\mathbf{A}$ such that defining $\varepsilon = \frac{1}{1+\|\mathbf{A}\|_\infty}$ leads to $\mathrm{sr}(\mathbf{W}) = 1$ (implying that the Neumann series $(\mathbf{I} - \mathbf{W})^{-1}$ does not converge).*

Proof: Here is the matrix I found by happenstance that works

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -4.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -4.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -4.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{8.A.19}$$

Rescaling a counterexample matrix by a constant does not appear to affect its status as a counterexample, nor does deleting (and/or adding) rows and/or columns consisting entirely of 0's. This leads to the simpler counterexample:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 7 \\ 9 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \end{bmatrix}, \tag{8.A.20}$$

for which $\varepsilon = \frac{1}{1+\|\mathbf{A}\|_\infty} = \frac{1}{10}$, and thus

$$\mathbf{W} = \begin{bmatrix} -\frac{9}{99} & 0 & \frac{90}{99} & 0 \\ 0 & -\frac{7}{99} & 0 & \frac{70}{99} \\ \frac{90}{99} & 0 & -\frac{9}{99} & 0 \\ 0 & \frac{90}{99} & 0 & -\frac{9}{99} \end{bmatrix}. \tag{8.A.21}$$

Very tedious calculations reveal that indeed the spectral radius of the matrix from (8.A.21) is 1. Related counterexamples appear to also include

$$\mathbf{A} = \begin{bmatrix} 0 & \pm 9 \\ \pm 9 & 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \pm 9 & 0 \\ \pm 9 & 0 & 0 \\ 0 & 0 & \pm 9 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \pm 9 & 0 \\ \pm 9 & 0 & 0 \\ 0 & 0 & \pm x \end{bmatrix} \ (\forall x \text{ s.t. } |x| < 9).$$
$$\tag{8.A.22}$$

Discerning which operations will preserve the property $\mathrm{sr}(\mathbf{W}) = \|\mathbf{W}\|_\infty$ is subtle, since e.g. the following matrices appear to ***not*** be counterexamples:

$$\mathbf{A} = \begin{bmatrix} 0 & \pm x \\ \pm 9 & 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \pm 9 \\ \pm x & 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \pm x & 0 \\ \pm 9 & 0 & 0 \\ 0 & 0 & \pm 9 \end{bmatrix} \quad (\forall x \text{ s.t. } |x| < 9). \quad (8.A.23)$$

In any case, counterexamples clearly exist, but also appear to not be generic. $\square$

Appendix 8.B   Technical Details

These sections document preliminary steps towards making the underlying ideas discussed in this chapter precise, rigorous, and falsifiable.

8.B.1   Node Correspondences and Comparison of Networks with Mixed-Sign Edge Weights

Let $\mathscr{G}_{[S]}$ denote the space of all networks with mixed-sign edge weights that share the same set of nodes $[S]$. Thus any two networks $\mathcal{G}_1, \mathcal{G}_2 \in \mathscr{G}_{[S]}$ differ at most by their edges, in the sense that possibly either $\mathcal{E}_{\mathcal{G}_1} \neq \mathcal{E}_{\mathcal{G}_2}$ and/or possibly $A_{\mathcal{G}_1} \neq A_{\mathcal{G}_2}$, and never differ by (the ordering or labelling) of their nodes, because by definition of $\mathscr{G}_{[S]}$ we always have that $\mathcal{N}_{\mathcal{G}_1} = [S] = \mathcal{N}_{\mathcal{G}_2}$.

Any network $\tilde{\mathcal{G}}$ with mixed-sign edge weights that has S nodes (not necessarily the integers in $[S]$) can be identified with a network $\mathcal{G} \in \mathscr{G}_{[S]}$. This identification is not unique and occurs by explicitly choosing a bijection between $[S]$ and the node set $\mathcal{N}_{\tilde{\mathcal{G}}}$ of $\tilde{\mathcal{G}}$. Such choice amounts to a fixed, "correct" ordering and labelling of the S nodes in $\mathcal{N}_{\tilde{\mathcal{G}}}$. Cf. Appendix A.3.1 of the Master's thesis [Kri20].

For ease of reference, given two networks with mixed-sign edge weights $\tilde{\mathcal{G}}_1$, $\tilde{\mathcal{G}}_2$ with S nodes and bijective identifications $\mathcal{N}_{\tilde{\mathcal{G}}_1} \xrightarrow{f_1} [S]$ and $\mathcal{N}_{\tilde{\mathcal{G}}_2} \xrightarrow{f_2} [S]$, we call the resulting identifications $\mathcal{N}_{\tilde{\mathcal{G}}_1} \xrightarrow{f_2^{-1} \circ f_1} \mathcal{N}_{\tilde{\mathcal{G}}_2}$ and $\mathcal{N}_{\tilde{\mathcal{G}}_2} \xrightarrow{f_1^{-1} \circ f_2} \mathcal{N}_{\tilde{\mathcal{G}}_1}$ between the node sets of $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$ the **node correspondence** between $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$. Of course for $\mathcal{G}_1, \mathcal{G}_2 \in \mathscr{G}_{[S]}$, the node correspondence is always the identity on $[S]$. Restricting our consideration to $\mathscr{G}_{[S]}$, rather than all networks with mixed-sign edge weights and S nodes, can correctly be thought of as indicating that all node correspondences have already been chosen and fixed.

A function $\mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ is called an **objective function**. An objective function we want to minimize is called a "penalty" or "loss" function, while an objective function we want to maximize is called a "reward" function. A **dissimilarity function** $\mathscr{D} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ is an objective function for which higher values are considered "less similar". Special classes of dissimilarity functions are often referred to as "divergences" or "distances". In contrast, a **similarity function** $\mathscr{S} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ is an objective function for which higher values are considered "more similar".

In [TITP19], evaluating (analogous) similarity $\mathscr{S}$ and/or dissimilarity $\mathscr{D}$ functions (for unsigned networks) is called the "known node correspondence" comparison problem. Evaluating such objective functions might occur as a subroutine in methods for the "unknown node correspondence" comparison problem as defined in [TITP19] or in methods for so-called graph alignment problems. Those problems, including whether or how to modify them for signed networks, are completely outside the scope of discussion of this paper. Herein we deal only with the known node correspondence problem, for which node correspondences are already given and thus require no computational expense.

If a function φ is monotone nondecreasing, i.e. $x_1 \leq x_2 \implies \varphi(x_1) \leq \varphi(x_2)$ for all x_1, x_2, then given a dissimilarity function $\mathscr{D} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ one has that $\varphi(\mathscr{D})$ is always again a dissimilarity function, and likewise given a similarity function $\mathscr{S} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ one has that $\varphi(\mathscr{S})$ is always again a similarity function. If a function ψ is monotone nonincreasing, i.e. $x_1 \leq x_2 \implies \psi(x_1) \geq \psi(x_2)$ for all x_1, x_2, then given a dissimilarity function $\mathscr{D} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ one has that $\psi(\mathscr{D})$ is always a *similarity* function, and likewise given a similarity function $\mathscr{S} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ one has that $\psi(\mathscr{S})$ is always a *dissimilarity* function. The analysis of dissimilarity functions can always be reduced to the analysis of similarity functions, and vice versa. Thus, when discussing the known node comparison problem, we can always restrict discussion to dissimilarity functions (or to similarity functions) without losing any generality.

This paper asks which qualities or properties make a dissimilarity function (or equivalently, in the sense explained above, a similarity function) for networks with mixed-sign edge weights "reasonable" or "well-behaved". The Double Penalization Principle (defined below) provides one possible answer.

A dissimilarity minimization problem (equivalently a similarity maximization problem) is of the form:

$$\underset{(\mathcal{G}_1, \mathcal{G}_2) \in \mathbb{G} \subseteq \mathscr{G}_{[S]}^2}{\operatorname{argmin}} \mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) \quad \text{or} \quad \underset{(\mathcal{G}_1, \mathcal{G}_2) \in \mathbb{G} \subseteq \mathscr{G}_{[S]}^2}{\operatorname{argmax}} \mathscr{S}(\mathcal{G}_1, \mathcal{G}_2), \qquad (8.B.1)$$

for some subset $\mathbb{G}$ of $\mathscr{G}_{[S]}^2 \overset{def}{=} \mathscr{G}_{[S]} \times \mathscr{G}_{[S]}$ containing "admissible" pairs of networks with mixed-sign edge weights, and for some dissimilarity function $\mathscr{D} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ or for some similarity function $\mathscr{S} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$. Using the same notation, a dissimilarity maximization problem (equivalently a similarity minimization problem) is of the form:

$$\underset{(\mathcal{G}_1, \mathcal{G}_2) \in \mathbb{G} \subseteq \mathscr{G}_{[S]}^2}{\operatorname{argmax}} \mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) \quad \text{or} \quad \underset{(\mathcal{G}_1, \mathcal{G}_2) \in \mathbb{G} \subseteq \mathscr{G}_{[S]}^2}{\operatorname{argmin}} \mathscr{S}(\mathcal{G}_1, \mathcal{G}_2). \qquad (8.B.2)$$

The specific problem motivating this paper is the dissimilarity minimization problem in the case that $\mathbb{G} = \{\hat{\mathcal{G}}_1, \ldots, \hat{\mathcal{G}}_N\} \times \{\mathcal{G}_*\}$, where $\mathcal{G}_*$ is a "ground truth" network and the networks $\hat{\mathcal{G}}_1, \ldots, \hat{\mathcal{G}}_N$ are "candidate estimate" networks. We want to make a *quantitative* statement about which of the N networks[24] $\hat{\mathcal{G}}_1, \ldots, \hat{\mathcal{G}}_N$ is the "least dissimilar" to the "ground truth"

---

[24] $\mathcal{G}_*$ is in all of the pairs belonging to $\mathbb{G}$, hence the optimization in this case is effectively only over the $\hat{\mathcal{G}}_1, \ldots, \hat{\mathcal{G}}_N$.

$\mathcal{G}_*$ and thus the "best" estimate of $\mathcal{G}_*$. Obviously whether the resulting claim, about which estimate is "best", is "reasonable" depends inherently on whether our choice of dissimilarity function $\mathscr{D}$ is itself "reasonable". Therefore we must first decide on principles characterizing "reasonable" dissimilarity functions in this context before we are able to "reasonably" decide which statistical estimators for networks with mixed-sign edge weights perform "best".

The computational complexity of solving the motivating optimization problem is $O(C(S)N)$, where $C(S)$ denotes the (maximum) number of operations required to evaluate our chosen dissimilarity function $\mathscr{D}$ on a single pair of networks in $\mathscr{G}_{[S]}$. Therefore the computational issues associated with the motivating example are trivial, and thus we will not further consider computational complexity issues herein.

8.B.2   Why (Most) Pre-Existing Comparison Methods are Insufficient

One might try to bypass the comparison problem for signed networks by using a comparison method for unsigned networks. The procedure is the following. Given a pair of signed networks, we could "project" both of them into a space of (unsigned) networks with less structure. (Cf. section 8.3.3 for examples.) Applying a comparison method for less structured networks to these "projected" counterparts then corresponds to defining a "new" comparison method on the space of signed networks. Cf. figure 8.B.1.
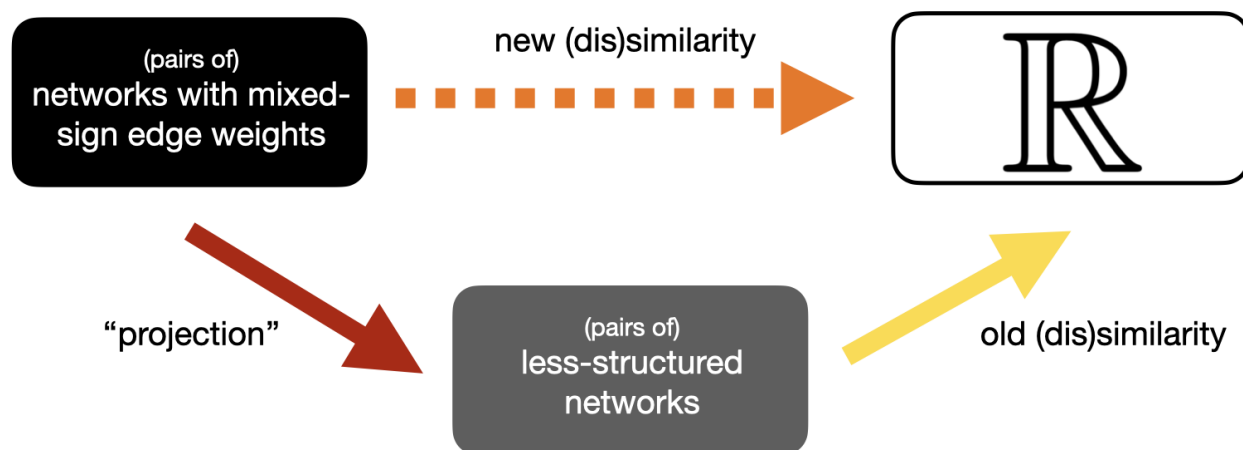
Figure 8.B.1: "Projection" procedure for applying comparison methods intended for networks with less structure to networks with mixed-sign edge weights. This is intended to (roughly) correspond to a so-called "commutative diagram", i.e. both paths from left to right are equal. Cf. [LS09] for an accessible introduction to commutative diagrams.

Using this "projection" procedure to apply comparison methods for unsigned networks to signed networks exposes us to potential pitfalls. We need to avoid situations where, for

example, positive and negative weights "cancel" in a scientifically meaningless way, or two edges with different signs are not considered distinct due to having similar magnitudes.

For example, positive interactions and negative interactions correspond to categorically and qualitatively distinct phenomena in microbial ecology. Cf. again section 8.2.2. If all growth-promoting (positive) interactions were changed to growth-suppressing (negative) interactions, ecologically speaking the result would be quite different. This is true even if, or in some cases even especially if, the magnitudes of the interactions were left unchanged, as happens when "projecting" onto the space of unsigned networks. This implies that the positive and negative edges of signed networks should always be treated as qualitatively distinct by any well-behaved comparison method, at least for applications to microbial ecology.

The distinction between positive and negative edges is presumably also scientifically important for other applications of signed networks. If the distinction weren't important for a given application, then there would seem to be no good reason to not use an unsigned network for that application instead. Cf. the discussion from the first part of section 8.1.1.

The reason why the "projection" procedure is inherently problematic is because two networks that are very distinct as signed and weighted networks may become difficult to distinguish after being "projected" into spaces of networks with less structure. This means that comparison methods resulting from the procedure are oblivious to what should be identified as important and obvious differences in the original signed and weighted networks. The "double penalization principle" is the most important sanity check for ruling out the possibility of such counter-intuitive behavior. Similar, but less important, principles that can also be used to assess the behavior of a comparison method are also discussed in sections 8.2.6 and 8.2.7. The double penalization principle allows us to check whether a given extension of a comparison method for unsigned networks to signed networks implicitly uses the "projection" procedure. This allows us to avoid the pitfalls of the "projection" procedure and to know when, and even how, to modify comparison methods so as to ensure they don't exhibit counter-intuitive behavior for signed networks.

8.B.3   Taxonomy of Single-Edge Comparisons

Any "reasonable" objective function should still be behaving "reasonably" when we "zoom in" on individual edges. In the following sections we try to make this intuition precise enough so that it can guide our choices about what constitute "reasonable" objective functions. First, we propose a way to categorize the kinds of differences that can exist between two networks with mixed-sign edge weights at the level of a single edge[25]. Next, we seek to describe how "reasonable" objectives should behave in response to perturbations that "shift" between these categories of differences. Finally, we aim to propose one way to make precise constraints on the "global" behavior of "reasonable" objective functions that would result from the earlier considerations about the "local" level of single edges.

---

[25] Edge here in the sense of an element of $[S] \times [S]$, i.e. a given input e to the "edge functions" $A_{\mathcal{G}_1}$ and $A_{\mathcal{G}_2}$ (not necessarily lying in either of their supports, i.e. not necessarily in either "edge set").

Let us first analyze the situation without taking magnitude differences into account. This can be interpreted as either (i) focusing on the signed but unweighted case, or (ii) as a first-pass, "naive" analysis of the general signed and weighted case. Either way, incorporating discussions of magnitude differences is easier once we have established the initial framework that neglects them. This is because how to correctly interpret any given magnitude difference depends on what the signs of the underlying edges are.

Not taking magnitude differences into account, there are $3^2 = 9$ different possible kinds of single-edge comparisons when taking the order of edges into account:

$$(+,+), (+,0), (+,-), (0,+), (0,0), (0,-), (-,+), (-,0), (-,-).\qquad(8.B.3)$$

In contrast, when disregarding the order of the edges, there are $\binom{3}{2} + \binom{3}{1} = 6$ different possible kinds of single-edge comparisons:

$$\{(+,-), (-,+)\}, \{(+,0), (0,+)\}, \{(-,0), (0,-)\}, \{(+,+)\}, \{(0,0)\}, \{(-,-)\}.\qquad(8.B.4)$$

Thus by a "taxonomy" or "classification scheme" of single-edge comparisons, we mean some partition [26] of either (8.B.3) or (8.B.4). Preferably any such proposed partition has a sensible organizing principles behind it, rather than being arbitrarily chosen.

We can categorize single-edge comparisons into $2 \times 2 = 4$ categories according to:

- whether both edge counterparts are non-zero (non-trivial vs. trivial),

- whether the signs of the edge counterparts coincide (sign agreement vs. sign disagreement).

	sign agreement	sign disagreement
trivial	$\{(0,0)\}$	$\{(0,+), (+,0), (0,-), (-,0)\}$
non-trivial	$\{(+,+), (-,-)\}$	$\{(+,-), (-,+)\}$

Table 8.B.1: Signs corresponding to each category of single-edge comparison.

8.B.4   Applying Framework to Constrain Dissimilarity Functions

Within this framework let us consider the properties that a "reasonable" dissimilarity $\mathscr{D}$ function or similarity $\mathscr{S}$ function would have. Specifically, we claim that an objective function can be designated "reasonable" or not based on its behavior in response to "perturbations" that shift single-edge comparisons from one of the above categories to another. This sets aside for now the separate issue of changes in magnitude, i.e. we assume that the magnitude of any non-zero edge is the same constant value, e.g. 1.

---

[26] Technically (8.B.4) itself is a partition of (8.B.3), such that a partition of (8.B.4) in practice amounts to nothing more than a coarsening of the corresponding partition of (8.B.3), but this is a pedantic, intuitively obvious, and thus not terribly important, point.

We start from $4 \times 4 = 16$ possible shifts between categories, but I claim we only need to consider $\binom{4}{2} = 6$ of these. Because all non-zero edges are assumed to have the same magnitude (or equivalently for now we neglect to consider changes in magnitude), any shift from the same category to itself can be assumed to be the "identity" shift that does nothing. Hence any objective function, reasonable or not, will register no change in values for any of these 4 "identity" shifts. This leaves $16 - 4 = 12$ shifts to consider.

		After Shift			
		trivial sign agreement	**trivial sign disagreement**	**non-trivial sign agreement**	**non-trivial sign disagreement**
Before Shift	**trivial sign agreement**	no change	?	?	?
	trivial sign disagreement	?	no change	?	?
	non-trivial sign agreement	?	?	no change	?
	non-trivial sign disagreement	?	?	?	no change

Table 8.B.2: Taking "identity" shifts into account

For the remaining 12 shifts, note that we only need to consider $12/2 = 6$ shifts due to the elementary fact that $x_1 \leq x_2$ if and only if $x_2 \geq x_1$. For example, how an objective function changes in response to a shift from a trivial sign agreement to a non-trivial sign disagreement completely determines how it changes in response to a shift from a non-trivial sign disagreement to a trivial sign agreement.

		After Shift			
		trivial sign agreement	**trivial sign disagreement**	**non-trivial sign agreement**	**non-trivial sign disagreement**
Before Shift	**trivial sign agreement**	no change	?	?	?
	trivial sign disagreement	opposite of value across diagonal	no change	?	?
	non-trivial sign agreement	opposite of value across diagonal	opposite of value across diagonal	no change	?
	non-trivial sign disagreement	opposite of value across diagonal	opposite of value across diagonal	opposite of value across diagonal	no change

Table 8.B.3: Pointing out redundant specifications.

Elementary considerations already determine how a "reasonable" objective should behave for 4 of the 6 non-redundant shifts. Namely, it's clear that when shifting from any "agreement" to any "disagreement" that a "reasonable" dissimilarity $\mathscr{D}$ function should not decrease and that a "reasonable" similarity $\mathscr{S}$ function should not increase. I will write "$\mathscr{D}{\uparrow}, \mathscr{S}{\downarrow}$" to indicate a non-decreasing change in value for a "reasonable" dissimilarity $\mathscr{D}$/a non-increasing change in value for a "reasonable" similarity $\mathscr{S}$.

This is equivalent to the stipulation that, when shifting from any "disagreement" to any "agreement", a "reasonable" dissimilarity $\mathscr{D}$ function should not increase and that a

"reasonable" similarity $\mathscr{S}$ function should not decrease. I will write "$\mathscr{D}\!\downarrow, \mathscr{S}\!\uparrow$" to indicate a non-increasing change in value for a "reasonable" dissimilarity $\mathscr{D}$/a non-decreasing change in value for a "reasonable" similarity $\mathscr{S}$.

		After Shift			
		trivial sign agreement	**trivial sign disagreement**	**non-trivial sign agreement**	**non-trivial sign disagreement**
Before Shift	**trivial sign agreement**	no change	$\mathscr{D}\!\uparrow, \mathscr{S}\!\downarrow$	?	$\mathscr{D}\!\uparrow, \mathscr{S}\!\downarrow$
	trivial sign disagreement	$\mathscr{D}\!\downarrow, \mathscr{S}\!\uparrow$	no change	$\mathscr{D}\!\downarrow, \mathscr{S}\!\uparrow$	?
	non-trivial sign agreement	opposite of value across diagonal	$\mathscr{D}\!\uparrow, \mathscr{S}\!\downarrow$	no change	$\mathscr{D}\!\uparrow, \mathscr{S}\!\downarrow$
	non-trivial sign disagreement	$\mathscr{D}\!\downarrow, \mathscr{S}\!\uparrow$	opposite of value across diagonal	$\mathscr{D}\!\downarrow, \mathscr{S}\!\uparrow$	no change

Table 8.B.4: Trivial considerations taken into account.

Note that technically it's redundant to specify the behavior of both a "reasonable" dissimilarity $\mathscr{D}$ and a "reasonable" similarity $\mathscr{S}$, because one should transform to the other via a non-increasing transformation[27]. I note both behaviors explicitly for the sake of expositional clarity, not mathematical necessity.

8.B.5 Choices Available After Imposing "Trivial" Constraints

Comparing with table 8.B.4, we see that this leaves only two kinds of shifts for which the behavior of a "reasonable" objective function is not trivially determined. Both remaining shifts are shifts "within agreement or disagreement", and "between non-triviality and triviality".

First, there is the shift from a trivial sign disagreement to a non-trivial sign disagreement (or equivalently in the sense of table 8.B.3 the shift from a non-trivial sign disagreement to a trivial sign disagreement).

Second, there is the shift from a trivial sign agreement to a non-trivial sign agreement (or equivalently in the sense of table 8.B.3 the shift from a non-trivial sign agreement to a trivial sign agreement).

Comparing the principles mentioned in section 8.2, we see that:

1. the "double penalization principle" from section 8.2.5 is the choice espoused herein to address shifts from a trivial sign disagreement to a non-trivial sign disagreement, and

2. the "sparsity savviness principle" from section 8.2.7 is the choice espoused herein to address shifts from a trivial sign agreement to a non-trivial sign agreement.

---

[27] If a function ψ is monotone nonincreasing, $x_1 \leq x_2 \implies \psi(x_1) \geq \psi(x_2)$ for all x_1, x_2, then given a dissimilarity function $\mathscr{D} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$, $\psi(\mathscr{D})$ is always a *similarity* function, and likewise given a similarity function $\mathscr{S} : \mathscr{G}_{[S]} \times \mathscr{G}_{[S]} \to \mathbb{R}$ one has that $\psi(\mathscr{S})$ is always a *dissimilarity* function. The analysis of dissimilarity functions can always be reduced to the analysis of similarity functions, and vice versa.

What the "correct" behavior for "reasonable" objectives for these two kinds of shifts is perhaps more arguable or subjective than for the others. Certainly the terminology "trivial" and "non-trivial" belies the opinions espoused herein. Namely, shifting from a "trivial" disagreement to a "non-trivial" disagreement should "amplify" the disagreement, and therefore correspond to a non-decrease for a dissimilarity $\mathscr{D}$ function/a non-increase for a similarity $\mathscr{S}$ function, i.e. $\mathscr{D}\underline{\uparrow}, \mathscr{S}\underline{\downarrow}$. Likewise, shifting from a "trivial" agreement to a "non-trivial" agreement[28] should "amplify" the agreement, and therefore correspond to a non-increase for a dissimilarity $\mathscr{D}$ function/a non-decrease for a similarity $\mathscr{S}$ function, i.e. $\mathscr{D}\underline{\downarrow}, \mathscr{S}\underline{\uparrow}$.

8.B.6 "Composability" Constraints for Objective Functions

One can argue fairly strongly on the basis of preferring "consistency", "composability", "continuity" of behavior of "rational" objectives for the double penalization principle.

Sparsity-savviness is compatible with the argument for "composability" constraints on "rational" objectives, but is also not supported by the argument.

This is analogous to another discrepancy between the double penalization principle and the sparsity-savviness principle. Namely, double penalization is straightforwardly compatible with how to reasonably penalize considering magnitude differences whereas sparsity-savviness is not. See section 8.B.7 below for a further discussion of this.

I believe that both discrepancies in fact share the same underlying cause. Unfortunately however I am not currently sure how to make any of this precise enough to falsify or confirm.

Consider the following undirected network:
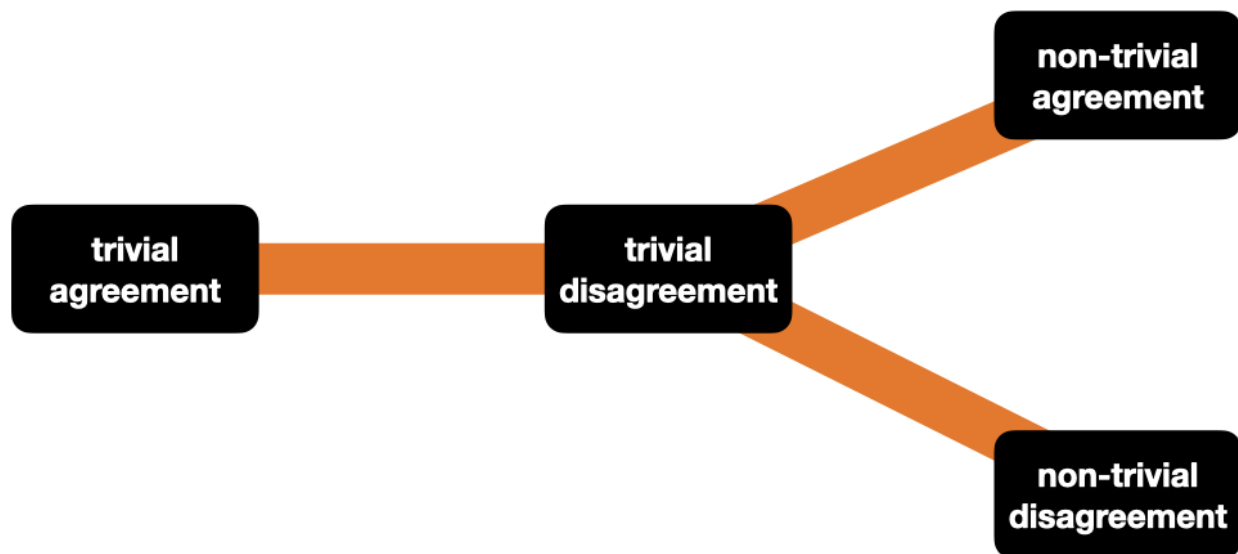
Figure 8.B.2: Spoke graph for the "edit" relationships of the four single-edge comparisons.

---

[28]Again, ceteris paribus, not (yet) taking magnitude differences into account.

Each node represents a different category of single-edge comparison, cf. again table 8.B.1. Two nodes are connected if one can be transformed to another by changing, or "editing", *one* of the two signs in the comparison to an "**adjacent sign**". By "adjacent sign" I mean $\{-, 0\}$ or $\{0, +\}$, corresponding to the ordering of the signs on the number line. For example, starting with a trivial agreement, $(0, 0)$, making the second edge positive $(0, +)$ corresponds to a trivial sign disagreement, as does making the first edge negative, $(-, 0)$. However making both the second edge positive *and* the first edge negative, $(-, +)$, is considered two changes or "edits", and therefore non-trivial sign disagreement is *not* adjacent to trivial sign agreement, but rather separated by a path of length two. This is similar in spirit to the "edit distance" discussed in [LPR19], but more abstract and with a different definition of "edits".

This spoke graph implies additional consistency conditions for a "reasonable" objective function to satisfy, besides the "trivial" ones from table 8.B.4. Namely, given a shift between categories corresponding to a path of length two, there are potentially two distinct "reasonable" behaviors implied by figure 8.B.2: (1) the behavior corresponding to the "direct jump" from the beginning of the path to the end of the path, (2) the behavior (if uniquely defined) corresponding to the composition of the behaviors of the two "edits" constituting the path. Requiring these two potentially distinct behaviors to coincide is the "composability" criterion for "rational" objectives alluded to above.

There are six directed paths of length two in the network from figure 8.B.2, corresponding to three undirected paths of length two (cf. table 8.B.3 again for why undirected suffices):

1. A path between "trivial agreements" and "non-trivial agreements",

2. a path between "trivial agreements" and "non-trivial disagreements", and

3. a path between "non-trivial agreements" and "non-trivial disagreements".

The last two belong to the "trivial" shifts from table 8.B.4 while the first does not.

Looking at all $\binom{3}{2} = 3$ pairs of these three undirected paths, we see that neither of the two pairs that include the relevant path provide unambiguous guidance for the behavior separating "trivial agreements" and "non-trivial agreements" via the "composability" criterion. Cf. figures 8.B.3a and 8.B.3b below. Hence, as mentioned before, "composability" neither provides an argument for, nor an argument against, the sparsity-savviness principle. Therefore one must look to other reasons to justify the Sparsity-Savviness Principle as a desideratum.

However, the third pair of undirected paths provides an unambiguous choice of behavior for one of the two shifts not decided by "trivial" considerations alone. Cf. figure 8.B.4 below. Thus "composability" argues strongly in favor of requiring the double penalization principle.
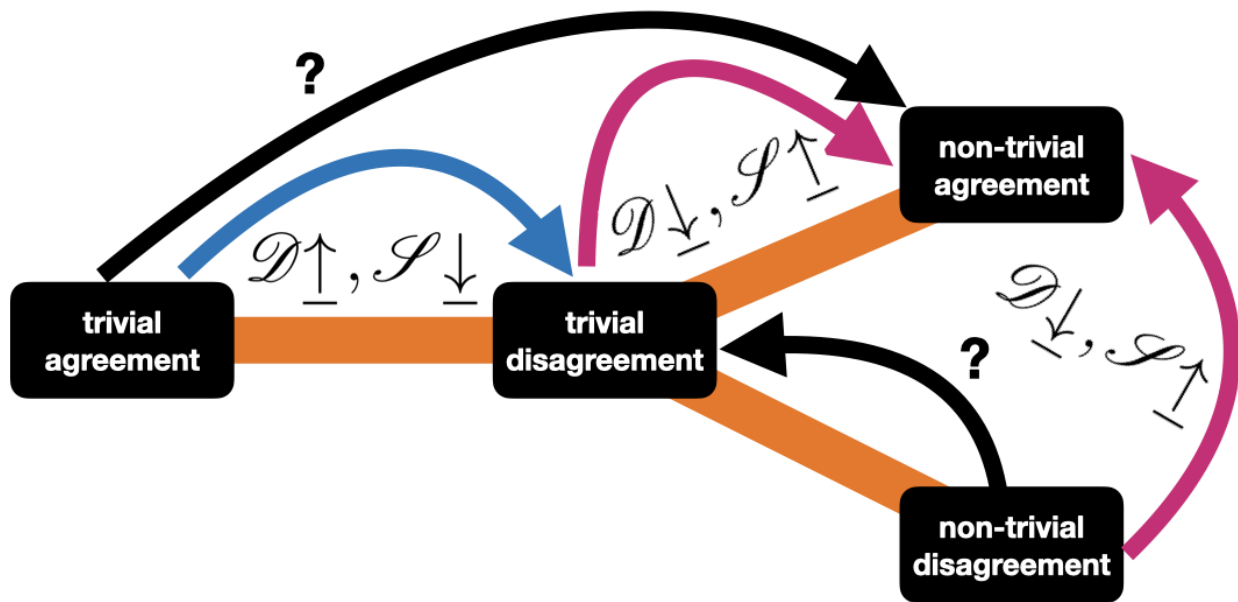
(a) These two paths fail to usefully constrain any unknown shift, even if we determined the shift between trivial and non-trivial disagreements.
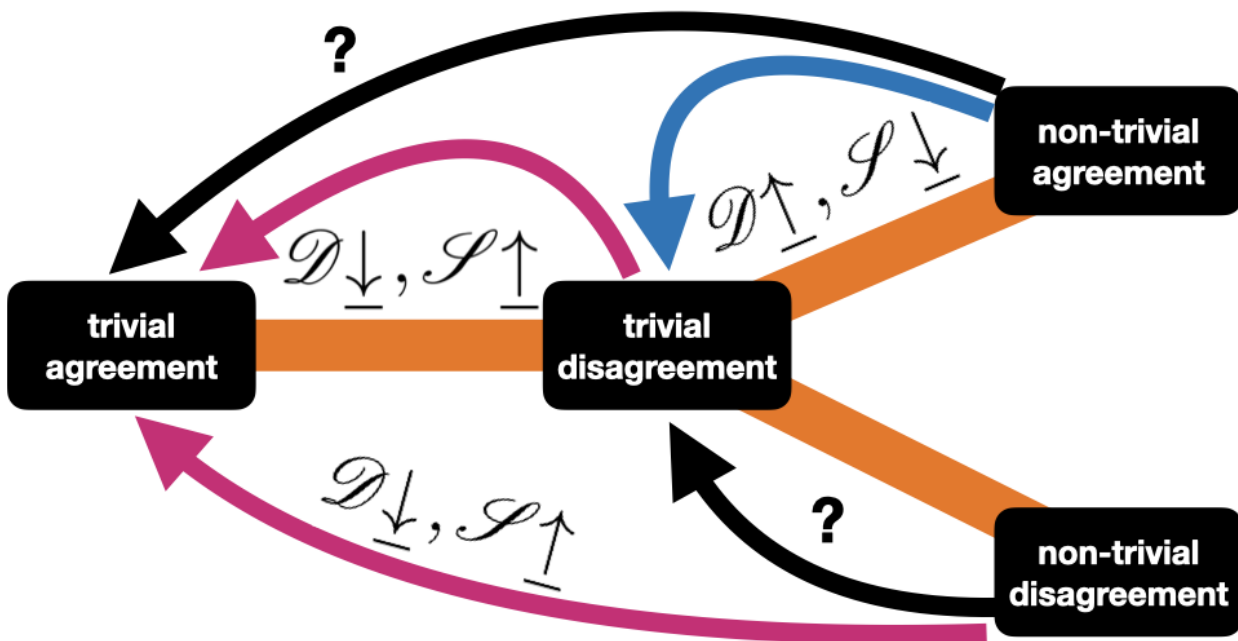
(b) These two paths *also* fail to usefully constrain any unknown shift, even if we determined the shift between trivial and non-trivial disagreements.

Figure 8.B.3: Using only the "trivial" behavior and the "composability" criterion, no useful constraints emerge for the shift between trivial agreement and non-trivial agreement.
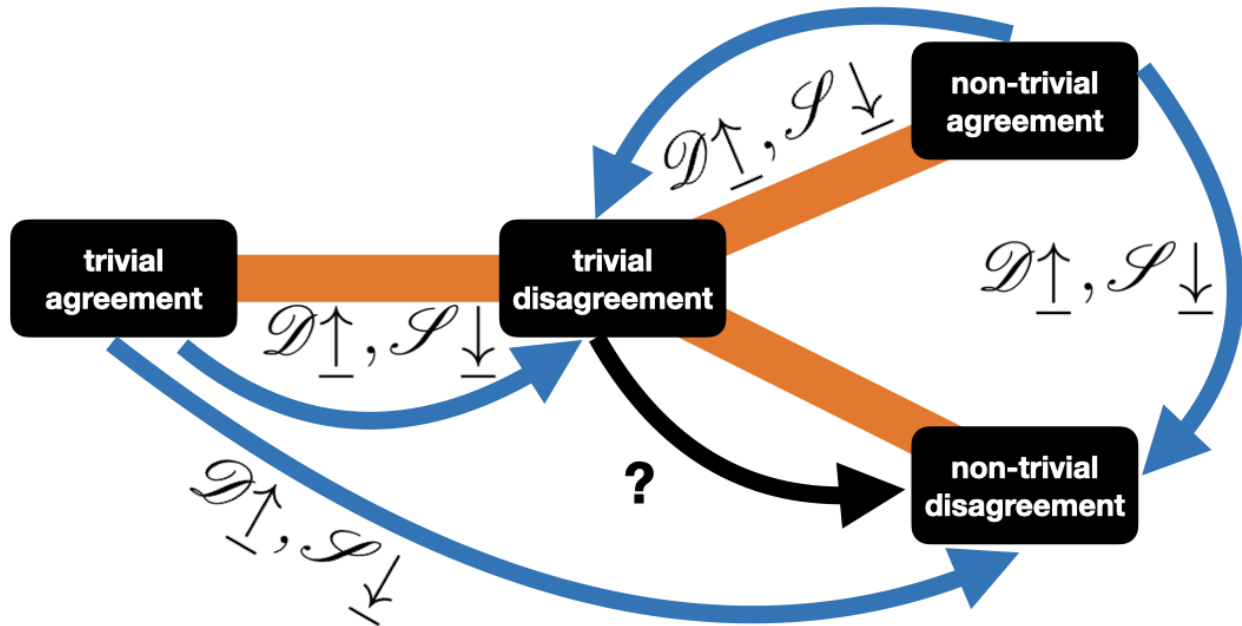
Figure 8.B.4: The "trivial" behavior combined with the "composability" criterion suggest only one possible option for the behavior of the unknown shift. Thus the double penalization principle is highly motivated.

Technically the constraint from figure 8.B.4, that going from trivial agreements to non-trivial disagreements never increases similarity, can be satisfied without imposing the double penalization principle. Doing so would require extra work however. Namely, for a given objective, one would have to demonstrate that the first step (trivial agreements to trivial disagreements) of the two part path from trivial agreements to non-trivial disagreements would always non-increase similarity more than any increase in similarity that might possibly occur during the second step (trivial disagreements to non-trivial disagreements), i.e. that the change corresponding to the first step would always be large enough to cancel out that from the second step and thus able to determine the overall trend by itself.

On the other hand, as figure 8.B.4 demonstrates, if we only consider objective functions to be "rational" when shifts from trivial disagreements to non-trivial disagreements non-decrease dissimilarity $\mathscr{D}\underline{\uparrow}$ (i.e. non-increase similarity $\mathscr{S}\underline{\downarrow}$), no extra work is ever required to satisfy the "composability" constraint. Restricting to such objective functions of course corresponds to requiring that the double penalization principle be satisfied.

8.B.7 Taking Magnitude Differences into Account

The above subsections limited the discussion to unweighted (weights ± 1) networks, both to simplify the discussion and to serve as a starting point for understanding the general weighted case. Preliminary steps towards the latter goal are taken in this section.

I conjecture (cf. section 8.B.6) that the greater compatibility of the double penalization principle, compared to the sparsity-savviness principle, with both (i) the "composability" constraint and (ii) consideration of magnitude differences, has one underlying cause.

(Assume in what follows that $x, y, z_1, z_2 > 0$.) In the general case where non-zero magnitudes can have values besides 1, a non-trivial sign agreement (between the corresponding edges of the compared networks) of the form $(+z_1, +z_2)$ could very well deserve a penalty, depending on the size of the magnitude difference $|z_1 - z_2|$. This is in explicit contrast to the unweighted case, where in the analogous situation $|z_1 - z_2| = 0$ always.

Of course, in both the unweighted and weighted cases, a non-trivial sign disagreement of the form $(+x, -y)$ always deserves a penalty. To summarize:

- In the unweighted case from before, there are never penalties arising from magnitude differences of non-trivial agreements, so there is no need to consider how to balance those with the penalties arising from sign differences of non-trivial disagreements.

- In the weighted case, we now must consider both kinds of penalties.

We need to consider whether, and to what extent, the ability to impose penalties for both of these kinds of differences might come at the expense of one another. This occurs especially when considering changes in the objective functions for shifts from non-trivial agreements $(+z_1, +z_2)$ to non-trivial disagreements $(+x, -y)$ (and vice versa). To answer that, we need to know when the penalty for $(+z_1, +z_2)$ smaller or larger than that for $(+x, -y)$.

According to the "composability" constraint, for "rational" objective functions the change in value when shifting from $(+z_1, +z_2)$ to $(+x, -y)$ should correspond to "composing"[29] the changes in value associated with any "path" from $(+z_1, +z_2)$ to $(+x, -y)$.

Figure 8.B.5a depicts what those possible paths are. Here are two explicitly:

$$
\begin{array}{ccccccc}
(+z_1, +z_2) & \rightarrow & (+x, +z_2) & \rightarrow & (+x, 0) & \rightarrow & (+x, -y) \\
\text{NTA} & \rightarrow & \text{NTA} & \rightarrow & \text{TDA} & \rightarrow & \text{NTDA}
\end{array}
\tag{8.B.5}
$$

$$
\begin{array}{ccccccccc}
(+z_1, +z_2) & \rightarrow & (0, +z_2) & \rightarrow & (0, 0) & \rightarrow & (+x, 0) & \rightarrow & (+x, -y) \\
\text{NTA} & \rightarrow & \text{TDA} & \rightarrow & \text{TA} & \rightarrow & \text{TDA} & \rightarrow & \text{NTDA}
\end{array}
\tag{8.B.6}
$$

(In the above, "NTA" stands for "non-trivial agreement", "NTDA" stands for "non-trivial disagreement", "TA" stands for "trivial agreement", and "TDA" stands for "trivial disagreement".) We can get the remaining paths by switching e.g. the roles of z_1 and z_2, or of x and y. I don't believe that doing so alters any of the following conclusions.

In figure 8.B.5a, all of the unknown values ("?") that depend on how the magnitudes of z_1, z_2, x, and y compare correspond to the following shifts (zero-step and one-step paths):

$$
\begin{array}{ccc}
(+z_1, +z_2) & \rightarrow & (+x, +z_2 \\
\text{NTA} & \rightarrow & \text{NTA}
\end{array}
\tag{8.B.7}
$$

---

[29]I.e. not necessarily additively, but at least monotonically.
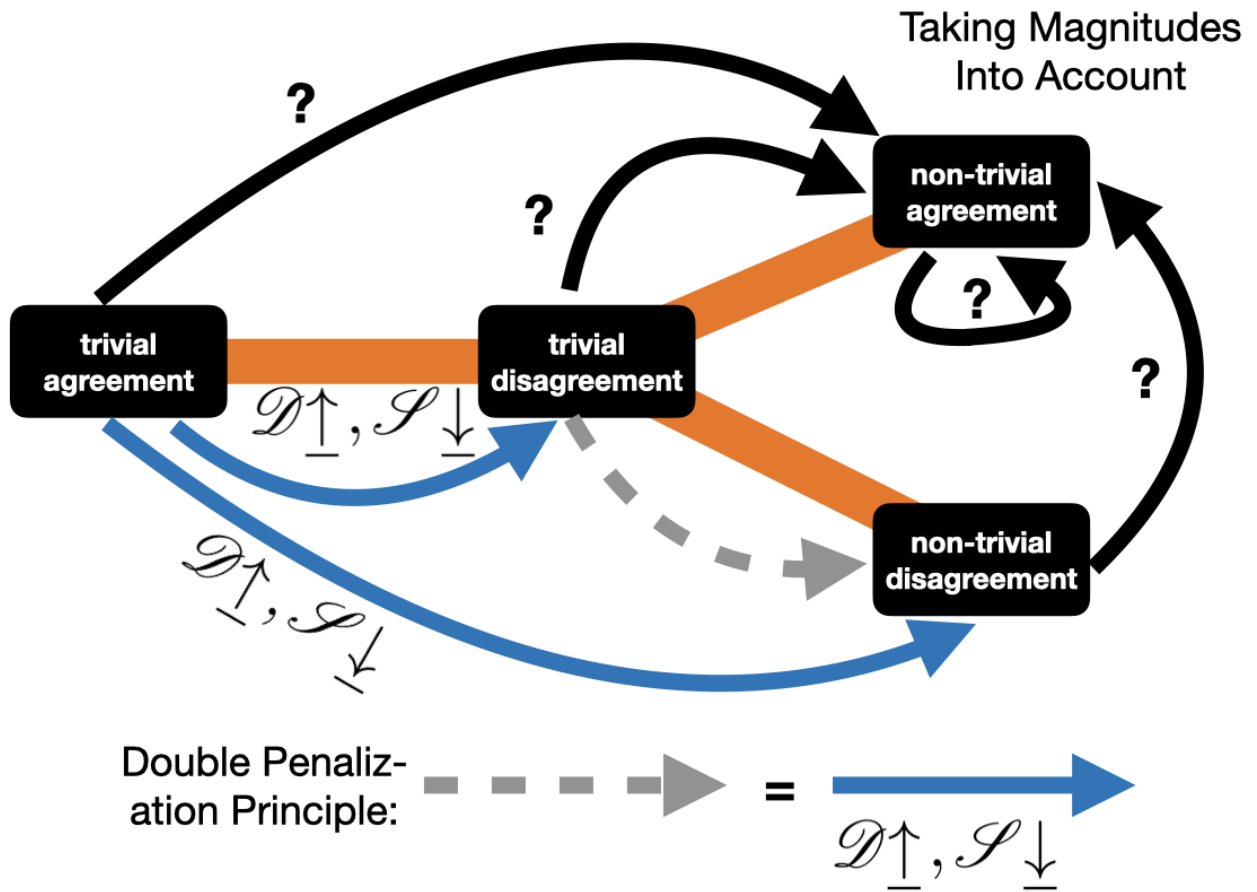
(a) Generalized "composability" criteria for weighted and signed networks.
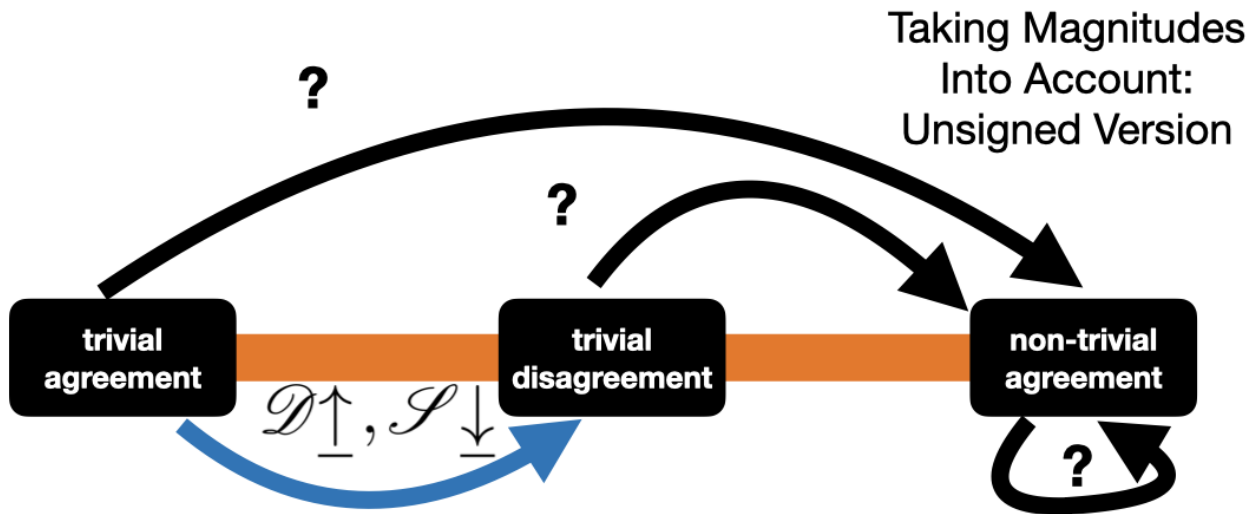
(b) Generalized "composability" criteria that already apply for weighted and *un*signed networks.

Figure 8.B.5: Generalized "composability" criteria for weighted networks.

$$(+z_1, +z_2) \quad \rightarrow \quad (0, +z_2)$$
$$\text{NTA} \quad \rightarrow \quad \text{TDA} \tag{8.B.8}$$

$$(+x, +z_2) \quad \rightarrow \quad (+x, 0)$$
$$\text{NTA} \quad \rightarrow \quad \text{TDA} \tag{8.B.9}$$

How to weight the changes in the objective function value corresponding to each shift, and whether the changes are increases or decreases in penalty, depends on what the specific magnitudes are, and which functions we choose for computing penalties as a function of magnitude differences. These are what determine the "?" arrows in figure 8.B.5a.

None of these choices have anything to do with the double penalization principle. The double penalization principle only dictates the behavior of the arrows on the bottom of figure 8.B.5a, namely that they should not decrease dissimilarity. This amounts to making the gray dashed arrow, which corresponds to both $(+x, 0) \rightarrow (+x, -y)$ and $(0, -y) \rightarrow (+x, -y)$, another blue, similarity non-increasing arrow, in addition to the arrows corresponding to $(0, 0) \rightarrow (+x, -y)$, $(0, 0) \rightarrow (+x, 0)$, and $(0, 0) \rightarrow (0, y)$.

In fact all of the uncertain arrows correspond to choices that already have to be made for purely unsigned networks. Cf. figure 8.B.5b. Thus, if we are already starting with a choice of objective function for weighted unsigned networks, we are already mostly done as long as we impose the "composability" constraint. In other words, in order to determine the "?" arrow from non-trivial disagreements to non-trivial agreements (and reverse) for weighted and signed networks, i.e. the shifts $(+z_1, +z_2) \leftrightarrow (+x, -y)$, all we need to do is:

- settle on a choice of comparison function for unsigned weighted networks to start with – this settles the other three "?" arrows (again cf. figure 8.B.5b),

- ensure that the gray dashed arrow at the bottom of figure 8.B.5a is determined (most likely by insisting that the double penalization principle be satisfied),

- and require the "composability" constraint from section 8.B.6 above be satisfied.

At *no* point do we ever have to explicitly or directly compare penalties for non-trivial agreements and non-trivial disagreements, because their implicit relative weighting follows from the three ingredients above, all of which we would have already demanded anyway.

Hence, whether the double penalization principle is satisfied is not directly relevant for determining implicit relative weighting of penalties for NTAs and and NTDAs.

What instead determines that implicit relative weighting is the original comparison function for unsigned and unweighted networks that we start with, as well as how we ensure that monotonic "composability" is satisfied. We should not need to *directly* compare NTA and NTDA penalties if we are starting from and extending an objective function for unsigned networks. More concretely:

- the weighting of the NTA penalties relative to trivial agreements and disagreements should be determined by the original objective function for weighted, unsigned networks that is being extended, and

- the weighting of the NTDA penalties relative to trivial agreements and disagreements should be determined by the means of satisfying the double penalization principle.

Then whatever ensures "composability" will then step in to use the above two points to implicitly determine the relative weighting of the NTA and NTDA penalties.

Of course, the devil is in the details. One would usually hope that monotonic "composability" can be ensured for the signed extension in the same way it was ensured for the unsigned version, thus requiring no extra work. However, given how vaguely "composability" is "defined" herein, this guidance is currently at best of limited use to a practitioner. To be more unambiguous, in the case of entrywise-L_1 relative error and Jaccard similarity, this monotone "composability" follows from the additivity of the expressions in the numerators. I am not even entirely certain it is satisfied for the other objectives, and may not be. Still, I hope the overall guiding principle is still at least heuristically useful.

It seems that the compatibility of the double penalization principle with the "composability" constraint implies that it can be satisfied without making any explicit choice of "relative weighting" between (i) penalties for sign differences and (ii) penalties for magnitude differences. In contrast, the fact that the sparsity-savviness principle does require such a choice of "relative weighting" seems to then imply its lack of inherent compatibility with the "composability" constraint (roughly the contrapositive statement). Again, as mentioned before, I am currently not sure how to make these conjectures precise enough to falsify.

8.B.8 Precise Formulation of Double Penalization Principle

Given a network with mixed-sign edge weights $\mathcal{G}$, as well as a subset $E \subseteq \mathcal{E}(\mathcal{G})$, define the "**erased edges** network" or "zeroed out edges network" $\mathcal{O}_E(\mathcal{G})$ as the network with mixed-sign edge weights such that

$$\mathcal{N}(\mathcal{O}_E(\mathcal{G})) = \mathcal{N}(\mathcal{G}), \quad \mathcal{E}(\mathcal{O}_E(\mathcal{G})) = \mathcal{E}(\mathcal{G}) \setminus E, \quad A_{\mathcal{O}_E(\mathcal{G})} := \begin{cases} A_{\mathcal{G}}(e) & e \notin E, \\ 0 & e \in E \end{cases}. \quad (8.B.10)$$

Sometimes it helps to think of $\mathcal{O}_E(\mathcal{G})$ as a "sub-network" of the original network $\mathcal{G}$. It can also be useful to think of $\mathcal{O}_E(\mathcal{G})$ as a "sparsifying perturbation" of the original network $\mathcal{G}$.

A "**non-trivial sign disagreement**" between two networks with mixed-sign edge weights $\mathcal{G}_1, \mathcal{G}_2$ occurs for a given edge $e \in \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2)$ if and only if $\text{sign}(A_{\mathcal{G}_1}(e)) \neq \text{sign}(A_{\mathcal{G}_2}(e))$. Recall how, *by definition*, that an edge e is in the edge set $\mathcal{E}(\mathcal{G})$ of a given network $\mathcal{G}$, $e \in \mathcal{E}(\mathcal{G})$, if and only if $A_{\mathcal{G}}(e) \neq 0$. In particular, a "**trivial sign disagreement**" between two networks with mixed-sign edge weights $\mathcal{G}_1, \mathcal{G}_2$ occurs for a given edge e if and only if $\text{sign}(A_{\mathcal{G}_1}(e)) \neq \text{sign}(A_{\mathcal{G}_2}(e))$ and either $\text{sign}(A_{\mathcal{G}_1}(e)) = 0$ or $\text{sign}(A_{\mathcal{G}_2}(e)) = 0$, *but not both*, i.e. either $e \notin \mathcal{E}(\mathcal{G}_1)$ or $e \notin \mathcal{E}(\mathcal{G}_2)$ (but not both). Similarly, a "**trivial sign agreement**"

between two networks with mixed-sign edge weights $\mathcal{G}_1, \mathcal{G}_2$ occurs for a given edge e if and only if both $e \notin \mathcal{E}(\mathcal{G}_1)$ *and* $e \notin \mathcal{E}(\mathcal{G}_2)$, i.e. $\mathrm{sign}(A_{\mathcal{G}_1}(e)) = 0 = \mathrm{sign}(A_{\mathcal{G}_2}(e))$. Finally, a "**non-trivial sign agreement**" between two networks with mixed-sign edge weights $\mathcal{G}_1, \mathcal{G}_2$ occurs for a given edge e if and only if $e \in \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2)$ and $\mathrm{sign}(A_{\mathcal{G}_1}(e)) = \mathrm{sign}(A_{\mathcal{G}_2}(e))$.

Observe how, given a non-trivial sign disagreement for an edge $e \in \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2)$, there are *two* erasures of e that would replace the non-trivial sign disagreement with a trivial sign disagreement, namely (1) erasing e in $\mathcal{G}_1$ but not in $\mathcal{G}_2$, and (2) erasing e in $\mathcal{G}_2$ but not in $\mathcal{G}_1$. The motivating idea behind the "Double Penalization Principle" is that we should never "reward" non-trivial sign disagreements over trivial sign disagreements. The penalty associated with a non-trivial sign disagreement should be no less severe than either of the two penalties that are associated with the two trivial sign disagreements resulting from erasing the corresponding edge. This idea is illustrated in figure 8.2 above.

For two networks $\mathcal{G}_1$ and $\mathcal{G}_2$ with mixed-sign edge weights, their "**set of non-trivial sign disagreements**" $\mathscr{E}_\pm(\mathcal{G}_1, \mathcal{G}_2) \subseteq \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2)$ is the set of "shared" edges such that

$$\mathscr{E}_\pm(\mathcal{G}_1, \mathcal{G}_2) := \{e \subseteq \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2) : \mathrm{sign}(A_{\mathcal{G}_1}(e)) \neq \mathrm{sign}(A_{\mathcal{G}_2}(e))\}. \qquad (8.B.11)$$

A dissimilarity function $\mathscr{D}$ satisfies the "weak double penalization principle" (figure 8.B.6, contrast figure 8.2 above) if and only if, given two networks with mixed-sign edge weights $\mathcal{G}_1$ and $\mathcal{G}_2$, for every pair of subsets $E_1, E_2 \subseteq \mathscr{E}_\pm(\mathcal{G}_1, \mathcal{G}_2)$ with $E_1 \cap E_2 = \emptyset$, one has that

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) \geq \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)), \qquad (8.B.12)$$

i.e. replacing non-trivial sign disagreements with trivial sign disagreements can only decrease the dissimilarity, but *never* increase it. Observe that the condition that $E_1 \cap E_2 = \emptyset$ is what enforces that we are always replacing non-trivial sign disagreements with trivial sign disagreements. If we removed the condition that $E_1 \cap E_2 = \emptyset$, we would also be including the replacement of non-trivial sign disagreements by trivial sign agreements.

Note that based on the weak double penalization property being satisfied alone it is a priori unclear for which $E_1, E_2 \subseteq \mathscr{E}_\pm(\mathcal{G}_1, \mathcal{G}_2)$, $E_1 \cap E_2 = \emptyset$ the minimum

$$\min_{\substack{E_1, E_2 \subseteq \mathscr{E}_\pm(\mathcal{G}_1, \mathcal{G}_2) \\ E_1 \cap E_2 = \emptyset}} \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)) \qquad (8.B.13)$$

is satisfied, although a minimizer must exist due to the finiteness of the optimization space.

Given two networks $\mathcal{G}_1, \mathcal{G}_2$ with mixed-sign edge weights, their "set of trivial sign disagreements" is $\mathcal{E}(\mathcal{G}_1) \Delta \mathcal{E}(\mathcal{G}_2)$, i.e. $(\mathcal{E}(\mathcal{G}_1) \setminus \mathcal{E}(\mathcal{G}_2)) \sqcup (\mathcal{E}(\mathcal{G}_2) \setminus \mathcal{E}(\mathcal{G}_1))$. Then a dissimilarity function $\mathscr{D}$ is called "**monotonic**" if and only if, for any two networks with mixed-sign edge weights $\mathcal{G}_1, \mathcal{G}_2$, and every $E_1 \subseteq \mathcal{E}(\mathcal{G}_1) \setminus \mathcal{E}(\mathcal{G}_2)$ and every $E_2 \subseteq \mathcal{E}(\mathcal{G}_2) \setminus \mathcal{E}(\mathcal{G}_1)$, one has

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) \geq \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)), \qquad (8.B.14)$$

i.e. replacing trivial sign disagreements with trivial sign agreements can only decrease the dissimilarity, but *never* increase it.
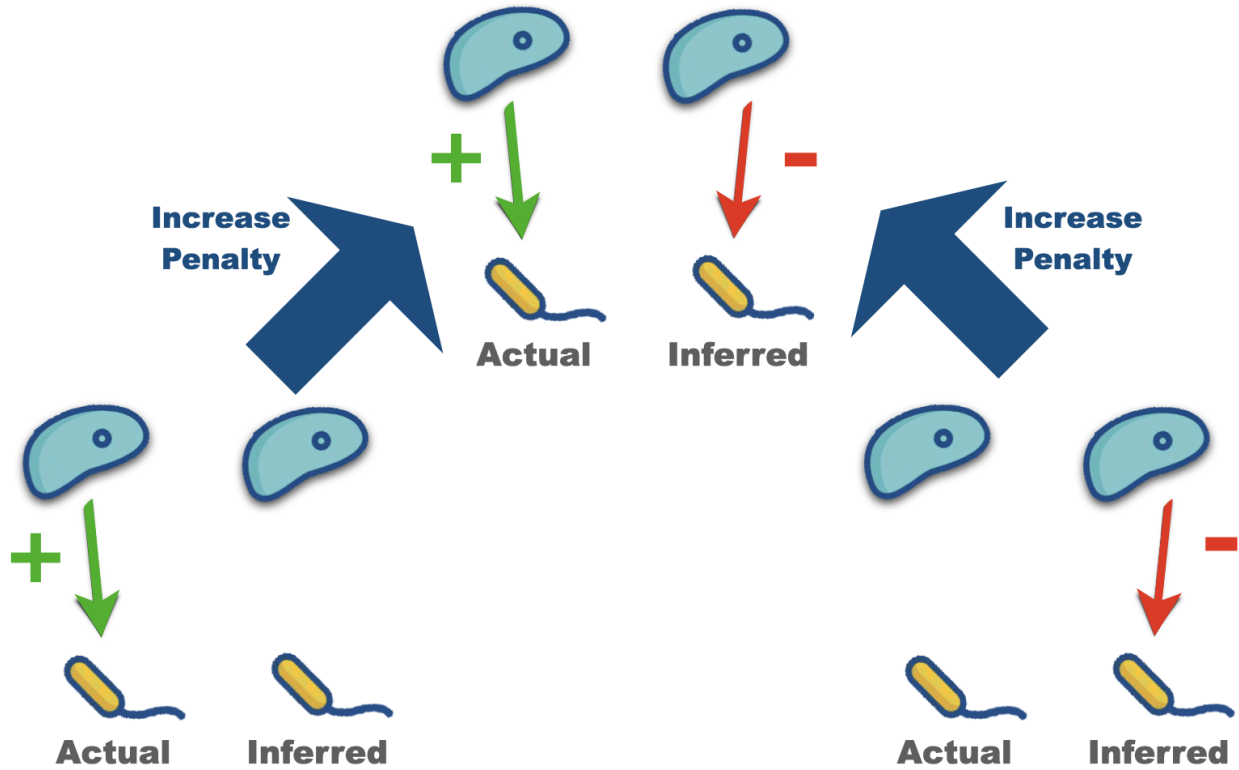
Figure 8.B.6: Schematic for weak double penalization principle.

Observe how, when considering unsigned (weighted) networks as special cases of networks with mixed-sign edge weights such that all nonzero edges are positive, all sign disagreements are trivial sign disagreements. Moreover, the monotonicity property is obviously desirable in that case. The monotonicity property says that replacing true misses with either false hits or false misses (cf. terminology from section 8.4.3.1) should always increase the dissimilarity. This interpretation remains valid in the case of general networks with mixed-sign edge weights, because trivial sign agreements always correspond to true misses, and trivial sign disagreements always correspond to false hits or false misses.

The monotonicity property implies (is it also equivalent to?) the relationship

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) = \max_{\substack{E_1 \subseteq \mathcal{E}(\mathcal{G}_1) \backslash \mathcal{E}(\mathcal{G}_2) \\ E_2 \subseteq \mathcal{E}(\mathcal{G}_2) \backslash \mathcal{E}(\mathcal{G}_1)}} \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)). \tag{8.B.15}$$

Claim: The monotonicity property implies (is it also equivalent to?)

$$\mathscr{D}(\mathcal{O}_{\mathcal{E}(\mathcal{G}_1) \backslash \mathcal{E}(\mathcal{G}_2)}(\mathcal{G}_1), \mathcal{O}_{\mathcal{E}(\mathcal{G}_2) \backslash \mathcal{E}(\mathcal{G}_1)}(\mathcal{G}_2)) = \min_{\substack{E_1 \subseteq \mathcal{E}(\mathcal{G}_1) \backslash \mathcal{E}(\mathcal{G}_2) \\ E_2 \subseteq \mathcal{E}(\mathcal{G}_2) \backslash \mathcal{E}(\mathcal{G}_1)}} \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)). \tag{8.B.16}$$

A dissimilarity function $\mathscr{D}$ is said to satisfy the "(strong) **double penalization principle**" if and only if, (1) $\mathscr{D}$ is monotonic, and (2) given any two networks with mixed-sign

edge weights $\mathcal{G}_1$ and $\mathcal{G}_2$, for *every* pair of subsets $E_1, E_2 \subseteq \mathscr{E}_{\pm}(\mathcal{G}_1, \mathcal{G}_2)$ (i.e. possibly but not necessarily $E_1 \cap E_2 \neq \emptyset$) one has that

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) \geq \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)), \tag{8.B.17}$$

i.e. replacing non-trivial sign disagreements with either trivial sign disagreements *or* trivial sign agreements can only decrease the similarity, but *never* increase it[30].

Claim: A dissimilarity function $\mathscr{D}$ satisfies the (strong) double penalization principle if and only if $\mathscr{D}$ both is monotonic and satisfies the weak double penalization principle.

More important is the following:

> **Claim:** A dissimilarity function $\mathscr{D}$ satisfies the (strong) double penalization principle *if and only if* $\mathscr{D}$ satisfies the monotonicity property on both the positive and negative parts simultaneously.

To the extent that the monotonicity property is highly motivated, perhaps even "trivially" so, for unsigned networks, the truth of the above claim would then imply that the (strong) double penalization principle is correspondingly strongly motivated for signed networks.

Observe how the double penalization principle is agnostic about whether (i) [the dissimilarity increase for replacing a trivial sign agreement with a trivial sign disagreement] is larger or smaller than (ii) [the dissimilarity increase for replacing a trivial sign disagreement with a non-trivial sign disagreement], i.e. it is agnostic about the "relative weighting" of the dissimilarity increases associated with each kind of replacement. The double penalization principle only requires that both kinds of replacement do not decrease the dissimilarity. Combined together, both requirements then imply that replacing a trivial sign agreement with a non-trivial sign disagreement also not decrease the dissimilarity.

Non-trivial sign disagreements correspond to a false hit or false miss for both the positive parts and the negative parts of the networks (cf. the sections 8.3.4 and 8.4.3.1 for relevant terminology). Thus the double penalization principle requires that a false hit/miss for the positive parts of the networks can not be combined with a false miss/hit for the negative parts of the network to decrease the dissimilarity.

The double penalization property implies (is it also equivalent to?) the relationship

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) = \max_{E_1, E_2 \subseteq \mathscr{E}_{\pm}(\mathcal{G}_1, \mathcal{G}_2)} \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)). \tag{8.B.18}$$

Claim: the double penalization property implies (is it also equivalent to?)

$$\mathscr{D}(\mathcal{O}_{\mathscr{E}_{\pm}(\mathcal{G}_1, \mathcal{G}_2)}(\mathcal{G}_1), \mathcal{O}_{\mathscr{E}_{\pm}(\mathcal{G}_1, \mathcal{G}_2)}(\mathcal{G}_2)) = \min_{E_1, E_2 \subseteq \mathscr{E}_{\pm}(\mathcal{G}_1, \mathcal{G}_2)} \mathscr{D}(\mathcal{O}_{E_1}(\mathcal{G}_1), \mathcal{O}_{E_2}(\mathcal{G}_2)). \tag{8.B.19}$$

---

[30] Without the possible loophole of replacing trivial sign disagreements with trivial sign agreements nevertheless somehow being able to increase the dissimilarity, which I think might technically be permitted if we only required (8.B.17) without also explicitly requiring $\mathscr{D}$ to be monotonic, albeit I'm not certain.

For two networks $\mathcal{G}_1$ and $\mathcal{G}_2$ with mixed-sign edge weights, their "**set of non-trivial sign agreements**" $\mathscr{E}_=(\mathcal{G}_1, \mathcal{G}_2) \subseteq \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2)$ is the set of "shared" edges such that

$$\mathscr{E}_=(\mathcal{G}_1, \mathcal{G}_2) := \{ e \subseteq \mathcal{E}(\mathcal{G}_1) \cap \mathcal{E}(\mathcal{G}_2) : \text{sign}(A_{\mathcal{G}_1}(e)) = \text{sign}(A_{\mathcal{G}_2}(e)) \} . \tag{8.B.20}$$

A *similarity* function $\mathscr{S}$ is "**sparsity-savvy**" if and only if, for any given networks with mixed-sign edge weights $\mathcal{G}_1$ and $\mathcal{G}_2$, for every subset $E \subseteq \mathscr{E}_=(\mathcal{G}_1, \mathcal{G}_2)$ one has that

$$\mathscr{S}(\mathcal{G}_1, \mathcal{G}_2) \geq \mathscr{S}(\mathcal{O}_E(\mathcal{G}_1), \mathcal{O}_E(\mathcal{G}_2)), \tag{8.B.21}$$

i.e. that replacing a trivial sign agreement with a non-trivial sign agreement at most only increase the similarity, but *never* decrease it. Cf. section 8.2.7. In particular we can't "game" the similarity by considering only sparse networks, or removing common edges, or "embedding networks" into spaces with higher numbers of nodes (i.e. adding completely disconnected nodes, or adding blocks of all 0's to the adjacency matrices). Keep in mind that this property is not actually entirely trivial, at least for weighted networks. This is because the difference in magnitudes for a trivial sign agreement is always 0, whereas the difference in magnitudes for a non-trivial sign agreement can be arbitrarily large. So the requirement is actually substantial/not to be taken for granted.

When considering unsigned (weighted) networks as special cases of signed networks whose nonzero edges are positive, there are no non-trivial sign disagreements, i.e. all sign disagreements are trivial sign disagreements. The sparsity-savviness principle says that replacing true hits with true misses should never increase the similarity. Such an interpretation remains valid in the case of general signed (weighted) networks. Even then non-trivial sign agreements always correspond to true hits for either the positive subnetworks or for the negative subnetworks (cf. the definitions in sections 8.3.4 and 8.4.3.1).

	sign agreement	sign disagreement
trivial	$\{(TM_+, TM_-)\}$	$\{(FH_+, TM_-), (TM_+, FH_-), (FM_+, TM_-), (TM_+, FM_-)\}$
non-trivial	$\{(TH_+, TM_-), (TM_+, TH_-)\}$	$\{(FH_+, FM_-), (FM_+, FH_-)\}$

For the sake of clarity and comparison, let's consider the analogous binary classification notions for unsigned networks:

	sign agreement	sign disagreement
trivial	$\{(TM)\}$	$\{(FH), (FM)\}$
non-trivial	$\{(TH)\}$	do not exist

Note that the characteristics for unsigned networks should emerge as a special case of those for general signed networks, and indeed they do.

Definition: Given $S \times S$ matrices, a **distance metric** $d : \mathbb{R}^{[S] \times [S]} \times \mathbb{R}^{[S] \times [S]} \to [0, \infty)$ is a function satisfy the axioms (1) for any matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ one has $d(\mathbf{M}_1, \mathbf{M}_2) = 0$ if and only if $\mathbf{M}_1 = \mathbf{M}_2$, (2) $d(\mathbf{M}_1, \mathbf{M}_2) = d(\mathbf{M}_2, \mathbf{M}_1)$ for any matrices $\mathbf{M}_1$ and $\mathbf{M}_2$, and

(3) $d(\mathbf{M}_1, \mathbf{M}_3) \leq d(\mathbf{M}_1, \mathbf{M}_2) + d(\mathbf{M}_2, \mathbf{M}_3)$ for any matrices $\mathbf{M}_1$, $\mathbf{M}_2$, and $\mathbf{M}_3$. The distance metric is said to turn $\mathbb{R}^{[S] \times [S]}$ into a "metric space".

Claim: Given a dissimilarity function $\mathscr{D}$ such that for any $\mathcal{G}_1, \mathcal{G}_2 \in \mathscr{G}_{[S]}$ one has that

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) = d(\mathbf{A}_{\mathcal{G}_1}, \mathbf{A}_{\mathcal{G}_2})$$

for some distance metric d on the space of matrices, then $\mathscr{D}$ (a) is monotone, and (b) satisfies the (strong) double penalization property.

In particular, when using dissimilarity functions directly derived from distance metrics on arbitrary matrices, we don't have to worry much. This isn't necessarily true if the distance metric is only defined for non-negative matrices (corresponding to unsigned networks).

Definition: A **distance metric** d on the space of unsigned networks with known node correspondences $\mathscr{G}_{[S]}^{\geq 0}$ satisfies the axioms (1) for any unsigned networks $\mathcal{U}_1$ and $\mathcal{U}_2$ one has $d(\mathcal{U}_1, \mathcal{U}_2) = 0$ if and only if $\mathcal{U}_1 = \mathcal{U}_2$, (2) $d(\mathcal{U}_1, \mathcal{U}_2) = d(\mathcal{U}_2, \mathcal{U}_1)$ for any unsigned networks $\mathcal{U}_1$ and $\mathcal{U}_2$, and (3) $d(\mathcal{U}_1, \mathcal{U}_3) \leq d(\mathcal{U}_1, \mathcal{U}_2) + d(\mathcal{U}_2, \mathcal{U}_3)$ for any unsigned networks $\mathcal{U}_1$, $\mathcal{U}_2$, and $\mathcal{U}_3$.

Claim: Given a dissimilarity function $\mathscr{D}$ such that for any $\mathcal{G}_1, \mathcal{G}_2$ one has that

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) = d(\mathfrak{M}(\mathcal{G}_1), \mathfrak{M}(\mathcal{G}_2))$$

for some distance metric d on unsigned networks, where $\mathfrak{M}(\mathcal{G})$ denotes the magnitude skeleton of $\mathcal{G}$, then in general $\mathscr{D}$ does **not** satisfy[31] the double penalization principle.

Even starting with the most "well-behaved" dissimilarity functions on unsigned networks, dissimilarity functions for general networks with mixed-sign edge weights derived from the "projection procedure" will not satisfy this basic "well-behavedness" property.

However, we should still have the following

Claim: Given a dissimilarity function $\mathscr{D}$ such that for any $\mathcal{G}_1, \mathcal{G}_2$ one has that

$$\mathscr{D}(\mathcal{G}_1, \mathcal{G}_2) = c_1 d(\mathcal{G}_1^+, \mathcal{G}_2^+) + c_2 d(\mathcal{G}_1^-, \mathcal{G}_2^-)$$

for $c_1, c_2 \geq 0$, c_1 and c_2 are allowed to depend on $\mathcal{G}_1$ and $\mathcal{G}_2$, and d is a distance metric on unsigned networks, then $\mathscr{D}$ *will* satisfy the double penalization principle.

The above claim only applies to conic (i.e. non-negative) combinations of dissimilarity functions. However I believe that the analogous claim for similarity functions is true as well (or at least have been implicitly assuming as much):

Claim: Given a $\mathscr{S}$ function such that for any $\mathcal{G}_1, \mathcal{G}_2$ one has that

$$\mathscr{S}(\mathcal{G}_1, \mathcal{G}_2) = c_1 \mathcal{S}(\mathcal{G}_1^+, \mathcal{G}_2^+) + c_2 \mathcal{S}(\mathcal{G}_1^-, \mathcal{G}_2^-)$$

for $c_1, c_2 \geq 0$, c_1 and c_2 are allowed to depend on $\mathcal{G}_1$ and $\mathcal{G}_2$, and $\mathcal{S}$ is a similarity function on unsigned networks satisfying the (similarity function version of) monotonicity, then $\mathscr{S}$ *will* satisfy the (strong) double penalization property.

---

[31] It may still satisfy monotonicity – honestly I'm not sure either way.

I think the above claim is true, or at the very least it seems I have been implicitly assuming as much. Similarly (pun unintended) it seems I have also been implicitly assuming

Claim: Given a $\mathscr{S}$ function such that for any $\mathcal{G}_1, \mathcal{G}_2$ one has that

$$\mathscr{S}(\mathcal{G}_1, \mathcal{G}_2) = c_1 \mathcal{S}(\mathcal{G}_1^+, \mathcal{G}_2^+) + c_2 \mathcal{S}(\mathcal{G}_1^-, \mathcal{G}_2^-)$$

for $c_1, c_2 \geq 0$, c_1 and c_2 are allowed to depend on $\mathcal{G}_1$ and $\mathcal{G}_2$, and $\mathcal{S}$ is a similarity function on unsigned networks satisfying the sparsity-savviness property, then $\mathscr{S}$ *will* satisfy the sparsity-savviness property.

8.B.9 Continuity with respect to Magnitude

Consider the plane ($\mathbb{R}^2$) with coordinates denoted (a, i), "a" for "actual", "i" for "inferred". The first coordinate corresponds to the weight of the edge of the actual network in a single-edge comparison, whereas the second coordinate corresponds to the weight of the edge of the inferred network in a single-edge comparison.

Then we can divide/partition the plane into 4 regions:

1.
$$\{(a, i) : (a > 0 \text{ and } i > 0) \text{ or } (a < 0 \text{ and } i < 0)\}. \qquad (8.B.22)$$

   These are the non-trivial sign agreements. Geometrically on the plane, these correspond to the interiors of the upper right and lower left quadrants.

2.
$$\{(a, i) : (a > 0 \text{ and } i < 0) \text{ or } (a < 0 \text{ and } i > 0)\}. \qquad (8.B.23)$$

   These are the non-trivial sign disagreements. Geometrically on the plane these correspond to the interiors of the upper left and lower right quadrants.

3.
$$\{(a, i) : (a = 0 \text{ and } i \neq 0) \text{ or } (a \neq 0 \text{ and } i = 0)\}. \qquad (8.B.24)$$

   These are the trivial sign disagreements. Geometrically on the plane these correspond to the union of the a and i axes with the origin removed.

4.
$$\{(a, i) : a = 0 \text{ and } i = 0\} = \{(0, 0)\}. \qquad (8.B.25)$$

   These are the trivial sign agreements. Geometrically on the plane these correspond to the origin.

Figure 8.B.7 illustrates this heuristically:

Figure 8.B.7: The "single-edge comparison plane" described in the main text.

Now consider any continuous function $[0,1] \to \mathbb{R}^2$, $t \mapsto (a(t), i(t))$, a so-called *path*. For any such path, we can look at the sequence of above regions that it passes through. Close inspection reveals that the set of all such allowable sequences corresponds to the set of all possible paths (in the graph-theoretic sense) in the undirected network from figure 8.B.2. (At least when prohibiting diagonal movements in the immediate vicinity of the origin, something which is admittedly better motivated in the unweighted case.)

The continuity/composability property, at least intuitively speaking or heuristically or roughly, corresponds to requiring that for any such path in the "single-edge comparison plane", and any given objective function, its values are continuous as a function of the t-parametrized family of graph comparisons corresponding to any given edge in the graph-comparison, and the start and end points.

Specifically, say we are comparing the actual graph $\mathcal{G}_*$ and the inferred graph $\hat{\mathcal{G}}$. We focus on some particular edge $e \in [S] \times [S]$. This corresponds to a tuple $(a, i) = (e_*, \hat{e})$, where $e_* := A_{\mathcal{G}_*}(e)$ and $\hat{e} := A_{\hat{\mathcal{G}}}(e)$. Then we can choose some path $[0,1] \to \mathbb{R}^2$ with $(e_*, \hat{e})$ as the starting point, i.e. $(a(0), i(0)) = (e_*, \hat{e})$. Then this gives us a parameterized family of tuples of networks, where each $t \in [0,1]$ gets sent to $(\mathcal{G}_*(t), \hat{\mathcal{G}}(t))$, again with $(\mathcal{G}_*(0), \hat{\mathcal{G}}(0)) = (\mathcal{G}_*, \hat{\mathcal{G}})$, and where $A_{\mathcal{G}_*(t)}(e) = a(t)$ and $A_{\hat{\mathcal{G}}(t)}(e) = i(t)$ (the values of the adjacency function are assumed unchanged for all other edges). Composing this parameterized family of tuples of

networks with our objective function, we are asking for the composite function $[0,1] \to \mathbb{R}$ to be continuous, irrespective of which edge we choose, and which path $[0,1] \to \mathbb{R}^2$ was chosen.

Call such a path "coordinate-wise non-decreasing" if, for all $t_1 < t_2$ in $[0,1]$, we have both $a(t_1) \le a(t_2)$ *and* $i(t_1) \le i(t_2)$.

Of course, when we consider unweighted network, there is no longer any "continuous plane" of pairs of edge weight values to send paths into. But we still would like to consider "discretized paths" where the only allowed "discretized paths" correspond again to paths (in the graph-theoretic sense) in the undirected network from figure 8.B.2.

Also of course in the case of weighted networks, non-trivial sign agreements are no longer straightforward. In other words, even when the signs (non-trivially) agree, we can now still have differences in magnitude which we would usually want to penalize as well. The only places in the region of non-trivial sign agreements where we don't have magnitude disagreements are along the intersection with the line $a = i$. This corresponds to a line with slope 1. In general, any line with slope bounded in between 0 and $+\infty$ passes through the two quadrants defining the region of non-trivial sign agreements and thus corresponds to a certain departure (possibly null when the slope is 1) from absence of magnitude of disagreement. (Hypothetically we could choose to associate this with the non-negative projective line, but that's probably a topic for another time.) In the limits as the slopes approach either 0 or $+\infty$, we have that the non-trivial sign agreements actually approach trivial sign disagreements. Therefore, continuity combined with the desire to penalize trivial sign disagreements *require us to penalize at least some non-trivial sign agreements*. (Again this is restricted only to the case of weighted networks, which is the only context where non-zero magnitude differences are possible for non-trivial sign agreements.)

While all of this certainly poses some complications for the sparsity-savviness principle, in particular probably requiring that it can only be satisfied in some asymptotic sense which we won't bother to attempt to precisely define, it actually does not pose any challenges or complications for satisfying a (continuous version of) the double penalization principle. To discuss this, it probably makes sense to begin by discussing a continuous version of monotonicity (i.e. comparing trivial sign agreements with trivial sign disagreements in the case of weighted networks). The main idea however is that (the continuous version of[32]) the double penalization principle does not require any constraints on the behavior of paths passing through the quadrants corresponding to non-trivial sign agreements. It only regulates paths which pass from trivial sign-disagreements to non-trivial sign disagreements, to ensure they display analogous monotonicity behavior to paths from trivial sign agreements to trivial sign disagreements. (I.e. how the objective function behaves in the upper right and lower left quadrants is irrelevant.)

Consider any path starting from $(0,0)$ and such that either $a(t) = 0$ or $i(t) = 0$ for all $t \in [0,1]$ (but not both). Then for the coordinate of the path that is not constant and identically 0, assume that it is monotone (non-decreasing or non-increasing, it does not matter which). Then any dissimilarity $\mathscr{D}$ satisfying the continuous monotonicity property will monotonically

---

[32]But then presumably also as a consequence the discretized version

non-decrease along any such path. I.e. making one of the edge counterparts being compared increasingly nonzero should only worsen the score.

Now the double penalization principle is motivated by the fact that, when looking at any point in the non-trivial sign disagreement quadrants (lower right or upper left), projecting the point onto the *closed* quadrant $\{a \geq 0 \text{ and } i \geq 0\}$ represents the point as a trivial disagreement for the positive part of the network, whereas projecting the point onto the *closed* quadrant $\{a \leq 0 \text{ and } i \leq 0\}$ portrays it as a trivial disagreement for the negative part of the network. So intuitively speaking, we would like, at least for paths which are suitably coordinate-wise monotone restricted within one of the non-trivial sign disagreement quadrants, for the objective function to "obey monotonicity for both the positive and negative parts", i.e. for a dissimilarity function to monotonically non-decrease. Again, this is not at all a restriction on the behavior in the non-trivial sign agreement quadrants and thus whether the objective satisfies the double penalization principle is completely independent of e.g. the relative weighting of sign errors versus magnitude errors. Cf. the figure 8.B.8a, where the lighter shaded areas indicate the regions that the (strong and continuous) double penalization principle implies should correspond to greater dissimilarity values than the corresponding large white dots.

In the lower right quadrant, a suitable path to which the continuous version of the double penalization principle applies is one for which $a(0) \geq 0$ and $i(0) \leq 0$, and then for all $t_1, t_2 \in [0,1]$ with $t_1 < t_2$, $a(t_1) \leq a(t_2)$ and $i(t_1) \geq i(t_2)$. (If the path coordinate functions are differentiable, then basically we are requiring that $a'(t) \geq 0$ whereas $i'(t) \leq 0$.) Likewise, in the upper left quadrant, suitable paths are those for which $a(0) \leq 0$ and $i(0) \geq 0$ and such that for all $t_1, t_2 \in [0,1]$ with $t_1 < t_2$ we have $a(t_1) \geq a(t_2)$ and $i(t_1) \leq i(t_2)$. (If the path coordinate functions are differentiable, then basically we are requiring that $a'(t) \leq 0$ and $i'(t) \geq 0$.) Then for all such paths, the (continuous version of) the double penalization principle requires that a dissimilarity function monotonically non-decrease with t. (I feel like requiring the objective to be continuous along with requiring the discretized definition above should be equivalent to this continuous version of the double penalization principle, but if it's true the formal argument would be very technical and boring so I at least currently will not bother.) Notice in particular that the continuous version of the double penalization principle does *not* require the dissimilarity to increase (non-decrease) fastest or slowest along a specific choice of this class of paths, only that it be monotonically non-decreasing along all of them. (If we assume the start and end points of all of the paths under consideration are the same, then only the "rate of change" along the paths can differ, because obviously then the net difference between the values of the dissimilarity function at the beginning and end of the path must be the same.)

The figure 8.B.8b indicates examples of such allowable trajectories.

8.B.10   Considerations for Generalized Networks

The arguments and notions of e.g. the monotonicity principle, double penalization principle, or sparsity-savviness principle, do not inherently need the comparison functions be defined
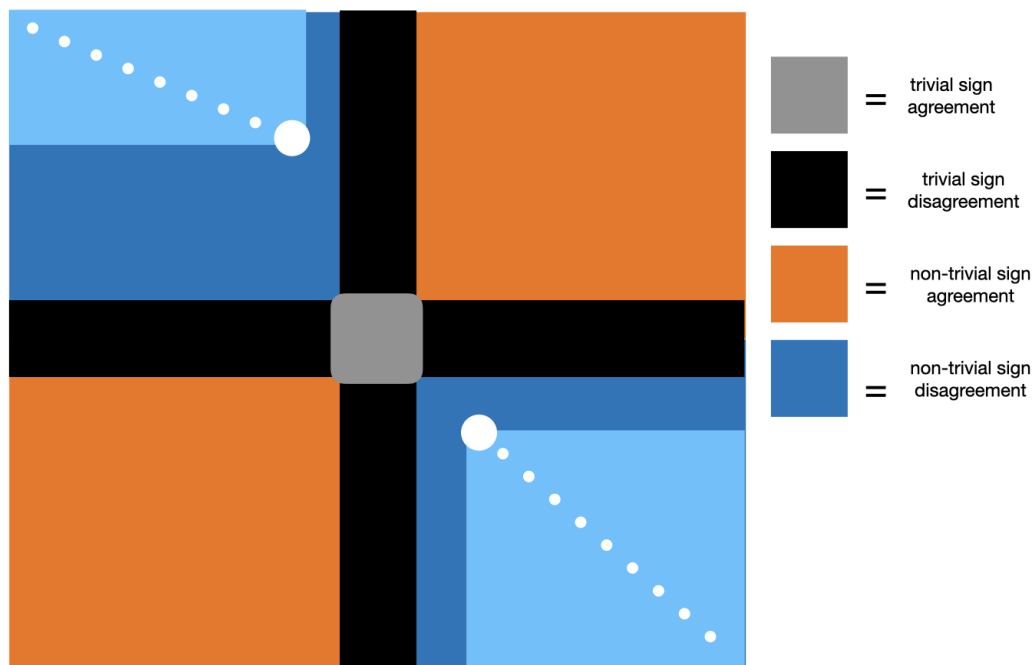
(a) Large white dot has smallest dissimilarity compared to everything in the corresponding lightened regions. Additionally white dots indicate the direction of "increasing isoclines of dissimilarity".
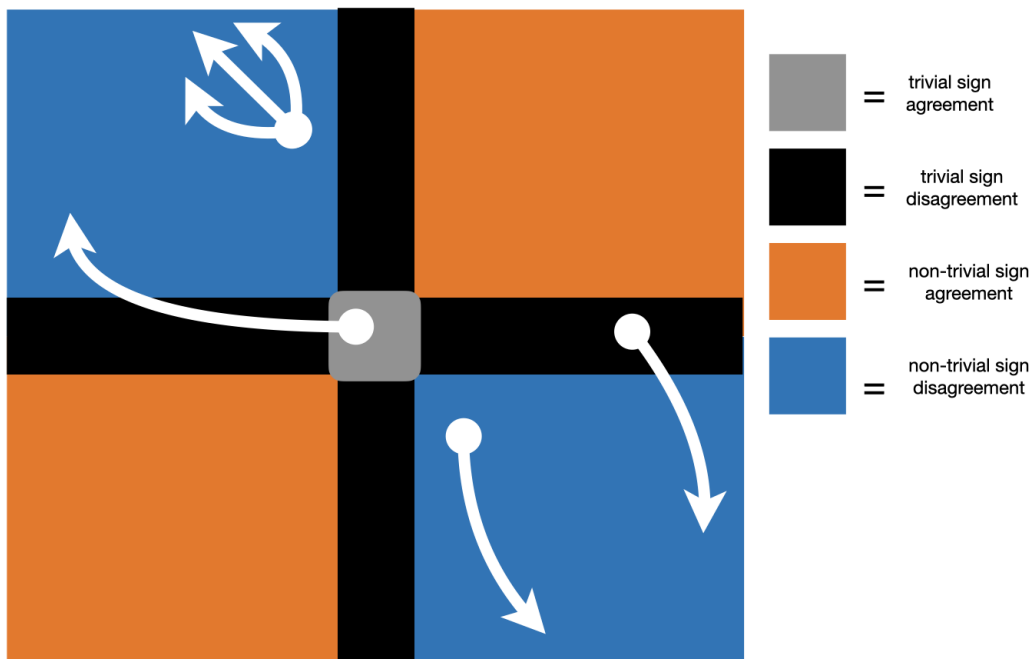
(b) Examples of paths to which the continuous double penalization principle applies. Large white dots are the start points of the paths, arrows indicate the end points and directions of the paths.

Figure 8.B.8: Schematic depiction of continuous version of double penalization principle.

for "discrete", finite networks. To clarify the "true" underlying ideas, what follows is a preliminary attempt to demonstrate their generalizability to other comparison functions. Hopefully this will be a preliminary step towards future work that makes all of the concepts and principles expressed in this chapter precise enough to be falsifiable.

8.B.10.1   Generalized Edge Spaces

Let a locally compact Polish space[33] E correspond to a "generalized edge space".
   Examples:

- For a (finite) network, $E = [S] \times [S]$ (equipped with the usual discrete[34] topology).

- For an infinite network, $E = \mathbb{N} \times \mathbb{N}$ (equipped with the usual discrete topology).

- For a graphon, $E = [0, 1] \times [0, 1]$ (equipped with the usual Euclidean subspace topology).

8.B.10.2   Generalized Edge/Adjacency Functions

Given a "generalized edge space E", i.e. a locally compact Polish space, a "generalized edge (adjacency) function" $A : E \to \mathbb{R}$ will be interpreted as a *continuous* function $E \to \mathbb{R}$.

Note that in the case that the topology of E is discrete, then all functions $E \to \mathbb{R}$ are continuous. However, this is a meaningful restriction e.g. in the graphon case. Analysis extending these notions to general Borel-measurable functions (or showing that such an extension is impossible) is left to future work.

Note that because E is locally compact Polish, and $\mathbb{R}$ is Polish, the space of all continuous functions $C(E, \mathbb{R})$ is itself a Polish space when equipped with the standard compact-open topology and, moreover, is an exponential object in the category of topological spaces (in particular evaluation of functions is continuous).

8.B.10.3   Indicators as Pointwise Limits

Lemma 8.7. *Let (X, τ) be a topological space, $Y \subseteq X$. Assume that the indicator function $\mathbb{I}\{Y\} : X \to [0, 1]$ can be written as the monotone non-decreasing pointwise limit of some sequence of continuous functions $f_n : X \to [0, 1]$, i.e. for all $x \in X$ one has $\lim_{n \to \infty} f_n(x) = \mathbb{I}\{Y\}(x)$ and for all $n_1, n_2 \in \mathbb{N}$, $n_1 \leq n_2$, $f_{n_1}(x) \leq f_{n_2}(x)$. Then Y must be (1) open and (2) F_σ.*

---

[33] Equivalent to a locally compact second-countable Hausdorff space (a.k.a. "LCCB" space), among other characterizations.

[34] Like the discrete topology for $\mathbb{N} \times \mathbb{N}$, this is also the same as the Euclidean subspace topology inherited from $\mathbb{R}^2$.

Proof of Lemma 8.7   First, note that $\{1\}$ and $\{0\}$ are closed subsets of $[0,1]$, whereas as $[0,1)$ and $(0,1]$ are open subsets $[0,1]$.

By definition, the inverse image of $\{1\}$ under $\mathbb{I}\{Y\}$ is exactly the set Y, so the sets of values where the f_n equal 1 must approach Y. Because the f_n are monotone increasing and bounded above by 1, for $n_1 \leq n_2$ we must have $f_{n_1}^{-1}(\{1\}) \subseteq f_{n_2}^{-1}(\{1\})$, i.e. the inverse images of $\{1\}$ under the f_n must be non-decreasing as well and thus "approach Y from the inside". More concretely:

$$Y = \bigcup_{n=1}^{\infty} f_n^{-1}(\{1\}) = \left(\bigcap_{n=1}^{\infty} f_n^{-1}([0,1)) \right)^c , \tag{8.B.26}$$

with the last equality following from DeMorgan's Law. Because all of the f_n are continuous, the $f_n^{-1}(\{1\})$ are all closed and the $f_n^{-1}([0,1))$ are all open. Hence Y is the countable union of closed sets, thus F_σ. (Equivalently Y is the complement of a countable intersection of open sets, i.e. the complement of a G_δ set.)

At the same time, the inverse image of $\{0\}$ under $\mathbb{I}\{Y\}$ is the set Y^c, so the sets of values where the f_n equal 0 must approach Y^c. Because the f_n are monotone increasing and bounded below by 0, for $n_1 \leq n_2$ we must have $f_{n_1}^{-1}(\{0\}) \supseteq f_{n_2}^{-1}(\{0\})$, i.e. the inverse images of $\{0\}$ under the f_n must be non-increasing and thus "approach Y^c from the outside". More concretely:

$$Y = (Y^c)^c = \left(\bigcap_{n=1}^{\infty} f_n^{-1}(\{0\}) \right)^c = \bigcup_{n=1}^{\infty} f_n^{-1}((0,1]) , \tag{8.B.27}$$

with the last equality again following from DeMorgan's Law. Again because all of the f_n are continuous, the $f_n^{-1}(\{0\})$ are all closed and the $f_n^{-1}((0,1])$ are all open. The arbitrary intersection of closed sets is closed, and the arbitrary union of open sets is open, so (8.B.27) means that Y is the complement of a closed set, or equivalently an open set. $\square$

Corollary 3. *Let (X, τ) be a topological space, $Z \subseteq X$. Assume that the indicator function $\mathbb{I}\{Z\} : X \to [0,1]$ can be written as the monotone non-increasing pointwise limit of some sequence of continuous functions $g_n : X \to [0,1]$, i.e. for all $x \in X$ one has $\lim_{n\to\infty} g_n(x) = \mathbb{I}\{Z\}(x)$ and for all $n_1, n_2 \in \mathbb{N}$, $n_1 \leq n_2$, $g_{n_1}(x) \geq g_{n_2}(x)$. Then Z must be (1) closed and (2) G_δ.*

Proof of Corollary 3: Take $Y := Z^c$ and $f_n := 1 - g_n$. These satisfy the hypotheses of Lemma 8.7. Because $Y = Z^c$ is open and F_σ, it follows that its complement Z must be closed and G_δ. $\square$

8.B.10.4 σ-compact subsets

Lemma 8.8. *Let X be a σ-compact and Hausdorff[35] space. Then a subset Y is σ-compact[36] if and only if Y is F_σ.*

Proof of Lemma 8.8 ($\implies$) Assume Y is σ-compact, i.e. $Y = \bigcup_{n=1}^{\infty} \tilde{K}_n$ with all $\tilde{K}_n$ compact. Because the $\tilde{K}_n$ are compact, they are closed. So Y is the countable union of closed sets, i.e. F_σ.

($\impliedby$) Assume Y is F_σ, i.e. $Y = \bigcup_{n=1}^{\infty} F_n$ with all F_n closed. Because X is σ-compact, $X = \bigcup_{m=1}^{\infty} K_m$ with all K_m compact. Then because $X \cap Y = Y$, we have that

$$Y = \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} (K_m \cap F_n).$$

Because the K_m are compact and the F_n are closed, for all m and n we have that $K_m \cap F_n$ is compact. Therefore Y is the countable union of compact sets, i.e. σ-compact. $\square$

8.B.10.5 Relevant Facts

- Locally compact Polish spaces are (by definition) metrizable, and metrizable spaces are perfectly normal (because they are normal and closed sets are G_δ).

- Locally compact and second-countable implies σ-compact, all Polish spaces are second-countable, so locally compact Polish spaces are σ-compact.

- In a metrizable space open sets are[37] F_σ, thus by Lemma 8.8 and the above facts we know that every open subset of a locally compact Polish space is σ-compact.

8.B.10.6 Approximating Open Indicators from Below

Lemma 8.9. *Let U be an open subset of a locally compact Polish space E. Then there exists a monotone non-decreasing sequence of continuous functions $f_n : E \to [0,1]$ such that the indicator $\mathbb{I}\{U\} : E \to [0,1]$ is the pointwise limit, i.e. for all $e \in E$ one has $\lim_{n \to \infty} f_n(e) = \mathbb{I}\{U\}(e)$ and for all $n_1, n_2 \in \mathbb{N}$, $n_1 \le n_2$, $f_{n_1}(e) \le f_{n_2}(e)$. (Using Lemma 8.7 it follows that the existence of such a sequence exactly characterizes open subsets of E.)*

---

[35] Hausdorffness is technically somewhat overkill. Because a closed subset of a compact set is always compact, all we need is for all compact sets to be closed (a so-called "KC space"), and even then only for one direction. That being said, Hausdorffness is a much weaker assumption than the assumption that E be a locally compact Polish space.

[36] Because compactness is an "intrinsic" topological property, i.e. $Z \subseteq X$ is compact in its subspace topology if and only if it is compact in the "ambient" topology of X, the same is also true of σ-compactness. For other properties we would need to be more careful and specify whether we meant in the subspace topology of Y or in the "ambient" topology of X.

[37] This is dual to the aforementioned fact that closed sets are G_δ.

Proof of Lemma 8.9 This seems to only[38] require that E be normal and that all open subsets be F_σ. First, using the fact that U open is F_σ, write $U = \bigcup_{n=1}^{\infty} F_n$ with all F_n closed[39]. Next follows an induction argument.

For the base case where $n = 1$, note that the fact that E is perfectly normal means Urysohn's lemma can be applied to the closed sets F_1 and U^c in such a way to give a continuous function $f_1 := g_1 : E \to [0, 1]$ such that $g_1(e) = 0$ for all $e \in U^c$, $g_1(e) = 1$ for all $e \in F_1$, and (this is the part requiring E to be perfectly normal) $0 < g_1(e) < 1$ for all $e \in U \setminus F_1$. Because there are no n smaller than 1, the monotonicity requirement is vacuously satisfied.

Now assume for $n \geq 1$ the induction hypothesis holds that for all $m \leq n$ there exists a continuous $f_m : E \to [0, 1]$ such that $f_m(e) = 0$ for all $e \in U^c$, $f_m(e) = 1$ for all $e \in F_m$, $0 < f_m(e) < 1$ for all $e \in U \setminus F_m$, and that for all $n_1, n_2 \leq n$ with $n_1 \leq n_2$ that one has for all $e \in E$ that $f_{n_1}(e) \leq f_{n_2}(e)$. To prove the claim, we now need to show that there exists a continuous $f_{n+1} : E \to [0, 1]$ such that $f_{m+1} = 0$ for all $e \in U^c$, $f_{n+1}(e) = 1$ for all $e \in F_{n+1}$, $0 < f_{n+1}(e) < 1$ for all $e \in U \setminus F_{n+1}$, and that for all $m \leq n$ that for all $e \in E$ one has $f_m(e) \leq f_{n+1}(e)$.

Because E is perfectly normal, again Urysohn's lemma can be applied to the closed sets F_{n+1} and U^c in such a way to give a continuous function $g_{n+1} : E \to [0, 1]$ such that $g_{n+1}(e) = 0$ for all $e \in U^c$, $g_{n+1}(e) = 1$ for all $e \in F_{n+1}$, and $0 < g_{n+1}(e) < 1$ for all $e \in U \setminus F_{n+1}$. Now define $f_{n+1} := \max\{g_{n+1}, f_n\}$. As the maximum of two continuous functions, f_{n+1} is itself continuous. Likewise, because for all $e \in U^c$ we have that $f_n(e) = g_{n+1}(e) = 0$, we also have that $f_{n+1}(e) = 0$ for all $e \in U^c$. Because $f_n(e) \leq 1$ for all $e \in E$, and $g_{n+1}(e) = 1$ for all $e \in F_{n+1}$, it follows that for all $e \in F_{n+1}$ one has

$$f_{n+1}(e) = \max\{f_n(e), g_{n+1}(e)\} = \max\{f_n(e), 1\} = 1 \,.$$

Finally, because $F_n \subseteq F_{n+1}$, we have that $U \setminus F_{n+1} \subseteq U \setminus F_n$, so $0 < f_n(e) < 1$ for all $e \in U \setminus F_n$ in turn implies that $0 < f_n(e) < 1$ for all $e \in U \setminus F_{n+1}$. Therefore

$$0 < f_{n+1}(e) = \max\{f_n(e), g_{n+1}(e)\} < 1$$

for all $e \in U \setminus F_{n+1}$. Therefore we have proven the induction step.

The only claim that remains to be verified is that for all $e \in E$ one has $\lim_{n \to \infty} f_n(e) = \mathbb{I}\{U\}(e)$. This of course splits into two cases: first verifying that, for all $e \in U^c$, that

---

[38] Specifically that E be perfectly normal. However, E is perfectly normal if and only if it is normal and every closed set is G_δ, every closed set is G_δ if and only if every open set is F_σ, so E is perfectly normal if and only if it is normal and all open subsets are F_σ. In particular metrizability is not directly relevant here, even though all metrizable spaces are perfectly normal. That being said, known counterexamples of perfectly normal spaces that are not metrizable, e.g. the Arens-Fort space or the Sorgenfrey line, are all fairly "weird" and thus not likely to be relevant to the practice of data science.

[39] We can assume without loss of generality that the sequence of closed sets is increasing, by taking $\tilde{F}_n := \bigcup_{m \leq n} F_m$ if necessary, noting the fact that the union of finitely many closed sets is again always closed.

$\lim_{n\to\infty} f_n(e) = \mathbb{I}\{U\}(e) = 0$, and second verifying that, for all $e \in U$, that $\lim_{n\to\infty} f_n(e) = \mathbb{I}\{U\}(e) = 1$.

For the first verification, note that by definition of the f_n we have that for all $e \in U^c$ one has $f_n(e) = 0$. Therefore for every $e \in U^c$ the corresponding sequence $\{f_n(e)\}$ is the constant sequence 0, which obviously converges and with 0 as its limit.

For the second verification observe that, because the F_n are assumed to be increasing, i.e. $F_{n+1} \supseteq F_n$ for all n, *and* that $\bigcup_{n=1}^{\infty} F_n = U$ (i.e. we can't have the F_n "stabilize" at some proper subset of U) that for every $e \in U$ there must exist some n_e such that for all $n \geq n_e$, $e \in F_n$. Thus for all $n \geq n_e$, by definition of the f_n we have that $f_n(e) = 1$. Thus for every $e \in U$ the sequence $\{f_n(e)\}$ equals 1 for all but finitely many terms, again obviously forcing the sequence to converge and to the limit 1.

Comments about Approximating Open Indicators from Below   In the case that E is locally compact Polish, as explained above it follows that E is both locally compact and σ-compact. Because the finite union of compact sets is again compact, being σ-compact is equivalent to equaling an *increasing* countable union of compact sets. Because any open subset of a locally compact space is again locally compact, any open subset of E is locally compact. Likewise, in a locally compact Polish space any open set is F_σ and F_σ subsets of all σ-compact spaces are themselves σ-compact, thus any open subset of E is also σ-compact. Therefore we can assume without loss of generality in the case that E is a locally compact Polish space that the increasing sequence of closed sets whose union equals the open subset U consists exclusively of compact sets, and that for each n we have $K_{n+1} \supseteq Int(K_{n+1}) \supseteq K_n$, where $Int(K_{n+1})$ denotes the topological interior (with respect to either E or U, both notions coincide because U is open). Such "(strong) exhaustions by compact sets" can be technically useful.

8.B.10.7   Use for Censoring Subsets

Let U be an open subset of a "generalized edge space" E, i.e. a locally compact Polish space E. Given a monotonically non-decreasing sequence of continuous functions $f_n : E \to [0, 1]$ which converge pointwise to the indicator function of U, it follows that the sequence of functions $1 - f_n$ are a non-increasing sequence of continuous functions which converge pointwise to the indicator function of U^c. If we replace/augment such a convergent sequence to get a convergent net indexed by the directed set $[0, 1]$, then we get non-increasing (as t increases) functions that converge pointwise to the indicator of U^c as $t \to 1$. Re-indexing by $1-t$, we get non-decreasing (as t increases) functions that converge pointwise to the indicator of U^c as $t \to 0$. The details of this will be explained below, but this is to be interpreted as progressively removing the "censoring" of the subset U as t increases, approaching no censoring as $t \to 1$, and full censoring as $t \to 0$. While the important aspect for definitions is that the "easing of censoring" is done monotonically, it is also nice to know that this can be constructed in a way that is continuous.

Specifically, first consider the sequence t_n defined on $[0, 1]$, and converging towards 0, defined for all $n \in \mathbb{N}$ such that[40] $t_n = 2^{-n}$. Define "censoring functions" $\xi_t : E \to [0, 1]$ for all $t \in [0, 1]$ as follows. First, ξ_0 is the constant function identically equal to 1 for all $e \in E$. Then for all $n \geq 1$, define ξ_{t_n} to be $\xi_{t_n} := 1 - f_n$. Then for any other value of $t \in [0, 1]$, first identify the $m \in \mathbb{N}$ such that $t_m = \sup_{n \in \mathbb{N}} \{t_n : t_n < t\}$. Such an m must exist because the t_n are monotonically decreasing and thus there can only be finitely many of them greater than t. Then basically define ξ_t as the corresponding "interpolation" between $\xi_{t_{m-1}}$ and ξ_{t_m}, i.e.

$$
\xi_t = \frac{(t_m - t)}{(t_m - t_{m-1})} \xi_{t_{m-1}} + \frac{(t_m - t_{m-1}) - (t_m - t)}{(t_m - t_{m-1})} \xi_{t_m} \ .
$$

From these definitions it follows that:

- the ξ_t are all continuous for $t > 0$ (being the convex combination of continuous functions),

- as $t \to 1$ the ξ_t approach the constant function identically equal to 1,

- as $t \to 0$ the ξ_t approach the (discontinuous) indicator function of U^c (which equals 0 on U and 1 everywhere else),

- the ξ_t are monotonically non-decreasing as t increases, i.e. for all $t_1 < t_2$ we have for all $e \in E$ that $\xi_{t_1}(e) \leq \xi_{t_2}(e)$.

Thus the ξ_t define a continuous "homotopy" from complete censoring of U as $t \to 0$ to no "censoring" at all as $t \to 1$. For $t > 0$, the $\xi_t \in C(E, \mathbb{R})$ (i.e. are continuous), so their graph-loci are closed subsets of $E \times \mathbb{R}$. Thus I expect, i.e. conjecture and have not rigorously verified, that we would have a homotopy, in the rigorous sense of a continuous function $[0, 1] \to (E \times \mathbb{R})$ connecting the graph-loci of the ξ_t for $t > 0$ with (as $t \to 0$) the closed subset of $E \times \mathbb{R}$ that equals $(U^c \times \{1\}) \cup (\partial(U) \times [0, 1]) \cup (U \times \{0\})$ (which contains, but is a strict superset of, the graph-locus of the discontinuous indicator function of U). Here $\partial(U)$ equals the topological boundary of U, i.e. the closure of U minus the interior of U, or equivalently U^c (which is closed and thus equals its own closure) minus its interior (because the topological boundary of any subset equals that of its complement).

8.B.10.8 Monotone Uncensoring of $U \subseteq E$

Definition 8.B.10.1. *A family of functions* $(\xi_t : E \to \mathbb{R})_{t \in [0,1]}$ *such that*

- *for all $t > 0$ we have $\xi_t \in C(E, \mathbb{R})$ (i.e. ξ_t is continuous),*

- *the graph-loci of the ξ_t "vary continuously with t,*

---

[40] This is probably the simplest example, but any monotonically decreasing sequence bounded within $[0, 1]$ with $t_0 = 1$ and whose limit is 0 will work.

- *as $t \to 1$ the ξ_t approach the constant function $1 : E \to \mathbb{R}$ pointwise,*

- *as $t \to 0$ the ξ_t approach the indicator function of U^c pointwise,*

- *the ξ_t are pointwise monotonically non-decreasing as t increases.*

I claim to have demonstrated previously that such a monotone uncensoring exists for any open subset $U \subseteq E$, and that satisfying the first condition (that the ξ_t be continuous for $t > 0$) requires that $U \subseteq E$ be open.

8.B.10.9 Comments about Monotone Uncensorings

Future work should probably seek to determine whether there is a sensible way to require the ξ_t for $t > 0$ to only be Borel-measurable, rather than necessarily continuous. Again, we showed above that "continuous" monotone uncensorings should only exist for open subsets of E, so one question such future work would have to address is which kinds of subsets of E would admit "measurable" monotone uncensorings.

The definition of (continuous) monotone uncensoring is used to define the (continuous versions of the) monotonicity property and double penalization property for comparison functions of (generalized) networks. Thus it is important to verify that such families of functions actually do exist.

When E has the discrete topology (and thus every subset is open), the condition that U be open is obviously not at all restrictive.

Bibliography

[AEE92]   Jorge Aragón, David Eberly, and Shelly Eberly, *Existence and uniqueness of the maximum likelihood estimator for the two-parameter negative binomial distribution*, Statistics & Probability Letters **15** (1992), no. 5, 375–379.

[AGH⁺17]   Rachael E. Antwis, Sarah M. Griffiths, Xavier A. Harrison, Paz Aranega-Bou, Andres Arce, Aimee S. Bettridge, Francesca L. Brailsford, Alexandre de Menezes, Andrew Devaynes, Kristian M. Forbes, Ellen L. Fry, Ian Goodhead, Erin Haskell, Chloe Heys, Chloe James, Sarah R. Johnston, Gillian R. Lewis, Zenobia Lewis, Michael C. Macey, Alan McCarthy, James E. McDonald, Nasmille L. Mejia-Florez, David O'Brien, Chloé Orland, Marco Pautasso, William D. K. Reid, Heather A. Robinson, Kenneth Wilson, and William J. Sutherland, *Fifty important research questions in microbial ecology*, FEMS Microbiology Ecology **93** (2017), no. 5, 1–10.

[AML⁺17]   Marco Tulio Angulo, Jaime A Moreno, Gabor Lippner, Albert-László Barabási, and Yang-Yu Liu, *Fundamental limitations of network reconstruction from temporal data*, J. R. Soc. Interface **14** (2017), no. 127, 20160966 (en).

[AML19]   Marco Tulio Angulo, Claude H. Moog, and Yang-Yu Liu, *A theoretical framework for controlling complex microbial communities*, Nature Communications **10** (2019), no. 1, 1–12.

[AN72]   Krishna B. Athreya and Peter E. Ney, *Multi-type branching processes*, Branching Processes, Springer Berlin Heidelberg, Berlin, Heidelberg, 1972, pp. 181–228.

[Ans50]   F. J. Anscombe, *Sampling Theory of the Negative Binomial and Logarithmic Series Distributions*, Biometrika **37** (1950), no. 3/4, 358–382.

[APRL⁺20]   Ricardo Andrade-Pacheco, Francois Rerolle, Jean Lemoine, Leda Hernandez, Aboulaye Meté, Lazarus Juziwelo, Aurélien F. Bibaut, Mark J. van der Laan, Benjamin F. Arnold, and Hugh J. W. Sturrock, *Finding hotspots: development of an adaptive spatial sampling approach*, Scientific Reports **10** (2020), no. 1, 1–12.

[Arn10]   Ludwig Arnold, *Random dynamical systems*, Springer monographs in mathematics, Springer, Berlin, Germany, December 2010 (en).

[BA02]   Kenneth P. Burnham and David R. Anderson, *Advanced issues and deeper insights*, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, Springer New York, New York, NY, 2002, pp. 267–351.

[BBI20]   Jacob T Barlow, Said R Bogatyrev, and Rustem F Ismagilov, *A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities*, Nat. Commun. **11** (2020), no. 1, 2590 (en).

[BBM99]   Andrew Barron, Lucien Birgé, and Pascal Massart, *Risk bounds for model selection via penalization*, Probability Theory and Related Fields **113** (1999), no. 3, 301–413.

[BEG+18]   Liana T Burghardt, Brendan Epstein, Joseph Guhlin, Matt S Nelson, Margaret R Taylor, Nevin D Young, Michael J Sadowsky, and Peter Tiffin, *Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments*, Proc. Natl. Acad. Sci. U. S. A. **115** (2018), no. 10, 2425–2430 (en).

[BGF01]   Dean Billheimer, Peter Guttorp, and William F Fagan, *Statistical interpretation of species composition*, Journal of the American Statistical Association **96** (2001), no. 456, 1205–1214.

[BKK16]   Eyal Bairey, Eric D Kelsic, and Roy Kishony, *High-order species interactions shape ecosystem diversity*, Nat. Commun. **7** (2016), no. 1, 12285 (en).

[BMK+18]   Alex Bishara, Eli L Moss, Mikhail Kolmogorov, Alma E Parada, Ziming Weng, Arend Sidow, Anne E Dekas, Serafim Batzoglou, and Ami S Bhatt, *High-quality genome sequences of uncultured microbes by assembly of read clouds*, Nat. Biotechnol. **36** (2018), no. 11, 1067–1075 (en).

[BML+15]   Surojit Biswas, Meredith McDonald, Derek S Lundberg, Jeffery L Dangl, and Vladimir Jojic, *Learning microbial interaction networks from metagenomic count data*, Lecture Notes in Computer Science, Lecture notes in computer science, Springer International Publishing, Cham, 2015, pp. 32–43.

[Bou20]   Djallel Bouneffouf, *Computing the dirichlet-multinomial log-likelihood function*, 2020.

[Bow84]   Kimiko O. Bowman, *Extended Moment Series and the Parameters of the Negative Binomial Distribution*, Biometrics **40** (1984), no. 1, 249–252.

[Bre73]   R. P. Brent, *Algorithms for Minimization Without Derivatives*, Prentice-Hall, Englewood Cliffs, N.J, 1973.

[BSJAB19]   A Barizien, M S Suryateja Jammalamadaka, G Amselem, and Charles N Baroud, *Growing from a few cells: combined effects of initial stochasticity and cell-to-cell variability*, J. R. Soc. Interface **16** (2019), no. 153, 20180935 (en).

[BSM⁺21]   Alexander P Browning, Jesse A Sharp, Tarunendu Mapder, Christopher M Baker, Kevin Burrage, and Matthew J Simpson, *Persistence as an optimal hedging strategy*, Biophys. J. **120** (2021), no. 1, 133–142 (en).

[CdC⁺17]   Daniel T. Chiu, Andrew J. deMello, Dino Di Carlo, Patrick S. Doyle, Carl Hansen, Richard M. Maceiczyk, and Robert C.R. Wootton, *Small but perfectly formed? successes, challenges, and opportunities for microfluidics in the chemical and biological sciences*, Chem **2** (2017), no. 2, 201–223.

[CG19]   Neil Convery and Nikolaj Gadegaard, *30 years of microfluidics*, Micro and Nano Engineering **2** (2019), 76–91.

[CHL⁺20]   Ashley R Coenen, Sarah K Hu, Elaine Luo, Daniel Muratore, and Joshua S Weitz, *A primer for microbiome time-series analysis*, Front. Genet. **11** (2020), 310 (en).

[CK21]   Genís Calderer and Marieke L Kuijjer, *Community detection in large-scale bipartite biological networks*, Front. Genet. **12** (2021), 649440 (en).

[CM12]   Dong Chen and Hans-Georg Müller, *Nonlinear manifold representations for functional data*, The Annals of Statistics **40** (2012), no. 1, 1–29.

[CM13]   Jeng-Min Chiou and Hans-Georg Müller, *Linear manifold modelling of multivariate functional data*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **76** (2013), no. 3, 605–626.

[CMR21]   Julien Chiquet, Mahendra Mariadassou, and Stéphane Robin, *The poisson-lognormal model as a versatile framework for the joint analysis of species abundances*, Frontiers in Ecology and Evolution **9** (2021), 1–17.

[CNd⁺15]   David J. Collins, Adrian Neild, Andrew deMello, Ai-Qun Liu, and Ye Ai, *The poisson distribution and beyond: methods for microfluidic droplet production and single cell encapsulation*, Lab on a Chip **15** (2015), no. 17, 3439–3459.

[Cor09]   Heather J Cordell, *Detecting gene-gene interactions that underlie human diseases*, Nat. Rev. Genet. **10** (2009), no. 6, 392–404 (en).

[Cow84]   Wayne R. Cowell (ed.), *Sources and development of mathematical software*, Prentice-Hall Series in Computational Mathematics, Cleve Moler, Advisor, Prentice-Hall, Upper Saddle River, NJ 07458, USA, 1984.

[CP89]   Suzanne J. Clark and Joe N. Perry, *Estimation of the Negative Binomial Parameter κ by Maximum Quasi -Likelihood*, Biometrics **45** (1989), no. 1, 309–316.

[Cra46]   Harald Cramer, *Mathematical methods of statistics (PMS-9), volume 9*, Princeton Landmarks in Mathematics and Physics, Princeton University Press, Princeton, NJ, March 1946 (en).

[CY18]   Kieran R Campbell and Christopher Yau, *Uncovering pseudotemporal trajectories with covariates from single cell and bulk expression data*, Nat. Commun. **9** (2018), no. 1, 1–12 (en).

[CZPA09] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, and Shun-Ichi Amari, *Nonnegative Matrix and Tensor Factorizations*, John Wiley & Sons, Ltd, September 2009.

[DA21]   Cyrille L Delley and Adam R Abate, *Modular barcode beads for microfluidic single cell genomics*, Sci. Rep. **11** (2021), no. 1, 10857 (en).

[DH15]   Robert M. Dorazio and Margaret E. Hunter, *Statistical models for the analysis and design of digital polymerase chain reaction (dpcr) experiments*, Analytical Chemistry **87** (2015), no. 21, 10886–10893.

[DiSd17] Oliver J. Dressler, Xavier Casadevall i Solvas, and Andrew J. deMello, *Chemical and biological dynamics using droplet-based microfluidics*, Annual Review of Analytical Chemistry **10** (2017), no. 1, 1–24.

[DM21]   Paromita Dubey and Hans-Georg Müller, *Modeling time-varying random objects and dynamic networks*, Journal of the American Statistical Association (2021), 1–16.

[ECR16]   Stephen P. Ellner, Dylan Z. Childs, and Mark Rees, *Data-driven modelling of structured populations*, Springer International Publishing, 2016.

[Efr82]   Bradley Efron, *Maximum Likelihood and Decision Theory*, The Annals of Statistics **10** (1982), no. 2, 340 – 356.

[ELW+20] Amir Erez, Jaime G Lopez, Benjamin G Weiner, Yigal Meir, and Ned S Wingreen, *Nutrient levels and trade-offs control diversity in a serial dilution ecosystem*, eLife **9** (2020), e57790.

[ESGVS21] Sylvie Estrela, Alicia Sanchez-Gorostiaga, Jean Cc Vila, and Alvaro Sanchez, *Nutrient dominance governs the assembly of microbial communities in mixed nutrient environments*, Elife **10** (2021), e65948 (en).

[EVL$^+$22] Sylvie Estrela, Jean C C Vila, Nanxi Lu, Djordje Bajić, Maria Rebolleda-Gómez, Chang-Yu Chang, Joshua E Goldford, Alicia Sanchez-Gorostiaga, and Álvaro Sánchez, *Functional attractors in microbial community assembly*, Cell Syst. **13** (2022), no. 1, 29–42.e7 (en).

[FFD08] Paul G. Falkowski, Tom Fenchel, and Edward F. Delong, *The microbial engines that drive earth's biogeochemical cycles*, Science **320** (2008), no. 5879, 1034–1039.

[FHT10] Jerome H. Friedman, Trevor Hastie, and Rob Tibshirani, *Regularization paths for generalized linear models via coordinate descent*, Journal of Statistical Software **33** (2010), no. 1, 122.

[Fie11] Stephen E Fienberg, *The analysis of contingency tables: From chi-squared tests and log-linear models to models of mixed membership*, Stat. Biopharm. Res. **3** (2011), no. 2, 173–184 (en).

[FM14] Charles K. Fisher and Pankaj Mehta, *Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression*, PLoS ONE **9** (2014), no. 7, e102451.

[FR12] Karoline Faust and Jeroen Raes, *Microbial interactions: from networks to models*, Nature Reviews Microbiology **10** (2012), no. 8, 538–550.

[GCLF18] Didier Gonze, Katharine Z Coyte, Leo Lahti, and Karoline Faust, *Microbial communities as dynamical systems*, Current Opinion in Microbiology **44** (2018), 41–49.

[GCS$^+$21] Xinzhou Ge, Yiling Elaine Chen, Dongyuan Song, Meilu McDermott, Kyla Woyshner, Antigoni Manousopoulou, Ning Wang, Wei Li, Leo D Wang, and Jingyi Jessica Li, *Clipper: p-value-free FDR control on high-throughput data from two conditions*, Genome Biol. **22** (2021), no. 1, 288 (en).

[GFP14] Lisa Goers, Paul Freemont, and Karen M. Polizzi, *Co-culture systems and technologies: taking synthetic biology to the next level*, Journal of The Royal Society Interface **11** (2014), no. 96, 20140065.

[GHG$^+$18] Xuefeng Gao, Bich-Tram Huynh, Didier Guillemot, Philippe Glaser, and Lulla Opatowski, *Inference of significant microbial interactions from longitudinal metagenomics data*, Front. Microbiol. **9** (2018), 2319 (en).

[GL12] Susan Gruber and Mark van der Laan, *tmle: An r package for targeted maximum likelihood estimation*, Journal of Statistical Software **51** (2012), no. 13, 135.

[GLB+18] Joshua E Goldford, Nanxi Lu, Djordje Bajić, Sylvie Estrela, Mikhail Tikhonov, Alicia Sanchez-Gorostiaga, Daniel Segrè, Pankaj Mehta, and Alvaro Sanchez, *Emergent simplicity in microbial community assembly*, Science **361** (2018), no. 6401, 469–474 (en).

[Gor94] Louis Gordon, *A Stochastic Approach to the Gamma Function*, The American Mathematical Monthly **101** (1994), no. 9, 858.

[GR18] Jacqueline Goordial and Jennifer Ronholm, *Metagenomics meets read clouds*, Nat. Biotechnol. **36** (2018), no. 11, 1049–1051 (en).

[Gre81] Ulf Grenander, *Abstract inference*, Probability & Mathematical Statistics S., John Wiley & Sons, Nashville, TN, March 1981 (en).

[GRHW12] Mira T. Guo, Assaf Rotem, John A. Heyman, and David A. Weitz, *Droplet microfluidics for high-throughput biological assays*, Lab on a Chip **12** (2012), no. 12, 2146–2155.

[GW11] Andrew Golightly and Darren J. Wilkinson, *Bayesian parameter inference for stochastic biochemical network models using particle markov chain monte carlo*, Interface Focus **1** (2011), no. 6, 807–820.

[GWDM21] Akshit Goyal, Tong Wang, Veronika Dubinkina, and Sergei Maslov, *Ecology-guided prediction of cross-feeding interactions in the human gut microbiome*, Nat. Commun. **12** (2021), no. 1, 1335 (en).

[GWV+11] David E Graham, Matthew D Wallenstein, Tatiana A Vishnivetskaya, Mark P Waldrop, Tommy J Phelps, Susan M Pfiffner, Tullis C Onstott, Lyle G Whyte, Elizaveta M Rivkina, David A Gilichinsky, Dwayne A Elias, Rachel Mackelprang, Nathan C VerBerkmoes, Robert L Hettich, Dirk Wagner, Stan D Wullschleger, and Janet K Jansson, *Microbes in thawing permafrost: the unknown variable in the climate change equation*, The ISME Journal **6** (2011), no. 4, 709–712.

[Har53] Frank Harary, *On the notion of balance of a signed graph.*, Michigan Mathematical Journal **2** (1953), no. 2, 143 – 146.

[HCBNM02] C Huber-Carol, N Balakrishnan, M Nikulin, and M Mesbah (eds.), *Goodness-of-fit tests and model validity*, 2002 ed., Statistics for Industry and Technology, Birkhäuser, Cambridge, MA, March 2002 (en).

[HCSB20] Joshua G Harrison, W John Calder, Vivaswat Shastry, and C Alex Buerkle, *Dirichlet-multinomial modelling outperforms alternatives for analysis of microbiome and other ecological count data*, Mol. Ecol. Resour. **20** (2020), no. 2, 481–497 (en).

[HCT+19]   Ryan H. Hsu, Ryan L. Clark, Jin Wen Tan, John C. Ahn, Sonali Gupta, Philip A. Romero, and Ophelia S. Venturelli, *Microbial interaction network inference in microfluidic droplets*, Cell Systems **9** (2019), no. 3, 229–242.e4.

[HdlTCD91]   J J Holland, J C de la Torre, D K Clarke, and E Duarte, *Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses*, J. Virol. **65** (1991), no. 6, 2960–2967 (en).

[HE15]   Tailen Hsing and Randall Eubank, *Theoretical foundations of functional data analysis, with an introduction to linear operators*, John Wiley & Sons, Ltd, May 2015.

[HFH21]   Shion Hosoda, Tsukasa Fukunaga, and Michiaki Hamada, *Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model*, Bioinformatics **37** (2021), no. Supplement_1, i16–i24.

[HFJS10]   Robin J. Hogan, Christopher A. T. Ferro, Ian T. Jolliffe, and David B. Stephenson, *Equitability Revisited: Why the 'Equitable Threat Score' Is Not Equitable*, Weather and Forecasting **25** (2010), no. 2, 710–726.

[HFZ18]   Ning Hao, Yang Feng, and Hao Helen Zhang, *Model selection for high-dimensional quadratic regression via regularization*, Journal of the American Statistical Association **113** (2018), no. 522, 615–625.

[HK12]   Lajos Horváth and Piotr Kokoszka, *Inference for functional data with applications*, Springer New York, 2012.

[HMvdW+20]   Charles R. Harris, K. Jarrod Millman, Stfan J van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernndez del Ro, Mark Wiebe, Pearu Peterson, Pierre Grard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant, *Array programming with NumPy*, Nature **585** (2020), 357362.

[HRM+19]   Spencer Havis, Jonathan Rangel, Sujina Mali, Abiodun Bodunrin, Zina Housammy, Rene Zimmerer, Jesse Murphy, William R Widger, and Steven J Bark, *A color-based competition assay for studying bacterial stress responses in micrococcus luteus*, FEMS Microbiol. Lett. **366** (2019), no. 5, fnz054 (en).

[HS02]   A. R. Humphries and A. M. Stuart, *Deterministic and random dynamical systems: theory and numerics*, Modern Methods in Scientific Computing and Applications (Anne Bourlioux, Martin J. Gander, and Gert Sabidussi, eds.), Springer Netherlands, Dordrecht, 2002, pp. 211–254.

[HTF09]   Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The elements of statistical learning*, Springer New York, 2009.

[Hun07]   J. D. Hunter, *Matplotlib: A 2d graphics environment*, Computing in Science & Engineering **9** (2007), no. 3, 90–95.

[HZ17]   Ning Hao and Hao Helen Zhang, *A note on high-dimensional linear regression with interactions*, The American Statistician **71** (2017), no. 4, 291–297.

[IBK06]   E. L. Ionides, C. Bretó, and A. A. King, *Inference for nonlinear dynamical systems*, Proceedings of the National Academy of Sciences **103** (2006), no. 49, 18438–18443.

[IYAR17]   David I. Inouye, Eunho Yang, Genevera I. Allen, and Pradeep Ravikumar, *A review of multivariate distributions for count data derived from the poisson distribution*, WIREs Computational Statistics **9** (2017), no. 3, 1–25.

[Joh04]   Valen E Johnson, *A bayesian $\chi 2$ test for goodness-of-fit*, Ann. Stat. **32** (2004), no. 6, 2361–2384 (en).

[Kee10]   Robert W. Keener, *Theoretical statistics*, Springer New York, 2010.

[KKK+11]   Danai Koutra, Tai-You Ke, U. Kang, Duen Horng (Polo) Chau, Hsing-Kuo Kenneth Pao, and Christos Faloutsos, *Unifying guilt-by-association approaches: Theorems and fast algorithms*, Machine Learning and Knowledge Discovery in Databases (Berlin, Heidelberg) (Dimitrios Gunopulos, Thomas Hofmann, Donato Malerba, and Michalis Vazirgiannis, eds.), Springer Berlin Heidelberg, 2011, pp. 245–260.

[KKO+19]   Jared Kehe, Anthony Kulesa, Anthony Ortiz, Cheri M. Ackerman, Sri Gowtham Thakku, Daniel Sellers, Seppe Kuehn, Jeff Gore, Jonathan Friedman, and Paul C. Blainey, *Massively parallel screening of synthetic microbial communities*, Proceedings of the National Academy of Sciences **116** (2019), no. 26, 12804–12809.

[KL08]   Charles Kooperberg and Michael Leblanc, *Increasing the power of identifying gene x gene interactions in genome-wide association studies*, Genet. Epidemiol. **32** (2008), no. 3, 255–263 (en).

[KMP+07]   Yann Karlen, Alan McNair, Sébastien Perseguers, Christian Mazza, and Nicolas Mermod, *Statistical significance of quantitative pcr*, BMC Bioinformatics **8** (2007), no. 1, 131.

[KOK+20]   Jared Kehe, Anthony Ortiz, Anthony Kulesa, Jeff Gore, Paul C. Blainey, and Jonathan Friedman, *Positive interactions are common among culturable bacteria*, bioRxiv (2020), 169474.

[KP10]   Peter E Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Stochastic Modelling and Applied Probability, Springer, Berlin, Germany, December 2010 (en).

[KR21]   Piotr Kokoszka and Matthew Reimherr, *Introduction to functional data analysis*, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis, London, England, June 2021.

[Kri20]   William E. Krinsman, *Tensor algebra and its applications to data science and statistics*, Master's thesis, University of California, Berkeley, 2020.

[KSV$^+$16] Danai Koutra, Neil Shah, Joshua T. Vogelstein, Brian Gallagher, and Christos Faloutsos, *DeltaCon*, ACM Transactions on Knowledge Discovery from Data **10** (2016), no. 3, 1–43.

[KTS$^+$15] Arto Klami, Abhishek Tripathi, Johannes Sirola, Lauri Vre, and Frederic Roulland, *Latent feature regression for multivariate count data*, Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics (San Diego, California, USA) (Guy Lebanon and S. V. N. Vishwanathan, eds.), Proceedings of Machine Learning Research, vol. 38, PMLR, 09–12 May 2015, pp. 462–470.

[KvVDH10]   Balint Kintses, Liisa D van Vliet, Sean RA Devenish, and Florian Hollfelder, *Microfluidic droplets: new integrated workflows for biological experiments*, Current Opinion in Chemical Biology **14** (2010), no. 5, 548–555.

[LBG10]   Fraser Lewis, Adam Butler, and Lucy Gilbert, *A unified approach to model selection using the likelihood ratio test*, Methods in Ecology and Evolution **2** (2010), no. 2, 155–162.

[LC98]   Erich L Lehmann and George Casella, *Theory of point estimation*, 2 ed., Springer Texts in Statistics, Springer, New York, NY, September 1998 (en).

[LE13]   Todd P Lagus and Jon F Edd, *A review of the theory, methods and recent applications of high-throughput single-cell droplet microfluidics*, Journal of Physics D: Applied Physics **46** (2013), no. 11, 114005.

[LHYA16]   Freeman Lan, John R Haliburton, Aaron Yuan, and Adam R Abate, *Droplet barcoding for massively parallel single-molecule deep sequencing*, Nat. Commun. **7** (2016), 11784 (en).

[Lid79]   William Z. Lidicker, *A clarification of interactions in ecological systems*, BioScience **29** (1979), no. 8, 475–477.

[LLLW11]   Tao Lu, Hua Liang, Hongzhe Li, and Hulin Wu, *High-dimensional ODEs coupled with mixed-effects modeling techniques for dynamic gene regulatory network identification*, Journal of the American Statistical Association **106** (2011), no. 496, 1242–1258.

[LPR19]   Jose V Lopez, Raquel S Peixoto, and Alexandre S Rosado, *Inevitable future: space colonization beyond earth with microbes first*, FEMS Microbiology Ecology **95** (2019), no. 10, 1–9.

[LR20]   Keli Liu and Feng Ruan, *A self-penalizing objective function for scalable interaction detection*, 2020.

[LS07]   James O. Lloyd-Smith, *Maximum Likelihood Estimation of the Negative Binomial Dispersion Parameter for Highly Overdispersed Data, with Applications to Infectious Diseases*, PLoS ONE **2** (2007), no. 2, e180.

[LS09]   F. William Lawvere and Stephen H. Schanuel, *Conceptual Mathematics: A First Introduction to Categories*, Cambridge University Press, Cambridge, UK New York, 2009.

[LS19]   Leo Lahti and Sudarshan Shetty, *microbiome r package*, 2012-2019.

[LYS+19]   Hsiao-Han Lin, Manda Yu, Manoj Kumar Sriramoju, Shang-Te Danny Hsu, Chi-Te Liu, and Erh-Min Lai, *A high-throughput interbacterial competition screen identifies ClpAP in enhancing recipient susceptibility to type VI secretion system-mediated attack by agrobacterium tumefaciens*, Front. Microbiol. **10** (2019), 3077 (en).

[MBH+14]   Simeone Marino, Nielson T Baxter, Gary B Huffnagle, Joseph F Petrosino, and Patrick D Schloss, *Mathematical modeling of primary succession of murine intestinal microbiota*, Proc. Natl. Acad. Sci. U. S. A. **111** (2014), no. 1, 439–444 (en).

[MGH80]   Jorge J. Moré, Burton S. Garbow, and Kenneth E. Hillstrom, *User guide for MINPACK-1*, Tech. Report ANL-80-74, Argonne National Laboratory, Argonne, IL, USA, August 1980.

[MGSD19]   Andrew McDavid, Raphael Gottardo, Noah Simon, and Mathias Drton, *Graphical models for zero-inflated single cell gene expression*, Ann. Appl. Stat. **13** (2019), no. 2, 848–873 (en).

[MH13]   Paul J. McMurdie and Susan Holmes, *phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data*, PLOS ONE **8** (2013), no. 4, 1–11.

[MJS20]   R S MacKay, S Johnson, and B Sansom, *How directed is a directed network?*, R. Soc. Open Sci. **7** (2020), no. 9, 201138 (en).

[MKBDB00]   A F Marée, W Keulen, C A Boucher, and R J De Boer, *Estimating relative fitness in viral competition experiments*, J. Virol. **74** (2000), no. 23, 11067–11072 (en).

[MMV+08]   Jerome Mounier, Christophe Monnet, Tatiana Vallaeys, Roger Arditi, Anne-Sophie Sarthou, Arnaud Helias, and Francoise Irlinger, *Microbial Interactions within a Cheese Microbial Community*, Applied and Environmental Microbiology **74** (2008), no. 1, 172–181.

[MRH20]   Kinga Matula, Francesca Rivello, and Wilhelm T. S. Huck, *Single-cell analysis using droplet microfluidics*, Advanced Biosystems **4** (2020), no. 1, 1900188.

[MSH+08]   Ramamurthy Mani, Robert P. St.Onge, John L. Hartman, Guri Giaever, and Frederick P. Roth, *Defining genetic interaction*, Proceedings of the National Academy of Sciences **105** (2008), no. 9, 3461–3466.

[MSH+19]   Clarisse Marotz, Anukriti Sharma, Greg Humphrey, Neil Gottel, Christopher Daum, Jack A Gilbert, Emiley Eloe-Fadrosh, and Rob Knight, *Triplicate PCR reactions for 16S rRNA gene amplicon sequencing are unnecessary*, Biotechniques **67** (2019), no. 1, 29–32 (en).

[MTFH16]   Ahmad Mahmood, Oğuz Can Turgay, Muhammad Farooq, and Rifat Hayat, *Seed biopriming with plant growth promoting rhizobacteria: a review*, FEMS Microbiology Ecology **92** (2016), no. 8, 1–14.

[MWC19]   Michael R McLaren, Amy D Willis, and Benjamin J Callahan, *Consistent and correctable bias in metagenomic sequencing experiments*, eLife **8** (2019), e46923.

[MXS17]   Babak Momeni, Li Xie, and Wenying Shou, *Lotka-volterra pairwise modeling fails to capture diverse pairwise microbial interactions*, eLife **6** (2017), e25051.

[Nak97]   Eiji Nakashima, *Some Methods for Estimation in a Negative-Binomial Model*, Annals of the Institute of Statistical Mathematics **49** (1997), no. 1, 101–115.

[NW18]   Deborah A Neher and Thomas R Weicht, *A plate competition assay as a quick preliminary assessment of disease suppression*, J. Vis. Exp. (2018), no. 140, e58767 (en).

[OTN+17]   Otso Ovaskainen, Gleb Tikhonov, Anna Norberg, F. Guillaume Blanchet, Leo Duan, David Dunson, Tomas Roslin, and Nerea Abrego, *How to make more out of community data? a conceptual framework and its implementation as models and software*, Ecology Letters **20** (2017), no. 5, 561–576.

[PE00]   Mercedes Pascual and Stephen P. Ellner, *Linking ecological patterns to environmental forcing via nonlinear time series models*, Ecology **81** (2000), no. 10, 2767–2780.

[PGB21]   Vincent Prost, Stéphane Gazut, and Thomas Brüls, *A zero inflated log-normal model for inference of sparse microbial association networks*, PLOS Computational Biology **17** (2021), no. 6, e1009089.

[Pie90]   Walter W. Piegorsch, *Maximum Likelihood Estimation for the Negative Binomial Dispersion Parameter*, Biometrics **46** (1990), no. 3, 863–867.

[PM17]   Nathan T. Porter and Eric C. Martens, *The Critical Roles of Polysaccharides in Gut Microbial Ecology and Physiology*, Annual Review of Microbiology **71** (2017), no. 1, 349–369.

[Pow70]   M. J. D. Powell, *A Hybrid Method for Nonlinear Equations*, Numerical Methods for Nonlinear Algebraic Equations (P. Rabinowitz, ed.), Gordon and Breach, 1970.

[PPG+10]   Maya L Petersen, Kristin E Porter, Susan Gruber, Yue Wang, and Mark J van der Laan, *Diagnosing and responding to violations in the positivity assumption*, Statistical Methods in Medical Research **21** (2010), no. 1, 31–54.

[PSG+21]   Prapti Pattanayak, Sachin Kumar Singh, Monica Gulati, Sukriti Vishwas, Bhupinder Kapoor, Dinesh Kumar Chellappan, Krishnan Anand, Gaurav Gupta, Niraj Kumar Jha, Piyush Kumar Gupta, Parteek Prasher, Kamal Dua, Harish Dureja, Deepak Kumar, and Vijay Kumar, *Microfluidic chips: recent advances, critical strategies in design, applications and future perspectives*, Microfluid. Nanofluidics **25** (2021), no. 12, 99 (en).

[QMT+17]   Xiaojie Qiu, Qi Mao, Ying Tang, Li Wang, Raghav Chawla, Hannah A Pliner, and Cole Trapnell, *Reversed graph embedding resolves complex single-cell trajectories*, Nat. Methods **14** (2017), no. 10, 979–982 (en).

[QSyL+08]   Yan Qi, Yasir Suhail, Yu yi Lin, Jef D. Boeke, and Joel S. Bader, *Finding friends and enemies in an enemies-only network: A graph diffusion kernel for predicting novel genetic interactions and co-complex membership from yeast genetic interactions*, Genome Research **18** (2008), no. 12, 1991–2004.

[RDC+06]   Daniel C. Reuman, Robert A. Desharnais, Robert F. Costantino, Omar S. Ahmad, and Joel E. Cohen, *Power spectra reveal the influence of stochasticity on nonlinear population dynamics*, Proceedings of the National Academy of Sciences **103** (2006), no. 49, 18860–18865.

[RGG+21]   Morteza Roodgar, Benjamin H Good, Nandita R Garud, Stephen Martis, Mohan Avula, Wenyu Zhou, Samuel M Lancaster, Hayan Lee, Afshin Babveyh, Sophia Nesamoney, Katherine S Pollard, and Michael P Snyder, *Longitudinal linked-read sequencing reveals ecological and evolutionary responses of a human gut microbiome during antibiotic treatment*, Genome Res. **31** (2021), no. 8, 1433–1446 (en).

[RH17]   James Ramsay and Giles Hooker, *Dynamic data analysis*, Springer New York, 2017.

[RHCC07]   J. O. Ramsay, G. Hooker, D. Campbell, and J. Cao, *Parameter estimation for differential equations: a generalized smoothing approach*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **69** (2007), no. 5, 741–796.

[RHG09]   James Ramsay, Giles Hooker, and Spencer Graves, *Functional data analysis with R and MATLAB*, Springer New York, 2009.

[Rob55]   Herbert Robbins, *A Remark on Stirling's Formula*, The American Mathematical Monthly **62** (1955), no. 1, 26–29.

[RP85]   G. J. S. Ross and D. A. Preece, *The Negative Binomial Distribution*, Journal of the Royal Statistical Society. Series D (The Statistician) **34** (1985), no. 3, 323–335.

[RS05]   J. O. Ramsay and B. W. Silverman, *Functional data analysis*, Springer New York, 2005.

[SBT+13]   Richard R Stein, Vanni Bucci, Nora C Toussaint, Charlie G Buffie, Gunnar Rätsch, Eric G Pamer, Chris Sander, and João B Xavier, *Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota*, PLoS Comput. Biol. **9** (2013), no. 12, e1003388 (en).

[Sco92]   David Scott, *Multivariate Density Estimation : Theory, Practice, and Visualization*, John Wiley & Sons, New York, Chicester, 1992.

[SCZ17]   Luoran Shang, Yao Cheng, and Yuanjin Zhao, *Emerging droplet microfluidics*, Chemical Reviews **117** (2017), no. 12, 7964–8040.

[Sel13]   Peter Selinger, *Lecture notes on the lambda calculus*, 2013.

[Ser80]   Robert J Serfling, *Approximation theorems of mathematical statistics*, Wiley Series in Probability and Statistics, Wiley-Interscience, Newy York, September 1980 (en).

[SFB14]   Eric K. Sackmann, Anna L. Fulton, and David J. Beebe, *The present and future role of microfluidics in biomedical research*, Nature **507** (2014), no. 7491, 181–189.

[Skl14]   Max Sklar, *Fast MLE Computation for the Dirichlet Multinomial*, 2014.

[SkM20]   Somayeh Sohrabi, Nour kassir, and Mostafa Keshavarz Moraveji, *Droplet microfluidics: fundamentals and its advanced applications*, RSC Advances **10** (2020), no. 46, 27560–27574.

[SL21]   Dongyuan Song and Jingyi Jessica Li, *PseudotimeDE: inference of differential gene expression along cell pseudotime with well-calibrated p-values from single-cell RNA sequencing data*, Genome Biol. **22** (2021), no. 1, 124 (en).

[SMY⁺12]   George Sugihara, Robert May, Hao Ye, Chih-Hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch, *Detecting causality in complex ecosystems*, Science **338** (2012), no. 6106, 496–500 (en).

[SP10]   Skipper Seabold and Josef Perktold, *statsmodels: Econometric and statistical modeling with python*, 9th Python in Science Conference, 2010.

[SS19]   Lauren Speare and Alecia N Septer, *Coincubation assay for quantifying competitive interactions between vibrio fischeri isolates*, J. Vis. Exp. (2019), no. 149, e59759 (en).

[SSB⁺17]   Yu Shang, Johannes Sikorski, Michael Bonkowski, Anna-Maria Fiore-Donno, Ellen Kandeler, Sven Marhan, Runa S Boeddinghaus, Emily F Solly, Marion Schrumpf, Ingo Schöning, Tesfaye Wubet, Francois Buscot, and Jörg Overmann, *Inferring interactions in complex microbial communities from nucleotide sequence data and environmental parameters*, PLoS One **12** (2017), no. 3, e0173765 (en).

[Str85]   Helmut Strasser, *Mathematical theory of statistics: Statistical experiments and asymptotic decision theory*, De Gruyter, 1985.

[SW62]   L. R. Shenton and P. A. Wallington, *The Bias of Moment Estimators with an Application to the Negative Binomial Distribution*, Biometrika **49** (1962), no. 1/2, 193–204.

[SWMD17]   Justin D Silverman, Alex D Washburne, Sayan Mukherjee, and Lawrence A David, *A phylogenetic transform enhances analysis of compositional microbiota data*, Elife **6** (2017), e21887 (en).

[TCG⁺14]   Cole Trapnell, Davide Cacchiarelli, Jonna Grimsby, Prapti Pokharel, Shuqiang Li, Michael Morse, Niall J Lennon, Kenneth J Livak, Tarjei S

Mikkelsen, and John L Rinn, *The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells*, Nat. Biotechnol. **32** (2014), no. 4, 381–386 (en).

[TF08]   Wei-Chang Tian and Erin Finehout, *Introduction to Microfluidics*, Microfluidics for Biological Applications, Springer US, 2008, pp. 1–34.

[TITP19]   Mattia Tantardini, Francesca Ieva, Lucia Tajoli, and Carlo Piccardi, *Comparing Methods for Comparing Networks*, Scientific Reports **9** (2019), no. 1, 1–19.

[TKA20]   Giuseppe Torrisi, Reimer Khn, and Alessia Annibale, *Percolation on the gene regulatory network*, Journal of Statistical Mechanics: Theory and Experiment **2020** (2020), no. 8, 083501.

[TLHL08]   Shia-Yen Teh, Robert Lin, Lung-Hsin Hung, and Abraham P. Lee, *Droplet microfluidics*, Lab on a Chip **8** (2008), no. 2, 198–220.

[Tre06]   Nickolay T Trendafilov, *The dynamical system approach to multivariate data analysis*, J. Comput. Graph. Stat. **15** (2006), no. 3, 628–650 (en).

[TRKS$^+$17]   Daniel Taylor-Rodríguez, Kimberly Kaufeld, Erin M. Schliep, James S. Clark, and Alan E. Gelfand, *Joint species distribution modeling: Dimension reduction using dirichlet processes*, Bayesian Analysis **12** (2017), no. 4, 939–967.

[TVL$^+$22]   Daniel Taylor, Nia Verdon, Peter Lomax, Rosalind J Allen, and Simon Titmuss, *Tracking the stochastic growth of bacterial populations in microfluidic droplets*, Physical Biology **19** (2022), no. 2, 026003.

[UF13]   Shahid Ullah and Caroline F Finch, *Applications of functional data analysis: A systematic review*, BMC Med. Res. Methodol. **13** (2013), no. 1, 43 (en).

[Vaa98]   A. W. van der Vaart, *Asymptotic statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1998.

[VCF$^+$18]   Ophelia S Venturelli, Alex V Carr, Garth Fisher, Ryan H Hsu, Rebecca Lau, Benjamin P Bowen, Susan Hromada, Trent Northen, and Adam P Arkin, *Deciphering microbial interactions in synthetic human gut microbiome communities*, Molecular Systems Biology **14** (2018), no. 6, 1–19.

[vdLPH07]   Mark J. van der Laan, Eric C Polley, and Alan E. Hubbard, *Super learner*, Statistical Applications in Genetics and Molecular Biology **6** (2007), no. 1, 1–21.

[vdLR11]   Mark J. van der Laan and Sherri Rose, *Targeted learning*, Springer New York, 2011.

[vdLR18]   ———, *Targeted learning in data science*, Springer International Publishing, 2018.

[VGO+20]   Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nature Methods **17** (2020), 261–272.

[Was21]   Michael L. Waskom, *seaborn: statistical data visualization*, Journal of Open Source Software **6** (2021), no. 60, 3021.

[WBGLR+21]   Deborah Weighill, Marouen Ben Guebila, Camila Lopes-Ramos, Kimberly Glass, John Quackenbush, John Platig, and Rebekka Burkholz, *Gene regulatory network inference as relaxed graph matching*, Proceedings of the AAAI Conference on Artificial Intelligence **35** (2021), no. 11, 10263–10272.

[WCM16]   Jane-Ling Wang, Jeng-Min Chiou, and Hans-Georg Mller, *Functional data analysis*, Annual Review of Statistics and Its Application **3** (2016), no. 1, 257–295.

[WFY84]   Linda J. Willson, J. Leroy Folks, and J. H. Young, *Multistage Estimation Compared with Fixed-Sample-Size Estimation of the Negative Binomial Parameter k*, Biometrics **40** (1984), no. 1, 109–117.

[WFY86]   L. J. Willson, J. L. Folks, and J. H. Young, *Complete sufficiency and maximum likelihood estimation for the two-parameter negative binomial distribution*, Metrika **33** (1986), no. 1, 349–362.

[Whi06]   George M. Whitesides, *The origins and the future of microfluidics*, Nature **442** (2006), no. 7101, 368–373.

[Wic16]   Hadley Wickham, *ggplot2: Elegant graphics for data analysis*, Springer-Verlag New York, 2016.

[Wil18]   Darren J Wilkinson, *Stochastic modelling for systems biology*, 3 ed., Chapman & Hall/CRC Mathematical and Computational Biology, CRC Press, London, England, November 2018.

[WL02]   Michael Wagner and Alexander Loy, *Bacterial community composition and function in sewage treatment systems*, Current Opinion in Biotechnology **13** (2002), no. 3, 218–227.

[WL15]   Michael J Wiser and Richard E Lenski, *A comparison of methods to measure fitness in escherichia coli*, PLoS One **10** (2015), no. 5, e0126210 (en).

[WM20]   Peter Wills and François G. Meyer, *Metrics for graph comparison: A practitioner's guide*, PLOS ONE **15** (2020), no. 2, e0228728.

[Woo10]   Simon N. Wood, *Statistical inference for noisy nonlinear ecological dynamic systems*, Nature **466** (2010), no. 7310, 1102–1104.

[WPW+15]   Kelly M Wetmore, Morgan N Price, Robert J Waters, Jacob S Lamson, Jennifer He, Cindi A Hoover, Matthew J Blow, James Bristow, Gareth Butland, Adam P Arkin, and Adam Deutschbauer, *Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons*, MBio **6** (2015), no. 3, e00306–15 (en).

[WSCR03]   Lance A Waller, David Smith, James E Childs, and Leslie A Real, *Monte carlo assessments of goodness-of-fit for ecological simulation models*, Ecol. Modell. **164** (2003), no. 1, 49–63 (en).

[XAF+17]   Yandong Xiao, Marco Tulio Angulo, Jonathan Friedman, Matthew K. Waldor, Scott T. Weiss, and Yang-Yu Liu, *Mapping the ecological networks of microbial communities*, Nature Communications **8** (2017), no. 1, 1–12.

[YEJ+07]   Takehito Yoshida, Stephen P Ellner, Laura E Jones, Brendan J. M Bohannan, Richard E Lenski, and Nelson G Hairston, *Cryptic population dynamics: Rapid evolution masks trophic interactions*, PLoS Biology **5** (2007), no. 9, e235.

[YHV13]   Danni Yu, Wolfgang Huber, and Olga Vitek, *Shrinkage estimation of dispersion in Negative Binomial models for RNA-seq experiments with small sample size*, Bioinformatics **29** (2013), no. 10, 1275–1282.

[YRCS06]   Joshua S. Yuan, Ann Reed, Feng Chen, and C. Neal Stewart, *Statistical analysis of real-time pcr data*, BMC Bioinformatics **7** (2006), no. 1, 85.

[YS14]   Peng Yu and Chad A. Shaw, *An efficient algorithm for accurate computation of the dirichlet-multinomial log-likelihood function*, Bioinformatics **30** (2014), no. 11, 1547–1554.

[YSP+19]   Sukhwan Yoon, Bongkeun Song, Rebecca L Phillips, Jin Chang, and Min Joon Song, *Ecological and physiological implications of nitrogen oxide reduction*

pathways on greenhouse gas emissions in agroecosystems, FEMS Microbiology Ecology **95** (2019), no. 6, 1–15.

[YWSJ08] Joshua S. Yuan, Donglin Wang, and C. Neal Stewart Jr., *Statistical methods for efficiency adjusted real-time pcr quantification*, Biotechnology Journal **3** (2008), no. 1, 112–123.

[Zho18] Mingyuan Zhou, *Nonparametric Bayesian Negative Binomial Factor Analysis*, Bayesian Analysis **13** (2018), no. 4, 1065–1093.

[ZHSL12] Or Zuk, Eliana Hechter, Shamil R Sunyaev, and Eric S Lander, *The mystery of missing heritability: Genetic interactions create phantom heritability*, Proc. Natl. Acad. Sci. U. S. A. **109** (2012), no. 4, 1193–1198 (en).

[ZW17] Pingan Zhu and Liqiu Wang, *Passive and active droplet generation with microfluidics: a review*, Lab on a Chip **17** (2017), no. 1, 34–75.

[ZYLS18] Chen Zhang, Hao Yan, Seungho Lee, and Jianjun Shi, *Dynamic multivariate functional data modeling via sparse subspace learning*, 2018.

[ZZZS17] Yiwen Zhang, Hua Zhou, Jin Zhou, and Wei Sun, *Regression models for multivariate count data*, Journal of Computational and Graphical Statistics **26** (2017), no. 1, 1–13.

Appendix A.1

Relevant Background on Differential Equations

The goals of the simulations were to follow the precedent of prior work [VCF$^+$18] [FM14] [MBH$^+$14] [SBT$^+$13] [MMV$^+$08] in postulating a likelihood model similar to the generalized Lotka-Volterra equations (cf. again section 2.2.5), while also being more biologically realistic in both (i) not having any restrictions on interaction coefficients (which would seem to be arbitrary from a biological point of view), and (ii) addressing the capacity for nutrient depletion and waste product accumulation to moderate changes in population size, something which seems especially relevant, and even inappropriate to ignore, inside of the highly resource-limited (oligotrophic) environment of a microfluidic droplet.

A.1.1   Exponential Growth: Unlimited Resources

The exponential differential equation describes the idealized situation where the rate of change in the population is directly proportional to the current number of cells:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = rn\,, \tag{A.1.1.1}$$

which is equivalent (for populations with more than zero cells) to the even simpler equation:

$$\frac{\mathrm{d}(\log n)}{\mathrm{d}t} = r\,. \tag{A.1.1.2}$$

The second form emphasizes how this idealized situation corresponds to what is often referred to as a constant growth (or decay) rate.

A.1.2   Nutrient Limitation and Indirect Interactions

The exponential model is often considered adequate for describing the growth under copiotrophic conditions immediately following the initial lag phase. However, since the effects of nutrient depletion and waste product accumulation increase with the size of the population, for large enough populations it becomes unrealistic to model the population size's effect on the rate of change as being constant. Instead an additional multiplicative factor can be included to describe how growth is expected to slow down as nutrients are depleted and waste products accumulate:

$$\frac{\mathrm{d}(\log x)}{\mathrm{d}t} = r \left(1 - \frac{x}{K} \right) . \tag{A.1.2.1}$$

This is called the logistic differential equation. As the population size continues to approach the constant K, growth slows down (asymptotically) towards nothing, causing K to represent an upper bound on the total possible population size. This upper bound is often referred to as the *carrying capacity* of the environment which supports the population.

Assuming that the population sizes of any given strain do not affect the population sizes of the other strains, it is straightforward to model the simultaneous exponential growth of multiple strains via a system of differential equations:

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = r_s . \tag{A.1.2.2}$$

When still assuming that the population sizes of different strains do not affect one another, it is also straightforward in exactly the same manner to model the simultaneous logistic growth of multiple strains using a system of differential equations:

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = r_s \left(1 - \frac{n^{(s)}}{K_s} \right) . \tag{A.1.2.3}$$

The above system of equations (A.1.2.3) is unrealistic, however, inasmuch as it assumes not only that the population sizes of different strains do not directly affect one another, but also because it assumes that the environment has a separate carrying capacity for each strain, independent of that for all other strains.

Especially in an environment like the interior of a single droplet in a microfluidics experiment, the same nutrient supply will often be common/shared among all strains, and the accumulation of waste products will affect all strains simultaneously. In other words, in a limited resource environment strains will almost always interact with each other at least *indirectly*, with the *indirect* interactions being mediated by the nutrient supply and concentration of waste products. Therefore it is often more realistic to assume a single carrying capacity for the total aggregate population size of all strains in the environment. The following system of equations better models the aforementioned situation:

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = r_s \left(1 - \frac{\sum_\sigma n^{(\sigma)}}{K}\right), \tag{A.1.2.4}$$

since as the total aggregate population size of all strains (asymptotically) approaches the common and shared carrying capacity K of the environment, the growth of all strains will simultaneously slow (asymptotically) towards zero. It is worth noting though that in practice different strains can have wildly varying metabolisms, such that what constitutes a nutrient or a harmful waste product could differ substantially between strains, and making the above description overly simplistic. For examples of models accounting for the potentially different metabolisms of different strains, see e.g. [EVL$^+$22], [GWDM21], [ESGVS21], [ELW$^+$20], [GLB$^+$18], or [MXS17]. (For an overview, see [GCLF18].) However, fully addressing this issue falls outside of the scope of the current manuscript and is left for future work.

A.1.2.1 Comparison with Previous Work

The paper [MBH$^+$14] uses equations analogous to (A.1.2.3), with $K_s = K$ for all s for a common K, and claims that this alone is sufficient alone to enforce the constraint that $\sum_\sigma n^{(\sigma)} \leq K$ always. However, that does not seem to be true. As far as I can tell, with an equation of that form alone, there is no reason at all why the (logarithmic) derivatives of all strains should approach 0 as $\sum_\sigma n^{(\sigma)} \to K$. If the growth rates do not asymptotically approach 0 as the total population sizes $\sum_\sigma n^{(\sigma)}$ of all strains approaches the supposedly common carrying capacity K, then there is no reason to expect in general that $\sum_\sigma n^{(\sigma)} \leq K$.

For example, based on (A.1.2.3), as $n^{(s_1)} \to K/2$ and $n^{(s_2)} \to K/2$ the growth rates of s_1 and s_2 would not seem to be approach 0. Yet under those circumstances $\sum_\sigma n^{(\sigma)} \leq K$. Hence, absent a flaw in the above reasoning, it seems clear that (A.1.2.3) is incapable of enforcing a common carrying capacity constraint shared by all strains simultaneously.

In contrast, it is obviously the case for the proposed equation (A.1.2.4) that the growth rates of all strains will asymptotically approach 0 as $\sum_\sigma n^{(\sigma)} \to K$.

A.1.3 Direct Interactions

Even focusing on the initial regime where diminishing nutrient supplies and increasing waste product concentrations have not yet become significant factors affecting growth, the assumption in (A.1.2.2) that the growth of any given strain's population is unaffected by the population sizes of other strains can still be unrealistic. This is of course because of *direct* interactions between different strains of microbes, which can either support or inhibit their growth, and which are the primary focus of interest of the current manuscript. The generalized Lotka-Volterra equations seek to account for direct interactions by adding to (A.1.2.2) a linear interaction term:

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)} \,. \tag{A.1.3.1}$$

The model describes interactions between different strains through the magnitude and sign of the coefficients in the portion of the affine expressions above which are a linear combination of the population sizes of all of the strains. Positive coefficients correspond to mutualistic interactions between strains whereas negative coefficients correspond to antagonistic interactions between strains.

A.1.4   Combining Indirect and Direct Interactions

The simulations used in this paper seek to capture both indirect and direct interactions between microbial strains by combining models (A.1.2.4) and (A.1.3.1) above in a straightforward way:

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = \left(r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)} \right) \left(1 - \frac{\sum_{\sigma} n^{(\sigma)}}{K} \right) \,. \tag{A.1.4.1}$$

This preserves the property that the change in population size will slow towards zero as the total aggregate population size approaches the environment's carrying capacity, while still retaining ability for the population sizes of various strains to directly affect one another. This makes sense even in the case of antagonistic interactions, since as the growth of the antagonizing strain slows towards zero (along with the growth of all growing strains), one would expect that the magnitude of the antagonizing interaction would also slow towards zero, making it sensible for the decay in the population size in the antagonized strain to slow towards zero as well.

Also, unlike the regular generalized Lotka-Volterra equations, it is intended that, for any possible choice of direct interaction coefficients, the population sizes of any solutions will remain bounded for any initial condition for which the carrying capacity of the aggregate population is not exceeded. (Cf. for example [VCF+18] which excludes certain combinations of coefficient values to prevent the possibility of blow-up. To me that seems arbitrary from a purely biological point of view.) This is an additional source of realism for (A.1.4.1), since there does not seem to be any a priori biological reason for why certain combinations of values of interaction coefficients should be prohibited. This quote from [MBH+14] is precedent for the above viewpoint on parameter constraints:

> One approach to improving parameter estimation is to superimpose constraints on the parameters; however, unless these constraints can be justified both biologically and mathematically, they should not be enforced because the results can be greatly affected.

A.1.4.1 Comparison with Previous Work

Note that the derivation of (A.1.4.1) given above is not the first attempt to combine logistic growth constraints with gLV-like direct interactions. Notably [MBH$^+$14] proposes an equation of the form (modulo logarithms and differences in notation):

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = r_s \left(1 - \frac{n^{(\sigma)}}{K} \right) + \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)} \,. \tag{A.1.4.2}$$

In section A.1.2.1 above I already explained why the proposed implementation of a common carrying capacity in (A.1.4.2) does not seem to enforce the intended constraint $\sum_\sigma n^{(\sigma)} \le K$. Let us apply the fix proposed in A.1.2.1, namely switching the attempted carrying capacity implementation from a form similar to (A.1.2.3) to a form similar instead to (A.1.2.4):

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = r_s \left(1 - \frac{\sum_\sigma n^{(\sigma)}}{K} \right) + \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)} \,. \tag{A.1.4.3}$$

However (A.1.4.3) also seems to fail to enforce the intended constraint $\sum_\sigma n^{(\sigma)} \le K$, for the same reason that (A.1.2.3) does. Namely, without distributing the "logistic factor" across the entire expression, which *is* done in (A.1.4.1), again there does not seem to be any reason why we should have for (A.1.4.3) that all growth rates simultaneously approach 0 as $\sum_\sigma n^{(\sigma)} \to K$. In contrast, (A.1.4.1) was explicitly designed so that all growth rates simultaneously approach 0 as $\sum_\sigma n^{(\sigma)} \to K$. Hence it seems much more plausible to me that (A.1.4.1) enforces the intended constraint $\sum_\sigma n^{(\sigma)} \le K$, whereas I do not see why either (A.1.4.2) or (A.1.4.3) would. Thus (A.1.4.1) seems to be an improvement on previous work.

A.1.5 Common Framework

All of the above systems of equations fit into the following framework:

$$\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = G_s(n^{(1)}, \ldots, n^{(S)}) \left(1 - \frac{H_s(n^{(1)}, \ldots, n^{(S)})}{K_s} \right) \,, \tag{A.1.5.1}$$

where G_s corresponds to the growth of the given strain under idealized conditions, and H_s is a quantity for which the external environment imposes some finite carrying capacity K_s, such that $H_s \to K_s$ causes all changes in the population size of strain s to slow towards zero. For equations (A.1.2.2), (A.1.2.3), and (A.1.2.4), one has that $G_s = r_s$, whereas for equations (A.1.3.1) and (A.1.4.1) one has instead that $G_s = r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)}$. For equations (A.1.2.2) and (A.1.3.1) one has that $H_s = 0$, for equation (A.1.2.3) one has that $H_s = n^{(s)}$, and for equations (A.1.2.4) and (A.1.4.1) one has instead that $H_s = \sum_\sigma n^{(\sigma)}$ as well as that $K_s = K$ for all $s \in [S]$ for a given fixed constant K.

Of course in order for its effect on changes in the population size of strain s to have any clear interpretation as an external limitation on the magnitude of changes in the system, one

needs that $0 \leq H_s \leq K_s$ so that the factor multiplying G_s is always a value between 0 and 1. As mentioned already before, $H_s = \sum_\sigma n^{(\sigma)}$ satisfies this property in equations (A.1.2.4) and (A.1.4.1), or at least for the region of phase space between the two hyperplanes $\sum_\sigma n^{(\sigma)} = 0$ and $\sum_\sigma n^{(\sigma)} = K$ (recall that $K_1 = \cdots = K_S = K$ for these equations). This is fine since the aformentioned region is the only biologically realistic/sensible/meaningful region of phase space anyway, and any solution of (A.1.4.1) beginning in that region will remain inside that region for all time.

A.1.6   "Competitive Lotka-Volterra Equations"

There is at least one other similar system of differential equations intended to model both the direct interactions corresponding to the system (A.1.3.1) and external limitations due to finiteness of nutrient supply and the accumulation of waste products, which is often called the "Competitive Lotka-Volterra equations". For the competitive Lotka-Volterra equations, using the general framework of (A.1.5.1) one has

$$G_s = r_s \,, \qquad H_s = \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)} \,. \tag{A.1.6.1}$$

Apart from the issue that K_s may have different values for different strains (which can easily be rectified), a bigger issue with these equations, and why they were not used for these simulations, is that there is no way to guarantee that $0 \leq H_s \leq 1$, at least not without restrictions on the interaction coefficients which from a biological point of view would be arbitrary. Moreover $H_s = \sum_{\sigma=1}^{S} \alpha_{\sigma s} n^{(\sigma)}$ does not correspond to any scientifically meaningful or experimentally observable quantity, making the meaning of the restrictions imposed on the growth of strain s by the $(1 - H_s/K_s)$ factor difficult, if not impossible, to give any scientifically meaningful interpretation. In contrast, $H_s = \sum_\sigma n^{(\sigma)}$ is not only possible to measure experimentally, but is probably easier to measure than any of the $n^{(s)}$ individually (since the accurate assignment of microbes' taxonomic identifications can be challenging to do in a high-throughput manner).

   This last reason for not using the competitive Lotka-Volterra equations is more subjective, but it seems that the purpose of the linear term in the generalized Lotka-Volterra equations (A.1.3.1) is to generalize the constant expression used for G_s in (A.1.2.2) with an affine expression instead, while still leaving the absence of constraints on growth unchanged when compared to (A.1.2.2). Therefore, the competitive Lotka-Volterra equations (A.1.6.1), in using the linear expression for the growth-constraining term H_s, and leaving the growth describing factor G_s as a constant and thus ungeneralized, seems to not follow the "spirit" of the generalized Lotka-Volterra equations (A.1.3.1), and thus did not seem adequately comparable to prior work [VCF+18] [FM14] [SBT+13] [MMV+08] using those equations.

Appendix A.2

Simulating Growth of Microbes in Droplets

In practice, observed measurements of the numbers of reads $\vec{\mathbf{R}}_d(t_b)$ for each strain in a droplet d from batch b, meant to correspond to the numbers of cells from each strain currently alive $\vec{\mathbf{N}}_d(t_b)$ in droplet d at time t_b, will instead better correspond to the numbers of cells which had been alive at any time before time t_b (and which have not yet had their genetic material scavenged by other microbes). Compare with Figure A.2.1. This discrepancy is because all cells, living or dead, must be lysed in order to be able to sequence the contents of the (previously) living cells. This leads to the contents of any cells which had died before being lysed to most likely also have their genetic material sequenced. One of the goals of the simulations was to explicitly account for this discrepancy between theory and measurements.
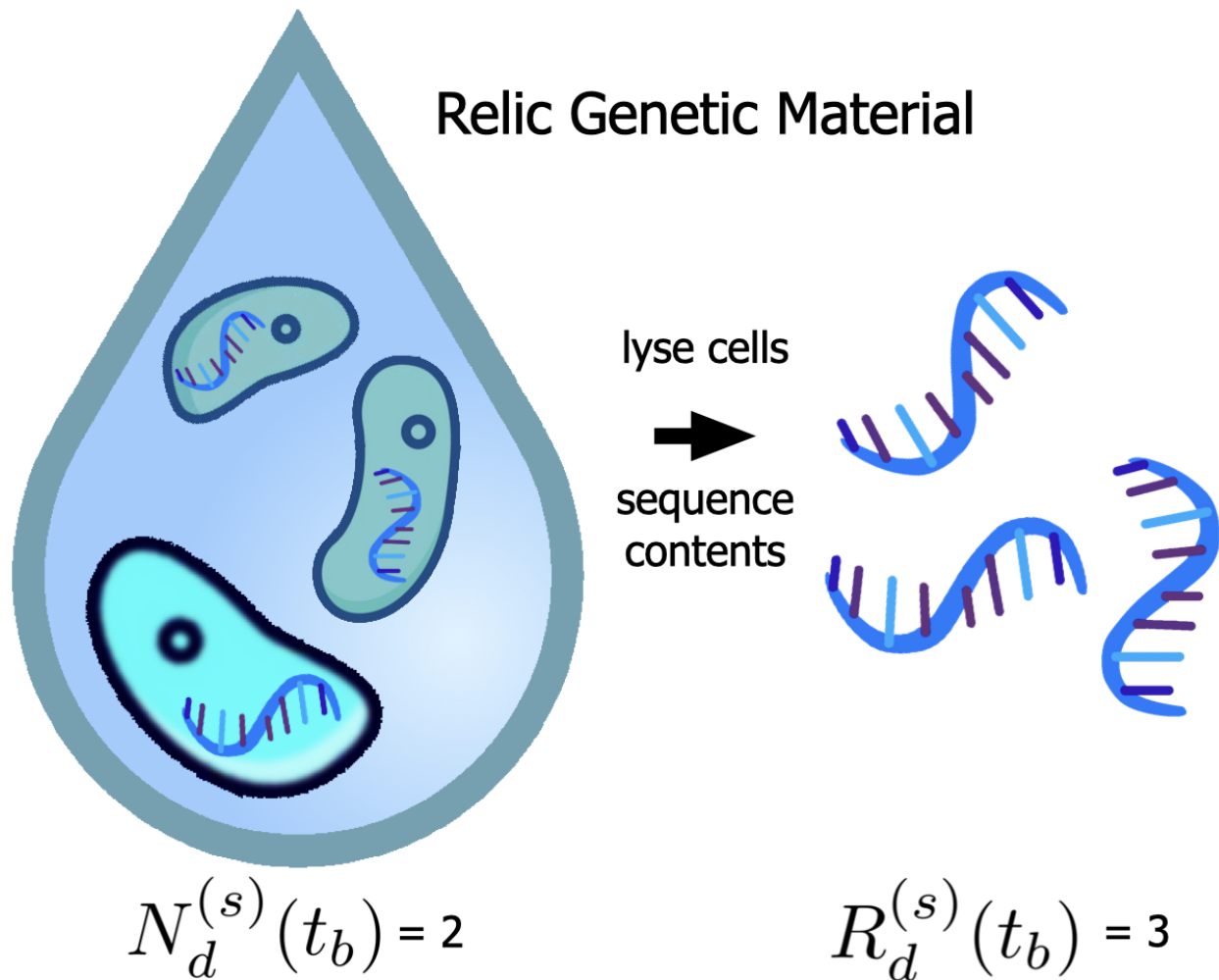
Figure A.2.1: At the end of the growth period, two cells are still alive in the droplet, but one is dead. After lysing the cells and sequencing their contents, the reported number of reads is three. The dead cell is effectively counted as if it were alive. Cf. figure 1.5.

Another goal of these simulations was to explicitly account for the reality of resource limitations in these small droplets, which clearly will not allow for indefinite exponential growth. We required that any solution achieving that goal also not impose any restrictions on the allowed interaction coefficients, since there would not seem to be any biological justification for any such restrictions. In particular, such restrictions seem to most often be made for no other reason than preventing unbounded growth. Thus, in principle, a solution correctly modelling resource limitations should work for any interaction coefficients. Resource limitations inherently limit growth and should be sufficient to do so without artificial constraints on the interaction coefficients. (Cf. the argument in section A.1.4.)

A further goal was to have a simulation model that was comparable to prior work

[VCF$^+$18] [SBT$^+$13] [MMV$^+$08] [FM14] [MBH$^+$14], while also not sacrificing accuracy or relevance for this problem in terms of failing to accomplish the first and second goals. For example, although it was found in [VCF$^+$18] that the generalized Lotka-Volterra equations provided a good approximation to the observed dynamics, this was for manually grown plate cultures which were continuously replated as necessary in order to ensure that the measured populations never encountered resource deprivations. Chapter A.1 describes in detail how the third and second goals were reconciled.

One benefit of satisfying the last goal above is that it allows looking at the performance of the parametric estimator described in appendix A.3 in a context which is both "fair" and "realistic". "Fair" in the sense of being reasonably similar to the models used implicitly in prior work [VCF$^+$18] [SBT$^+$13] [MBH$^+$14] [MMV$^+$08] [FM14] from which it was derived. "Realistic" in the sense that the original models are still significantly misspecified[1].

A.2.1 Modelling Cell Counts under Nutrient Limitation

Starting from equation (A.1.4.1), a discretization is applied to define an Euler step update, and then a stochastic noise term ε is added to yield

$$
\log\left(N^{(s)}(t+\Delta t)\right) - \log\left(N^{(s)}(t)\right)
$$
$$
= \begin{cases} \left(\Delta t\left(r_s + \sum_{\sigma=1}^{S}\alpha_{\sigma s}N^{(\sigma)}(t)\right) + \varepsilon\right)\dfrac{\left(K - \sum_{\sigma=1}^{S}N^{(\sigma)}(t)\right)}{K} & \text{when } N^{(s)}(t) \neq 0 \\ 0 & \text{when } N^{(s)}(t) = 0 \end{cases} \quad \text{(A.2.1.1)}
$$

where all of the ε are i.i.d. Gaussian$(0, \eta^2\Delta t)$ for some "noise scale" $\eta > 0$ (the second parameter is the variance, so e.g. the standard deviation is $\eta\sqrt{\Delta t}$).

The equation (A.2.1.1) has a relationship with (A.1.4.1) that is analogous to the relationship that the assumed updates (2.2.15) from previous work [VCF$^+$18] [MMV$^+$08] [FM14] [SBT$^+$13] have with the generalized Lotka-Volterra equations (A.1.3.1).

This may be equivalent (cf. [KP10, section 6.2 or part IV] to check my work, because I

---

[1] I assume that parametric models are almost always misspecified in practice. Cf. [vdLR11] for discussion.

doubt this is entirely correct) to a discretization of the following SDE model:

$$
\mathrm{d}N^{(s)}(t)
$$
$$
= \begin{cases} N^{(s)}(t)\left(\mathrm{d}t\left(r_s + \displaystyle\sum_{\sigma=1}^{S}\alpha_{\sigma s}N^{(\sigma)}(t)\right) + \eta\mathrm{d}W(t)\right)\dfrac{\left(K - \displaystyle\sum_{\sigma=1}^{S}N^{(\sigma)}(t)\right)}{K} & \text{when } N^{(s)}(t) \neq 0 \\ 0 & \text{when } N^{(s)}(t) = 0 \end{cases}
$$
$$
\text{(A.2.1.2)}
$$

where $\mathrm{d}W(t)$ denotes a standard Wiener process/Brownian motion.

Using the notation of equation (A.1.5.1) from chapter A.1, the update (A.2.1.1) can be written as

$$
\log\left(N^{(s)}(t+\Delta t)\right) - \log\left(N^{(s)}(t)\right) = \begin{cases} ((\Delta t)G_s + \varepsilon)\,\frac{K-H_s}{K} & \text{when } N^{(s)}(t) \neq 0 \\ 0 & \text{when } N^{(s)}(t) = 0 \end{cases} \quad\text{(A.2.1.3)}
$$

using the definitions

$$
\begin{aligned}
G_s &= r_s + \sum_{\sigma=1}^{S}\alpha_{\sigma s}N^{(\sigma)}(t) \\
H_s &= \sum_{\sigma=1}^{S}N^{(\sigma)}(t)
\end{aligned}
\qquad\text{(A.2.1.4)}
$$

It is possible, due to the errors introduced by both discretization and stochasticization, that the population of a droplet in the simulation might temporarily exceed the carrying capacity, which corresponds to a special case of $\mathrm{sign}(K-H_s) = -1$. Without modification to equation (A.2.1.1), this would lead to a biologically extremely unrealistic indefinite explosion in population growth for any strain s whose growth due to microbial interactions G_s was negative (i.e. decreasing growth). Defining

$$
\varphi(x,y) := \begin{cases} -\,\mathrm{sign}(x)\,\mathrm{sign}(y) & \mathrm{sign}(x) = \mathrm{sign}(y) = -1 \\ 1 & \text{else} \end{cases}, \qquad\text{(A.2.1.5)}
$$

the aforementioned problem is fixed by using the update equation

$$
\log\left(N^{(s)}(t+\Delta t)\right) - \log\left(N^{(s)}(t)\right) = \begin{cases} \varphi(G_s,(K-H_s))\,((\Delta t)G_s + \varepsilon)\,\frac{K-H_s}{K} & N^{(s)}(t) \neq 0 \\ 0 & N^{(s)}(t) = 0 \end{cases}
$$
$$
\text{(A.2.1.6)}
$$

whose corresponding system of differential equations

$$
\frac{\mathrm{d}(\log n^{(s)})}{\mathrm{d}t} = \varphi(G_s,(K-H_s))G_s(n^{(1)},\ldots,n^{(S)})\left(1 - \frac{H_s(n^{(1)},\ldots,n^{(S)})}{K}\right), \qquad\text{(A.2.1.7)}
$$

has solutions which vary smoothly in phase space except for a region of measure zero, where they still vary continuously. Thus any concerns about the realism of equation (A.2.1.3) due to non-smoothness seem likely to be minor in comparison to e.g. the known potential issues for correctly modelling the dynamics of even the unmodified generalized Lotka-Volterra equations (A.1.3.1) in low dimensions [LE01] (i.e. few strains, e.g. 2-3) introduced by discretization (especially via Euler steps, cf. e.g. [Ush82]) or stochasticization (discussed e.g. in the case of continuous time SDE models in [Ara03]). Empirically, it seems that with Δt chosen to be small enough, the dynamics should (almost) always be constrained to those of (A.2.1.3), so the dynamics in (A.2.1.6) can be considered a fallback or last-resort failsafe that usually is irrelevant but nevertheless guarantees no unbounded growth.

A.2.2   Read Counts vs Cell Counts

In the simulation used for the subsequent analysis, the updates used in (A.2.1.3) are used to define what is assumed to be the true (pseudo)count of the numbers of cells of each strain in each droplet. (Fractional values of the pseudocounts can, for the sake of argument, be assumed to represent incomplete cell division.) We can make further changes so that the simulation keeps track of both (1) the actual number of cells $\vec{\mathbf{N}}_d(t)$ (until $t = t_b$) that would be predicted to be observed as a consequence of the updates (A.2.1.3) using a perfect experimental apparatus, as well as (2) the number of (16S V4 region) reads $\vec{\mathbf{R}}_d(t)$ that are likely to be observed in practice.

Three modifications are used to model the number of reads for each strain which the researcher is likely to measure.

One modification accounts for the accumulation of relic genetic material (RNA/DNA) in the droplets, a second modification accounts for the fact that different strains are likely to have different copy numbers of the 16S rRNA gene, and a final modification accounts for the possibility of nonzero droplet merging error at the end of the experiment. Note that no attempts were made to directly account for stochasticities in the amount of marker introduced in each droplet (which is assumed to have a low signal-to-noise ratio $\frac{\sigma}{\mu}$ due to the efforts of the experimenter), or for potential PCR amplification biases. However the modification employed to account for 16S rRNA copy number variations resembles (at least superficially) the modelling of PCR amplification bias proposed in previous work [MWC19], and thus can likely without loss of generality be assumed to account for both.

$$N^{(s_1)}(t_b) = 2$$

$$N^{(s_2)}(t_b) = 1$$

$$R^{(s_1)}(t_b) = 2$$

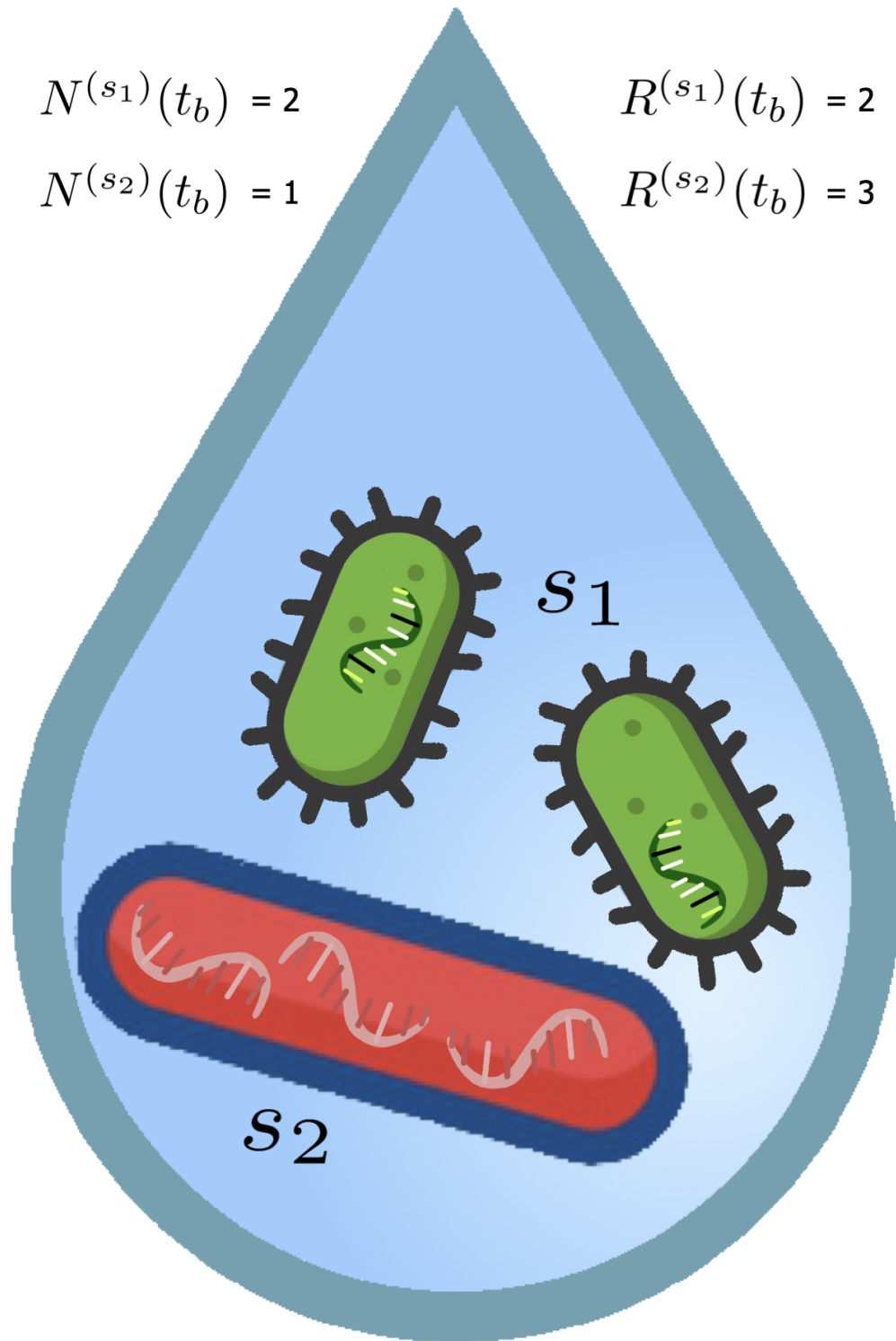$$R^{(s_2)}(t_b) = 3$$

s_1

s_2

Figure A.2.2: One strain outnumbers the other strain 2:1. Yet because the latter strain has a larger 16S gene copy number, the observed ratio of reads is 2:3.

For each time step, the change in the observed number of reads $R_d^{(s)}(t)$ for each strain s in a given droplet d is assumed to follow from the true (pseudo)counts $N_d^{(s)}(t)$ of the numbers of cells for each strain using the update steps:

$$\vec{\mathbf{R}}_d(t_{\beta+1}) - \vec{\mathbf{R}}_d(t) = \left(\vec{\mathbf{B}} * \left(\vec{\mathbf{N}}_d(t+\Delta t) - \vec{\mathbf{N}}_d(t) \right) \right)_+, \tag{A.2.2.1}$$

where $*$ denotes the entrywise or Hadamard product, $(x)_+$ denotes the positive part of x, i.e. $\max\{x, 0\}$, and $\vec{\mathbf{B}}$ denotes the vector whose entries contain the 16S rRNA copy numbers (and/or PCR amplification biases) for each strain. The idea is to take the total number of cells of each strain which have ever been alive in the droplet and then multiply those numbers by the corresponding 16S copy numbers/PCR amplification biases. This is since the total number of currently living cells $n^{(s)}(t) = \int_0^t \frac{dn^{(s)}(\tau)}{d\tau} d\tau$ is equal to the total number of cells which have ever lived $\int_0^t \left(\frac{dn^{(s)}(\tau)}{d\tau} \right)_+ d\tau$ minus the total number of cells which have ever died $\int_0^t \left(\frac{dn^{(s)}(\tau)}{d\tau} \right)_- d\tau$, $(x)_- := \max\{-x, 0\}$. Compare with Figure A.2.3.
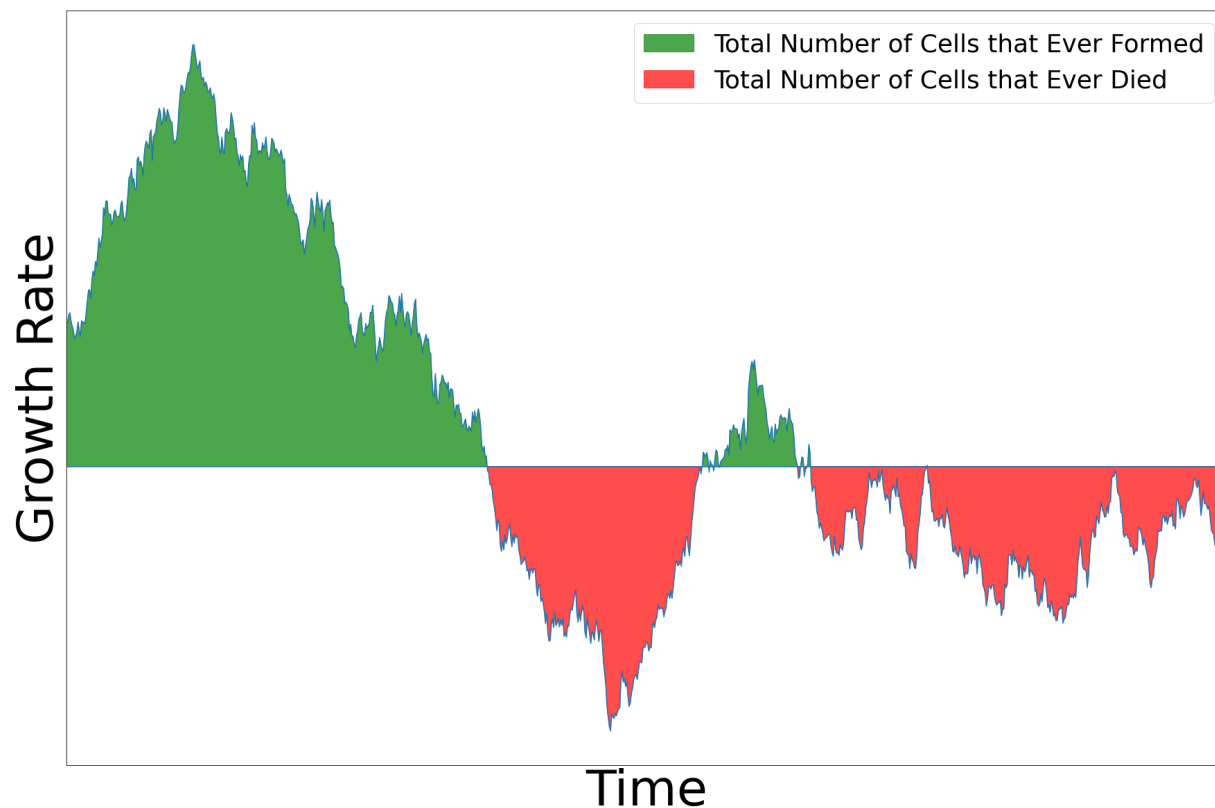
Figure A.2.3: The true number of cells observed at the end of the simulation is the green area minus the red area, whereas the number of reads actually observed at the end of the simulation is proportional to the green area only. *(NB: This is a heuristic plot illustrating the idea and not data from an actual simulation.)*

Thus the difference between the total number of cells which have ever lived and the total number of currently living cells, after being multiplied by the copy numbers and/or PCR amplification biases $\vec{\mathbf{B}}$, can be assumed to approximately correspond to the amount of relic genetic material (RNA/DNA) in the droplet. (At least assuming little to no cell lysis and/or nucleoside scavenging has occurred before the time b when the droplet's contents are finally measured.) Thus vector $\vec{\mathbf{B}}$ (entrywise) multiplied by the total number of cells which have ever lived can be assumed to correspond to the total number of reads observed from both living cells and from relic genetic material, thus (A.2.2.1) encompasses both of the first two aforementioned modifications.

Regarding the last of the three aforementioned modifications, at the end of the simulation, if a non-zero merging error rate m is assumed, then $\lfloor D \cdot (m/2) \rfloor$ droplets are randomly selected without replacement to be merged into other droplets, and another $\lfloor D \cdot (m/2) \rfloor$ droplets are randomly selected without replacement to be the droplets which are randomly merged into. Then for each of the $\lfloor D \cdot (m/2) \rfloor$ pairs of droplets, a new droplet consisting of the sum of the

observed contents from each droplet of the pair is created, with the values from the two older droplets discarded (as they are assumed to be unobserved). (The indices of which droplets were merged with which other droplets are recorded, and the original droplets for the cell pseudocounts are left unchanged.)

A.2.3   Modelling PCR Amplification

A (naively) simple model assumes that the PCR amplification factor for each kind k of genetic material (either the spikein gene or the 16S rRNA gene of one of the S strains) equals a constant a multiplied by a kind-specific instance of a log-normal random variable $\xi^{(k)}$. (This assumes that some sort of multiplicative central limit theorem is applicable.) Thus after amplification the (amplified) reads $\vec{\mathbf{A}}_d(t_b)$ corresponding to droplet d would equal

$$\vec{\mathbf{A}}_d(t_b) := \vec{\mathbf{R}}_d(t_b) * (a_d \vec{\boldsymbol{\xi}}) = \left(a_d \xi_d^{(1)} R_d^{(1)}(t_b), \ldots, a_d \xi_d^{(s)} R_d^{(s)}(t_b), \ldots, a_d \xi_d^{(S)} R_d^{(S)}(t_b) \right) . \quad \text{(A.2.3.1)}$$

Letting M denote the copy number for the spikein gene (which recall is Poisson distributed, with a high expectation, cf. footnote 13 from section 1.4.3), it follows that after normalizing by the observed number of spikein gene amplicons, the observed number of reads $\vec{\mathbf{O}}_d(t_b)$ corresponding to droplet d is

$$\vec{\mathbf{O}}_d(t_b) := \left(\frac{1}{a_d \xi_d^M M_d} \right) \cdot \vec{\mathbf{A}}_d(t_b) = \left(\frac{\xi_d^{(1)}}{\xi_d^M M_d} R_d^{(1)}(t_b), \ldots \frac{\xi_d^{(s)}}{\xi_d^M M_d} R_d^{(s)}(t_b), \ldots, \frac{\xi_d^{(S)}}{\xi_d^M M_d} R_d^{(S)}(t_b) \right) .$$
$$\text{(A.2.3.2)}$$

Then the hope is that $\mathbb{E}[M] \cdot \vec{\mathbf{O}}_d(t_b) \approx \vec{\mathbf{N}}_d(t_b)$, or at least $\mathbb{E}[M] \cdot \vec{\mathbf{O}}_d(t_b) \approx \vec{\mathbf{R}}_d(t_b)$, "on average" (i.e. the variance of the $\log(\xi_d)$'s is not too large).

A.2.4   Example Cell and Read Count Trajectories

For all simulations, the carrying capacity K was set to $K = 268$ (based on what my collaborator said was realistic). Likewise for all simulations, the timestep Δt was set to 0.0001, the "noise scale" was set to $\eta = 8$, and the cells were assumed to be grown "in the fifth batch" with 200 timesteps separating each batch (so the simulation was ran for $1,000$ time steps total). All simulations were initialized with one cell each of the two strains.

The copy numbers for both strains were always set to be 1.
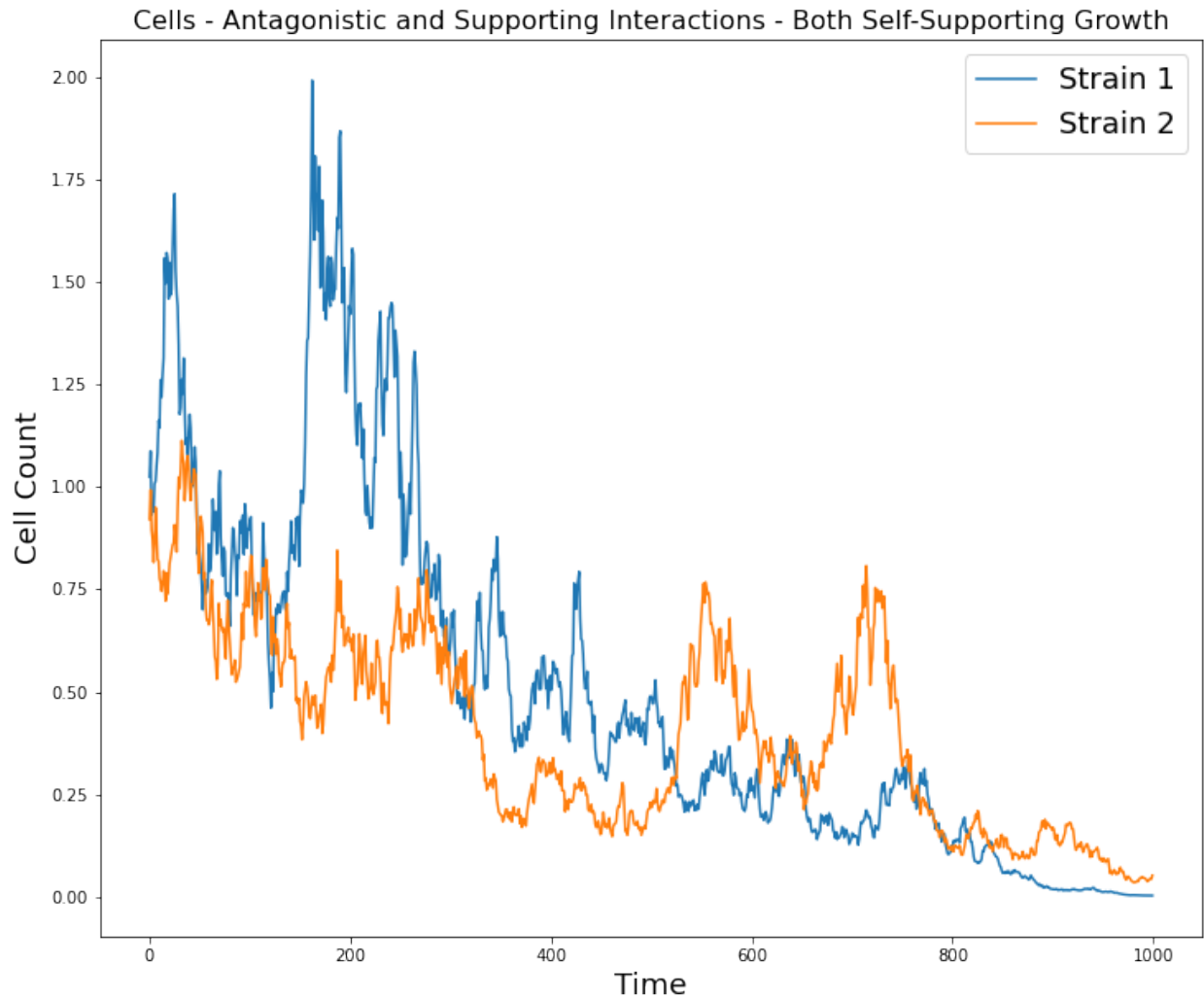
A.2.4.1   Cell Count Trajectories

Figure A.2.4:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.09417735 \\ 0.97562235 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0.5479121 & -0.12224312 \\ 0.71719584 & 0.39473606 \end{bmatrix}$$
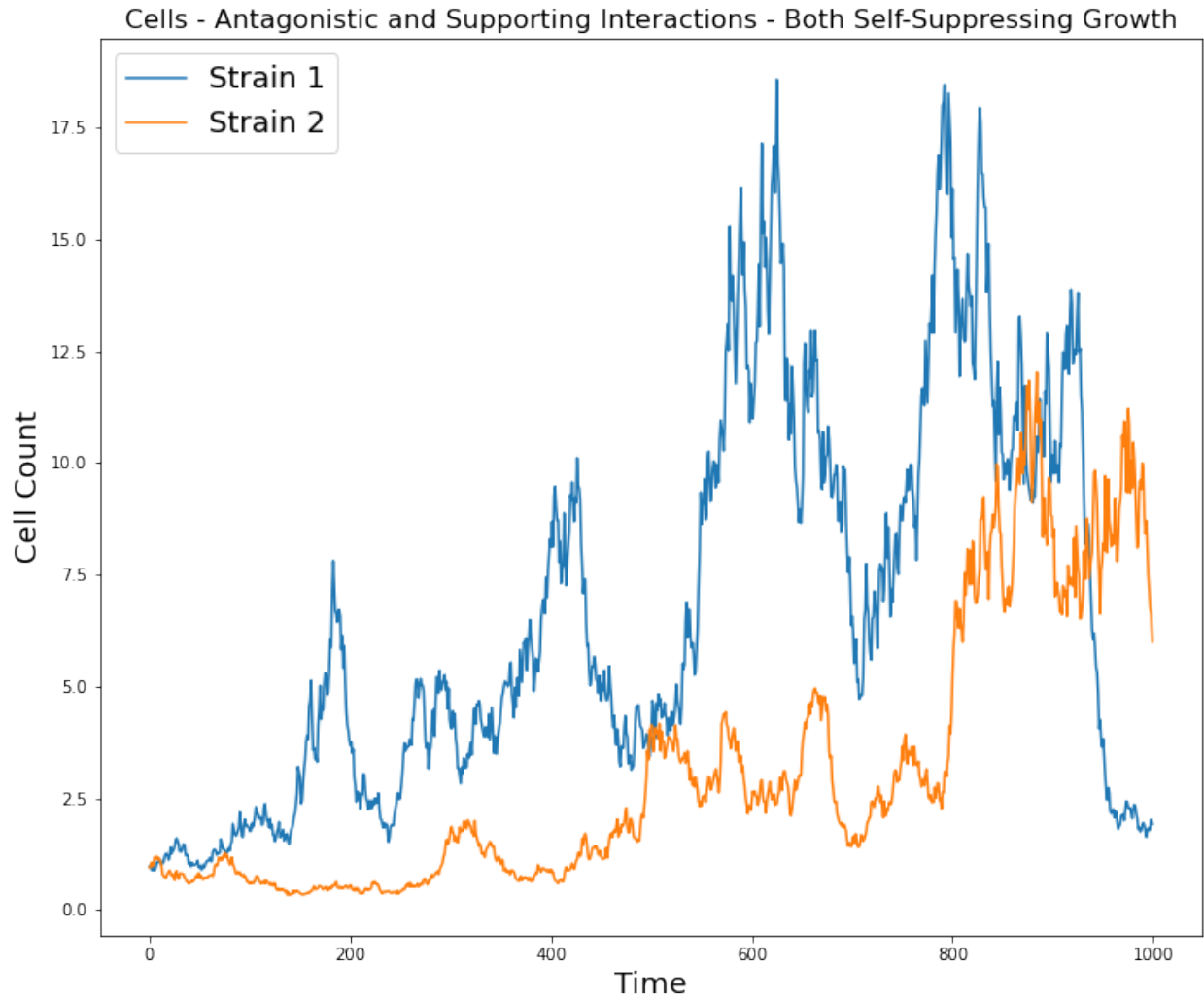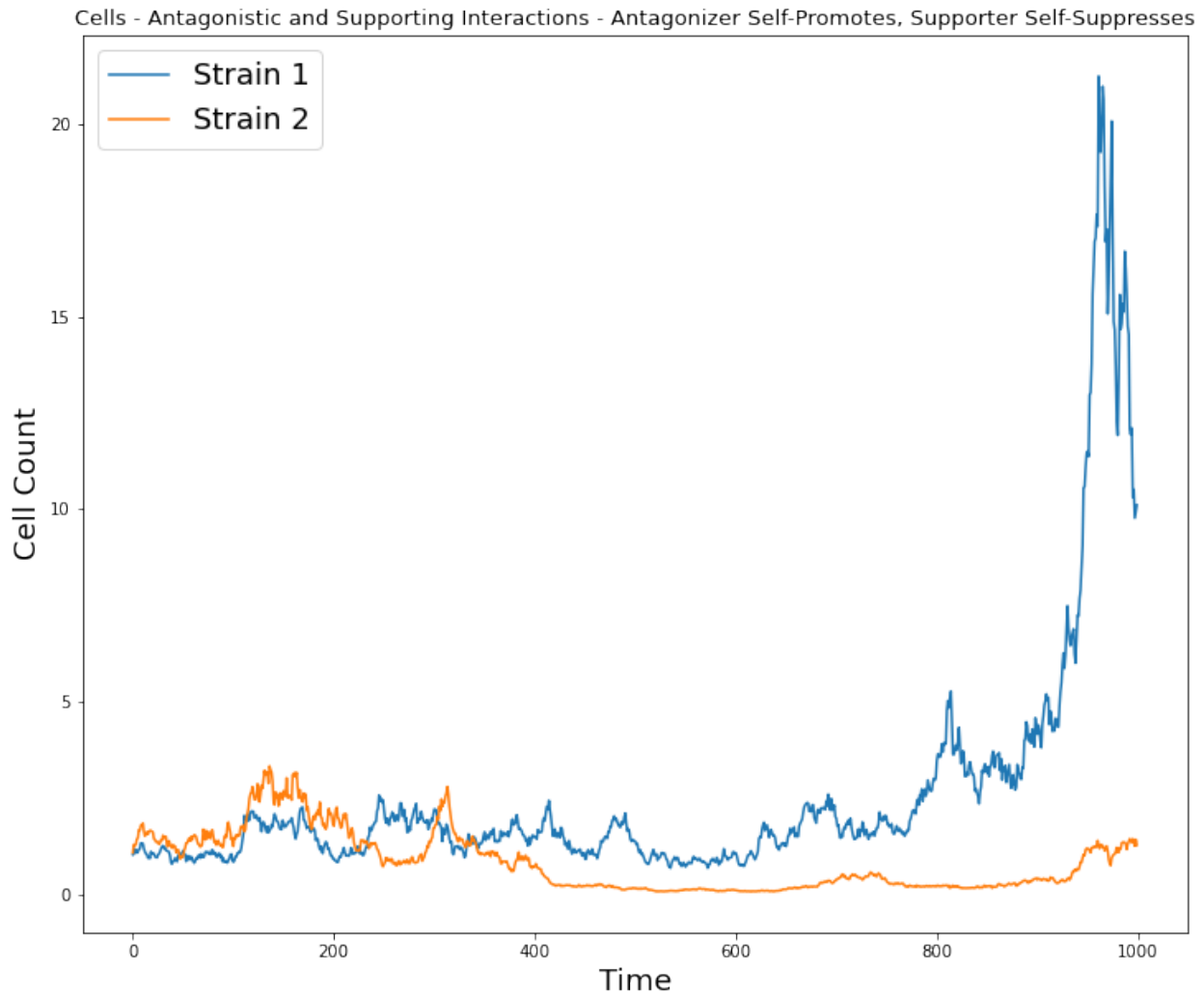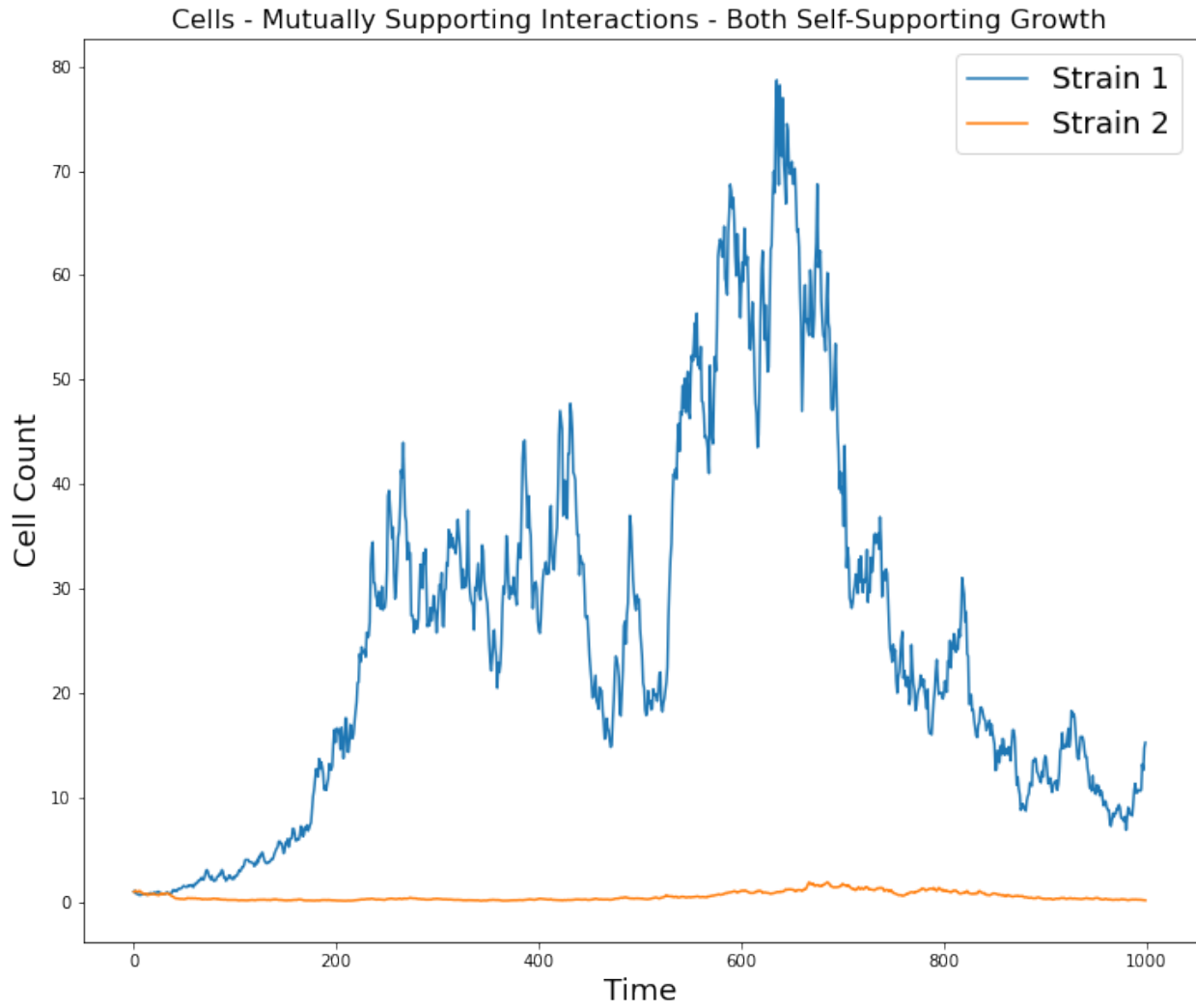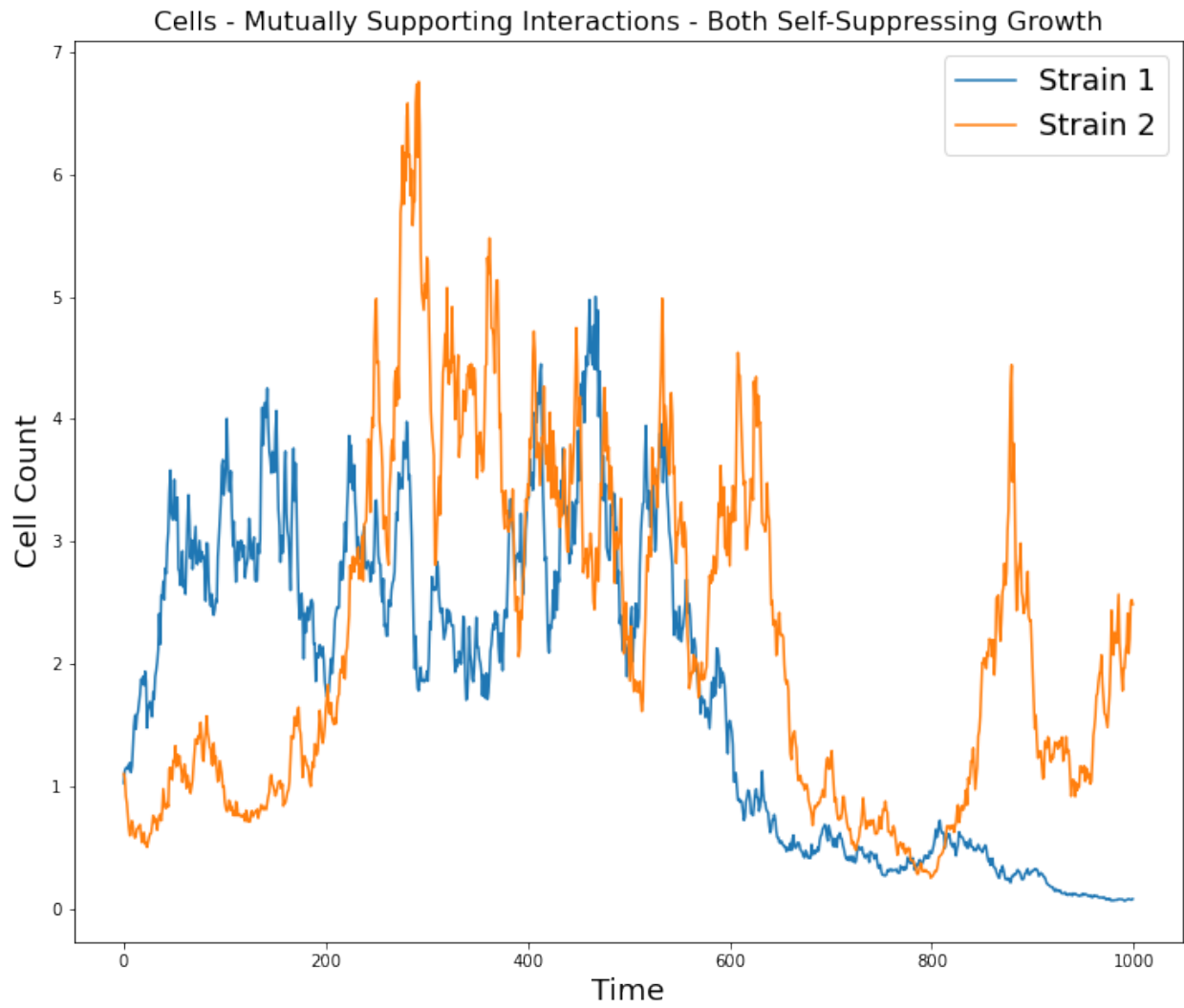
Figure A.2.5:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.4797985 \\ 0.97804575 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.96753909 & -0.09498414 \\ 0.68530924 & -0.35219747 \end{bmatrix}$$

Figure A.2.6:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.92088119 \\ 0.26796687 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0.62393464 & -0.37883923 \\ 0.56723451 & -0.34407233 \end{bmatrix}$$
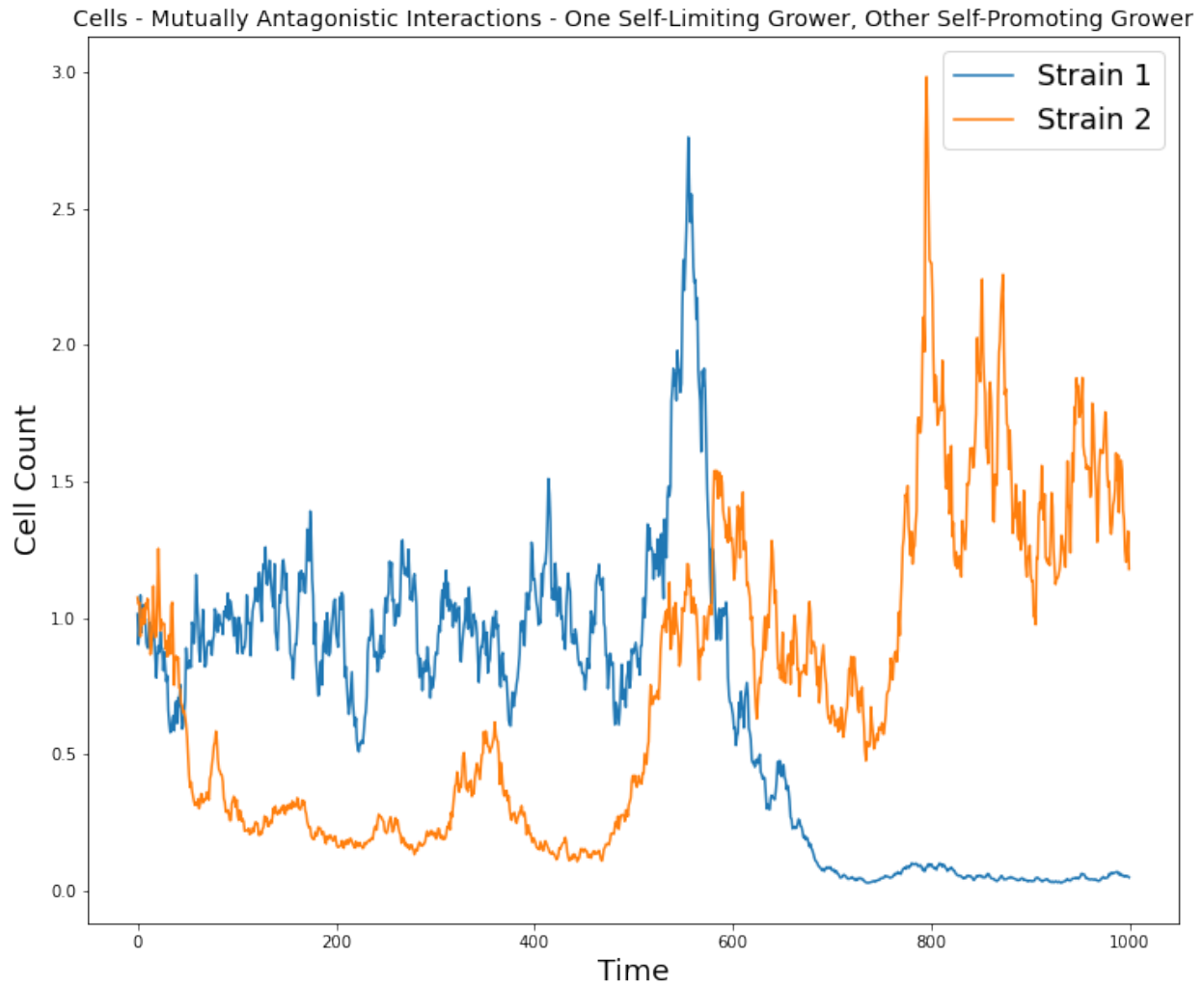
Figure A.2.7:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.32162234 \\ 0.82774067 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0.42400205 & 0.45574743 \\ 0.83065306 & 0.03716118 \end{bmatrix}$$

Figure A.2.8:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.4602045 \\ 0.34569703 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.14107603 & 0.78315155 \\ 0.30476027 & -0.204964 \end{bmatrix}$$

Figure A.2.9:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.22324414 \\ 0.69651731 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.15129588 & -0.05373991 \\ -0.41705514 & 0.58090085 \end{bmatrix}$$
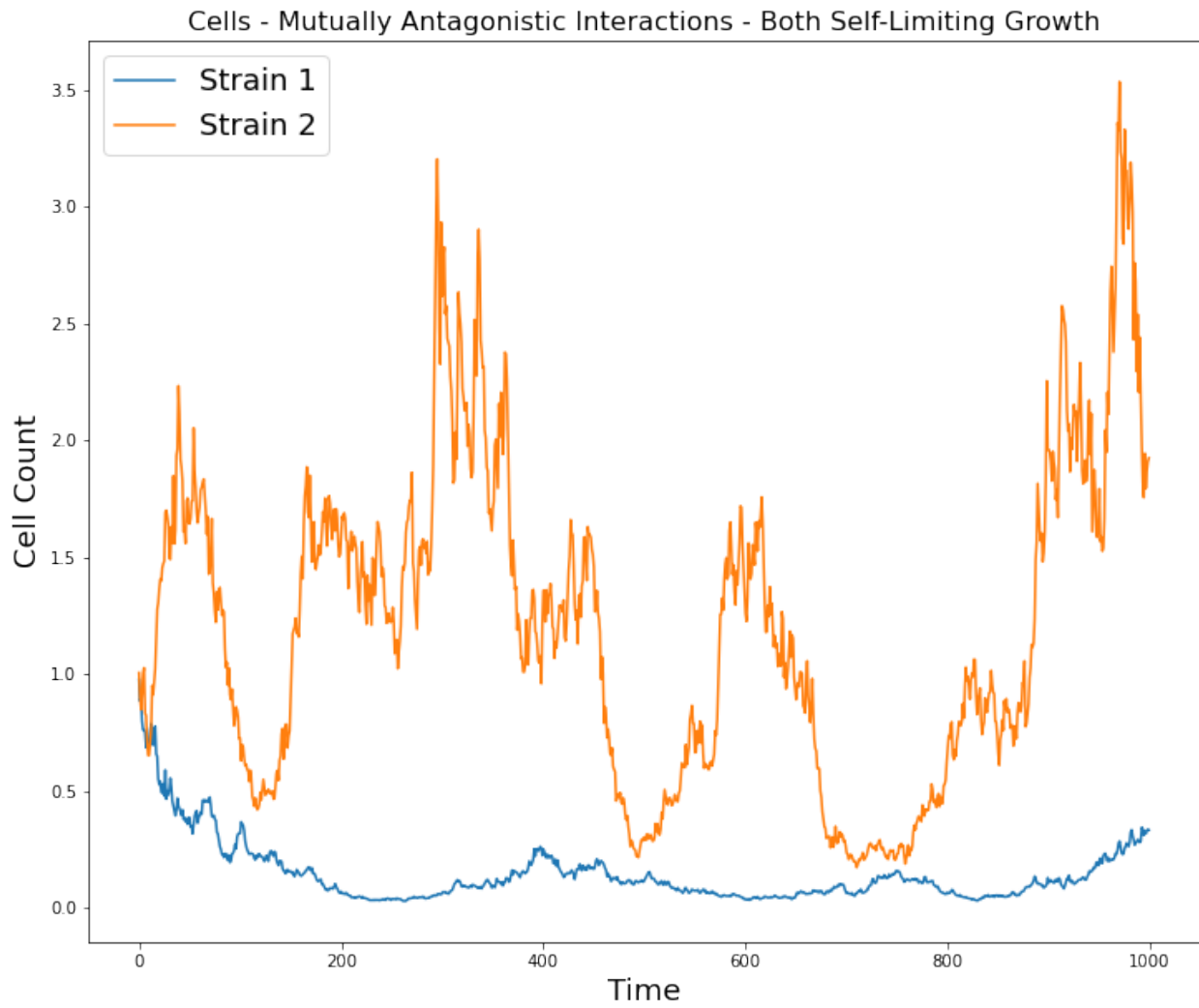
Figure A.2.10:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.3970483 \\ 0.18062232 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.07416439 & -0.17117972 \\ -0.79256721 & -0.27501624 \end{bmatrix}$$
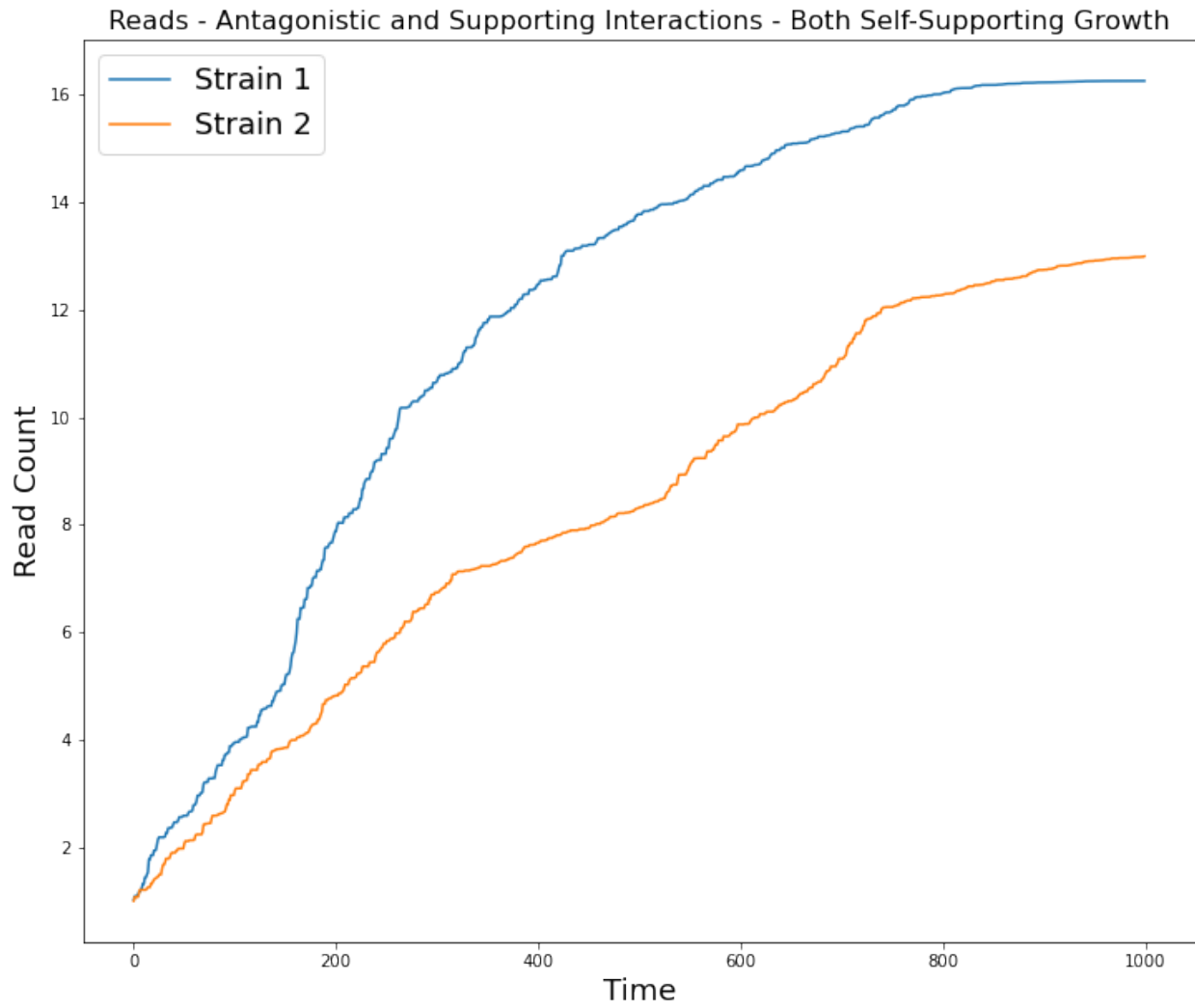
A.2.4.2 Read Count Trajectories

Figure A.2.11:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.09417735 \\ 0.97562235 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0.5479121 & -0.12224312 \\ 0.71719584 & 0.39473606 \end{bmatrix}$$
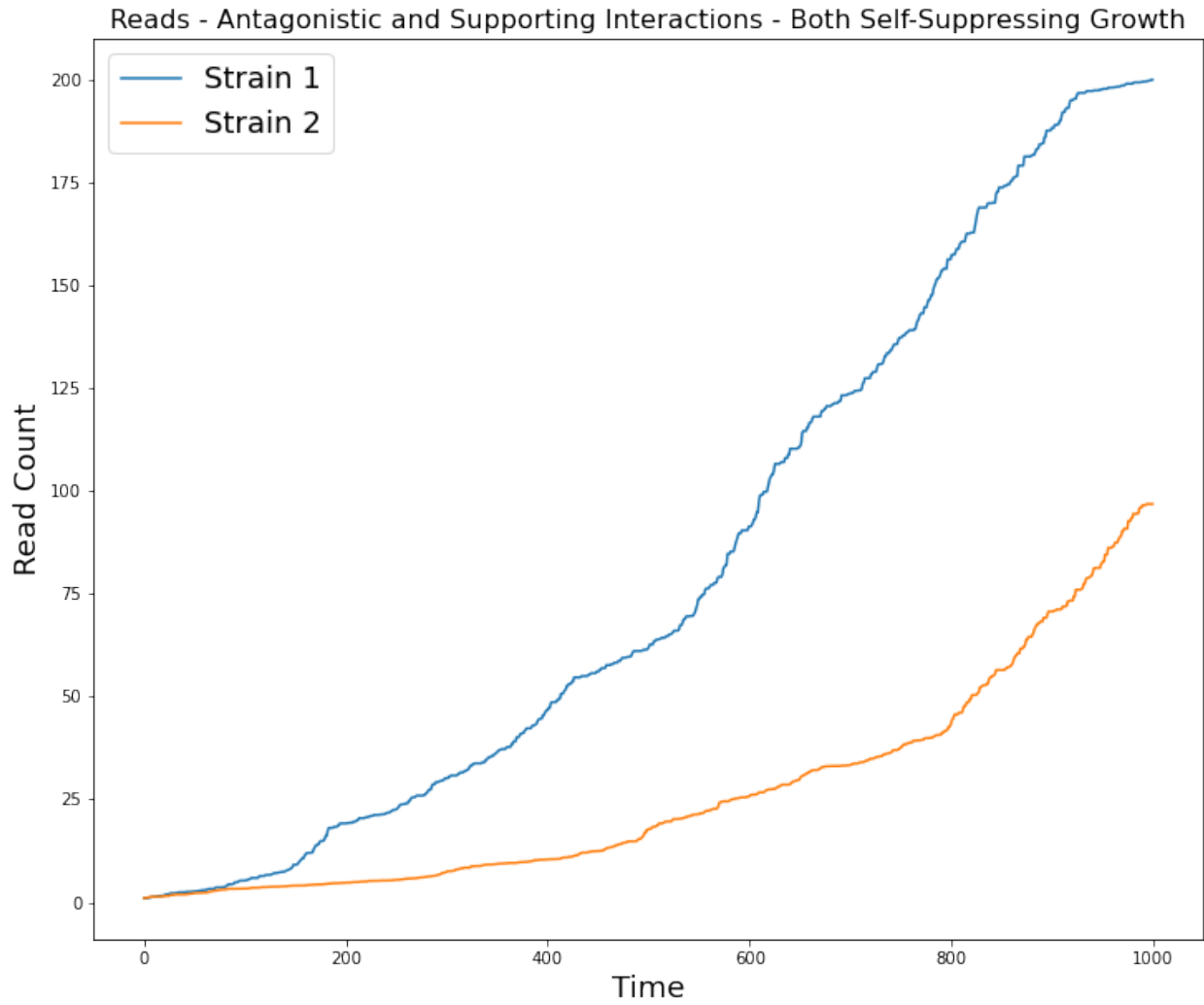
Figure A.2.12:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.4797985 \\ 0.97804575 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.96753909 & -0.09498414 \\ 0.68530924 & -0.35219747 \end{bmatrix}$$

Reads - Antagonistic and Supporting Interactions - Antagonizer Self-Promotes, Supporter Self-Suppresses
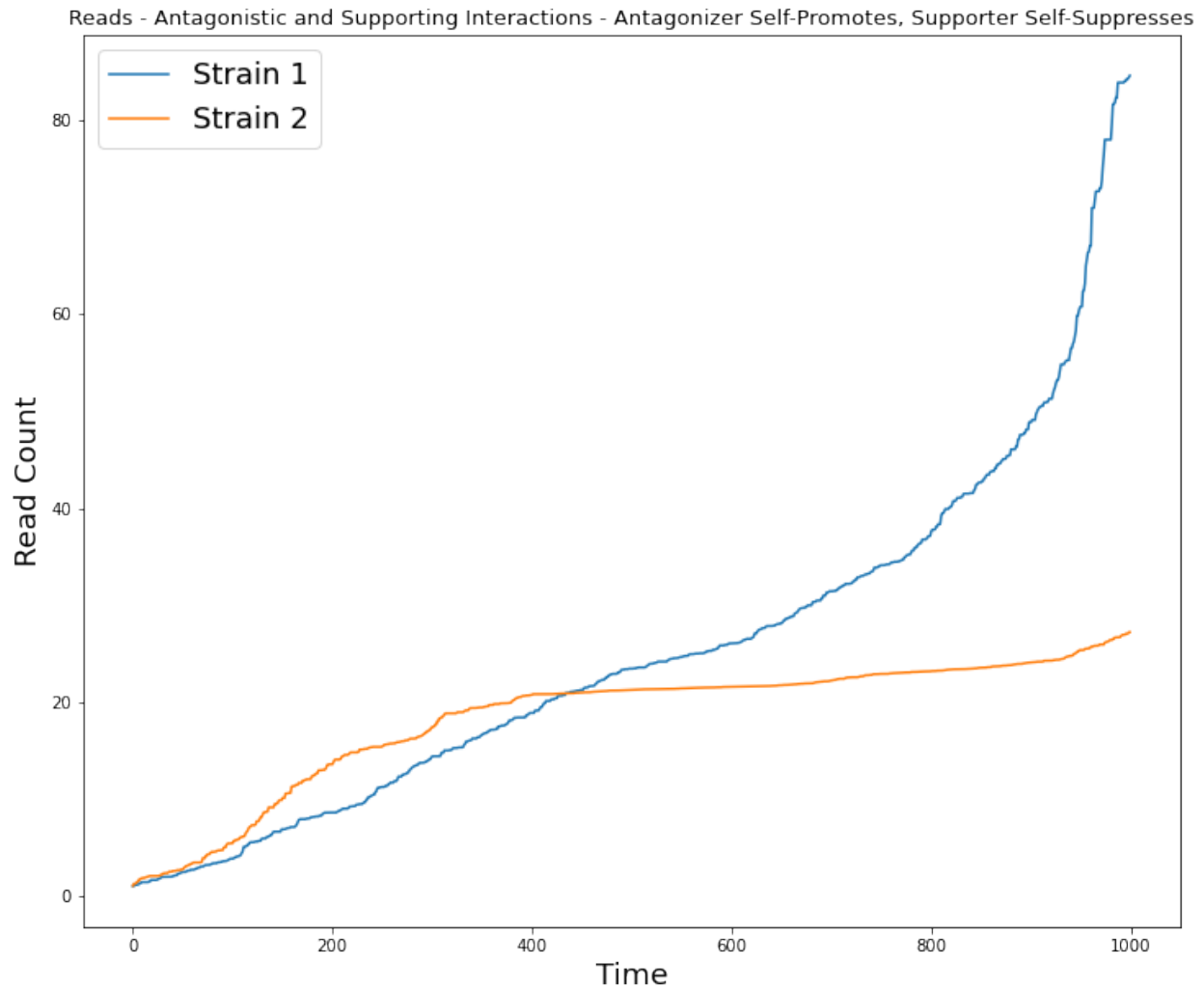
Figure A.2.13:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.92088119 \\ 0.26796687 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0.62393464 & -0.37883923 \\ 0.56723451 & -0.34407233 \end{bmatrix}$$
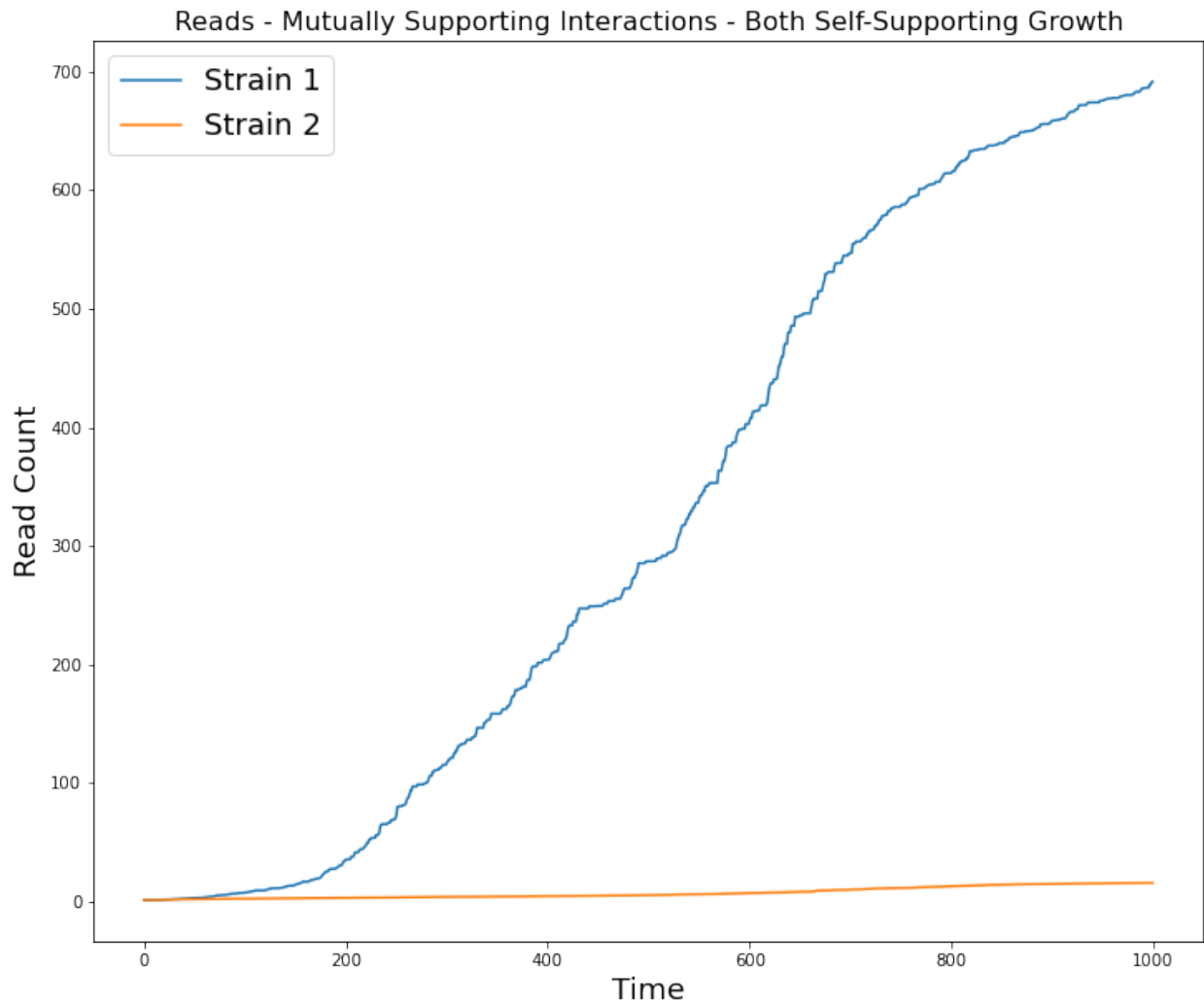
Figure A.2.14:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.32162234 \\ 0.82774067 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0.42400205 & 0.45574743 \\ 0.83065306 & 0.03716118 \end{bmatrix}$$
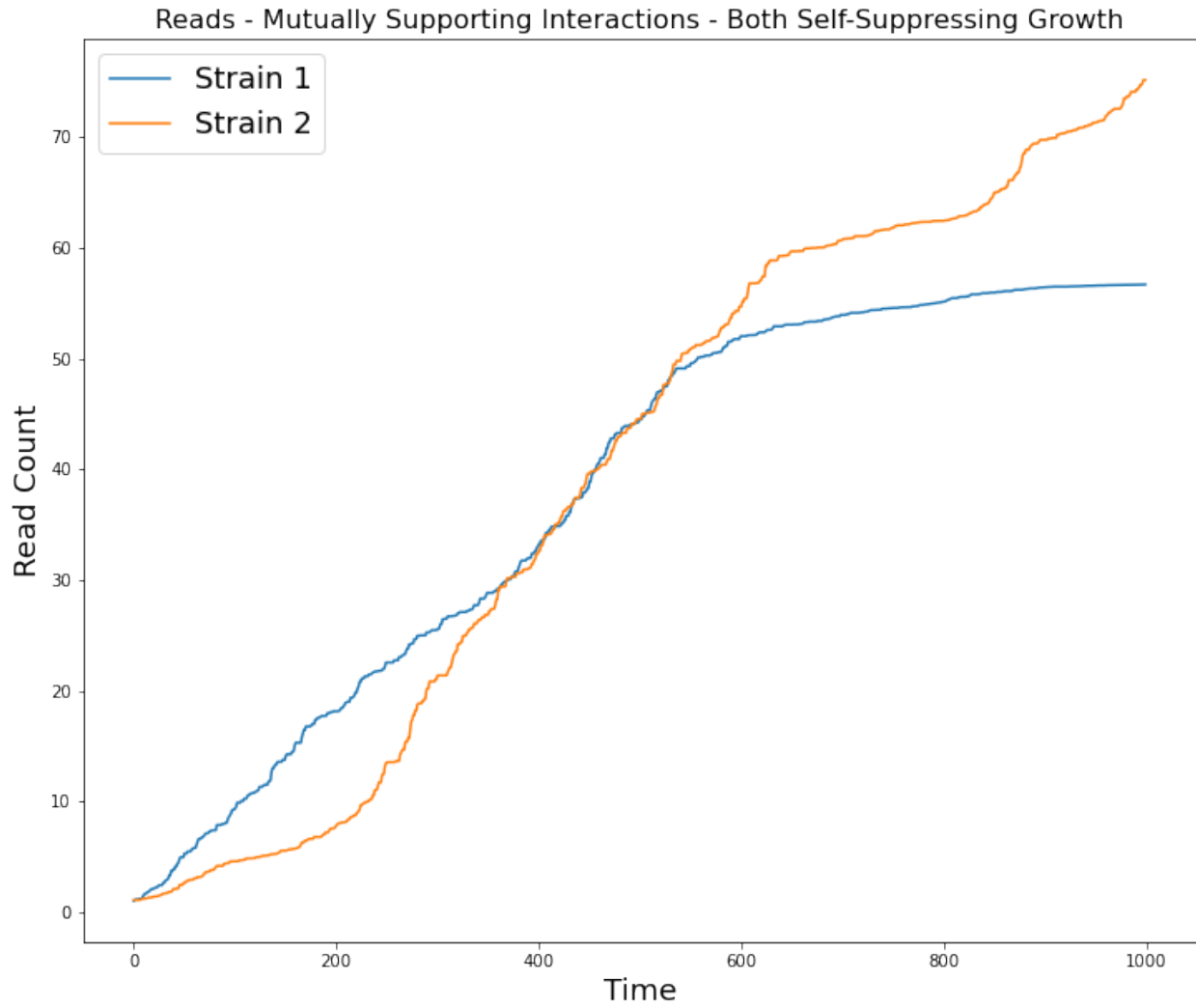
Figure A.2.15:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.4602045 \\ 0.34569703 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.14107603 & 0.78315155 \\ 0.30476027 & -0.204964 \end{bmatrix}$$
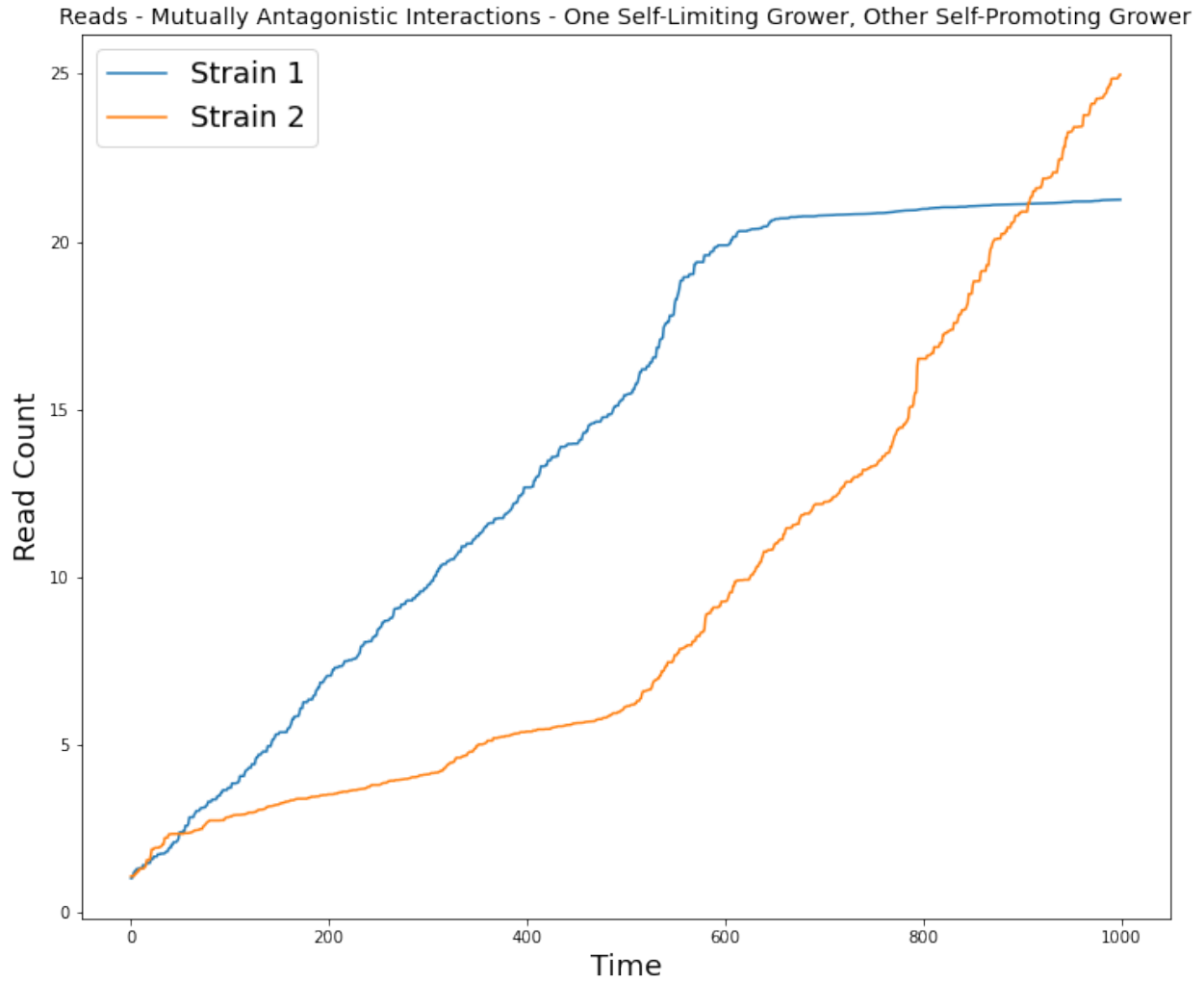
Figure A.2.16:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.22324414 \\ 0.69651731 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.15129588 & -0.05373991 \\ -0.41705514 & 0.58090085 \end{bmatrix}$$
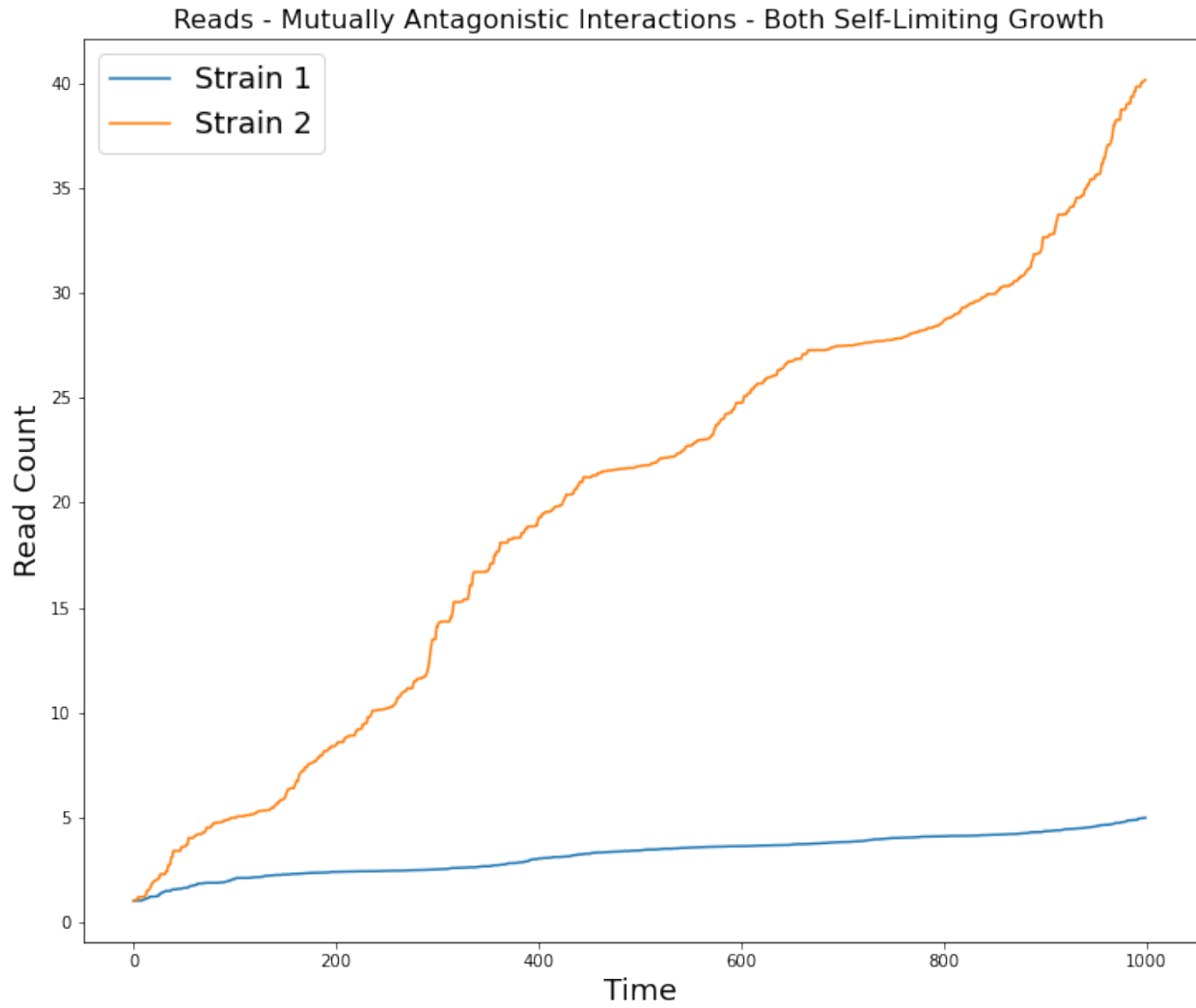
Figure A.2.17:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0.3970483 \\ 0.18062232 \end{bmatrix}, \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} -0.07416439 & -0.17117972 \\ -0.79256721 & -0.27501624 \end{bmatrix}$$

Appendix A.3

Parametric Estimation of Interaction Coefficients

The goal of this section is to modify the regression-based methodology used in [VCF⁺18] [SBT⁺13] [MBH⁺14] [FM14] [MMV⁺08] to not require longitudinal data, and to be applicable to pseudo-longitudinal data. (Unfortunately preliminary data suggests that the resulting methodology does not work well in practice.) Future work should compare the ideas herein with the ideas in [SSB⁺17], which proposes modifying the aforementioned methodology to be applicable to a different[1] kind of cross-sectional data. Cf. section 2.1.2.

A.3.1 Details

Given any estimand which is a conditional expectation, i.e. of the form $\mathbb{E}[X \mid A]$ for some random variable X and some event A in the underlying sigma-algebra, in what follows the corresponding empirical (arithmetic) mean will be denoted

$$\hat{\boldsymbol{\mu}}[X \mid A] := \frac{\sum_{d \in [D]} X_d \mathbb{I}\{A\}}{\sum_{d \in [D]} \mathbb{I}\{A\}} \,. \tag{A.3.1.1}$$

In particular the unconditional empirical mean is

$$\hat{\boldsymbol{\mu}}[X] := \frac{1}{D} \sum_{d \in [D]} X_d \,. \tag{A.3.1.2}$$

---

[1]I.e. as opposed to the kind of cross-sectional data that pseudo-longitudinal data is.

If one assumes the transition likelihoods are implicitly specified by the relationships:

$$
\begin{cases}
\log(N^{(s)}(t_\beta)) - \log(N^{(s)}(t_{\beta-1})) = r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} N^{(\sigma)}(t_{\beta-1}) + \varepsilon & N^{(s)}(0) > 0 \\
N^{(s)}(t_\beta) = 0 & N^{(s)}(0) = 0
\end{cases},
$$

$$((2.2.15) \text{ revisited})$$

for all $\beta \in [b]$ and all $s \in [S]$ (as discussed in further detail in section 2.2.5), then as a result of the linearity of (conditional) expectation it follows that for any $b \in [B]$

$$\mathbb{E}\big[\log(N^{(s)}(t_b)) \mid N^{(s)}(0) > 0\big] - \mathbb{E}\big[\log(N^{(s)}(t_{b-1})) \mid N^{(s)}(0) > 0\big] \tag{A.3.1.3}$$

$$= r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} \mathbb{E}\big[N^{(\sigma)}(t_{b-1}) \mid N^{(s)}(0) > 0\big]. \tag{A.3.1.4}$$

Thus one would then hope that for any $b \in [B]$ as the number of observations increases

$$\hat{\boldsymbol{\mu}}\big[\log(N^{(s)}(t_b)) \mid N^{(s)}(0) > 0\big] - \hat{\boldsymbol{\mu}}\big[\log(N^{(s)}(t_{b-1})) \mid N^{(s)}(0) > 0\big] \tag{A.3.1.5}$$

$$\approx r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} \hat{\boldsymbol{\mu}}\big[N^{(\sigma)}(t_{b-1}) \mid N^{(s)}(0) > 0\big]. \tag{A.3.1.6}$$

On this basis, below, for all $s \in [S]$ the least squares estimators

$$(\hat{r}_s, \hat{\alpha_{1s}}, \ldots, \hat{\alpha_{Ss}}) := \operatorname*{argmin}_{(\tilde{r}_s, \tilde{\alpha_{1s}}, \ldots, \tilde{\alpha_{Ss}})} \sum_{b=2}^{B} \Bigg\| \Big(\hat{\boldsymbol{\mu}}\big[\log(N^{(s)}(t_b)) \mid N^{(s)}(0) > 0\big] \tag{A.3.1.7}$$

$$- \hat{\boldsymbol{\mu}}\big[\log(N^{(s)}(t_{b-1})) \mid N^{(s)}(0) > 0\big] \Big) \tag{A.3.1.8}$$

$$- \Big(\tilde{r}_s + \sum_{\sigma=1}^{S} \tilde{\alpha_{\sigma s}} \hat{\boldsymbol{\mu}}\big[N^{(\sigma)}(t_{b-1}) \mid N^{(s)}(0) > 0\big] \Big) \Bigg\|_2^2. \tag{A.3.1.9}$$

are considered.

Technically speaking, it is actually necessary to consider the set:

$$\mathscr{B}_s := \big\{ b \in [B] \setminus \{1\} : \text{ the events } \{N^{(s)}(t_b) > 0\} \text{ and } \{N^{(s)}(t_{b-1}) > 0\} \tag{A.3.1.10}$$

$$\text{are both observed at least once} \big\}, \tag{A.3.1.11}$$

such that the computed estimators actually are

$$(\hat{r}_s, \hat{\alpha_{1s}}, \dots, \hat{\alpha_{Ss}}) := \operatorname*{argmin}_{(\tilde{r}_s, \tilde{\alpha_{1s}}, \dots, \tilde{\alpha_{Ss}})} \sum_{b \in \mathscr{B}_s} \left\| \left(\hat{\boldsymbol{\mu}} \left[\log(N^{(s)}(t_b)) \mid N^{(s)}(0) > 0 \right] \right. \right. \tag{A.3.1.12}$$

$$\left. - \hat{\boldsymbol{\mu}} \left[\log(N^{(s)}(t_{b-1})) \mid N^{(s)}(0) > 0 \right] \right) \tag{A.3.1.13}$$

$$\left. - \left(\tilde{r}_s + \sum_{\sigma=1}^{S} \tilde{\alpha}_{\sigma s} \hat{\boldsymbol{\mu}} \left[N^{(\sigma)}(t_{b-1}) \mid N^{(s)}(0) > 0 \right] \right) \right\|_2^2. \tag{A.3.1.14}$$

When $\mathscr{B}_s = \emptyset$, the estimates are marked as missing. Just like for the estimands from chapter 7, in analyses of simulation data, when estimates of the interaction coefficients are missing, they are conservatively set to 0. (I.e. in the absence of any evidence for an interaction, the choice is made to prefer the possibility of making the fase negative error of assuming no interaction exists when one might exist, rather than possibly making the false positive error of assuming an interaction exists when in fact none exists.)

Note that starting from (2.2.15) and considering instead the relationship:

$$\mathbb{E} \left[\log(N^{(s)}(t_b)) \mid N^{(s)}(t_{b-1}) = n, N^{(s)}(0) > 0 \right] \tag{A.3.1.15}$$

$$- \mathbb{E} \left[\log(N^{(s)}(t_{b-1})) \mid N^{(s)}(t_{b-1}) = n, N^{(s)}(0) > 0 \right] \tag{A.3.1.16}$$

$$= r_s + \sum_{\sigma=1}^{S} \alpha_{\sigma s} \mathbb{E} \left[N^{(\sigma)}(t_{b-1}) \mid N^{(s)}(t_{b-1}) = n, N^{(s)}(0) > 0 \right], \tag{A.3.1.17}$$

following reasoning analogous to (A.3.1.12) would lead directly to the linear regression used in e.g. [VCF$^+$18] [MBH$^+$14] [FM14] [SBT$^+$13] [MMV$^+$08] to estimate interaction coefficients. However, estimating the quantities of the form:

$$\mathbb{E} \left[\log(N^{(s)}(t_b)) \mid N^{(s)}(t_{b-1}) = n, N^{(s)}(0) > 0 \right] \tag{A.3.1.18}$$

requires being able to observe the value of $N^{(s)}$ at more than one time point for a single droplet, which is impossible for the data relevant to this study. As a reminder, this data is not actually longitudinal. Thus it seems one is limited to using (A.3.1.12), or something conceptually similar, when attempting to modify the approach from previous work [VCF$^+$18] [FM14] [MBH$^+$14] [SBT$^+$13] [MMV$^+$08] to be applicable to this data.

A.3.2 Comments

The only reason for including the parameters r_s in the model (2.2.15) used in this work is simply to ensure that "baseline" or "constant" growth effects are not misattributed as being part of the effects of the microbial interactions. Indeed, unlike the papers [VCF$^+$18],

[MBH$^+$14], and [MMV$^+$08], the previous work [FM14] did not include (analogues of) the parameters r_s in their model, so there would be precedent for omitting them.

The assumption that the temporal dynamics can be adequately described by the equations (2.2.15) also implies at least two important assumptions:

1. Because the equations (2.2.15) allow for indefinite population growth, using them implicitly assumes that all of our observations correspond only to the regime before resource scarcity and other factors begin to constrain the growth of microbial cells. Cf. the discussion from section 1.4.2.

2. The equations (2.2.15) implicitly assume that only pairwise interactions between different microbial species are relevant for determining their rates of growth, and that there is no non-negligible effect from higher-order interactions involving three or more distinct species. This is reflected in the absence of any terms that are a function of the abundances of more than two species. This has been recognized as a major limitation of these equations for accurately describing microbial interactions in previous work [MXS17]. Cf. also [BKK16] or the discussion from section 1.2.

The second assumption was also reflected in the experimental design used in [VCF$^+$18], which nevertheless found the equations (2.2.15) adequate for describing the microbial interactions which they observed. In previous work the first assumption has been enforced via constraints on β_s and $\alpha_{s_1 s_2}$ which are not applied here, since such constraints seem to lack a biological justification, and they are unnecessary for the specific simulation model (A.2.1.3) that I used.

This quote from [MBH$^+$14] is precedent for the above viewpoint on parameter constraints:

> One approach to improving parameter estimation is to superimpose constraints on the parameters; however, unless these constraints can be justified both biologically and mathematically, they should not be enforced because the results can be greatly affected.

Additional References for Appendices

[Ara03] Matyas Arato, *A famous nonlinear stochastic equation (lotka-volterra model with diffusion)*, Mathematical and Computer Modelling **38** (2003), 709–726.

[LE01] Pingzhou Liu and Saber N. Elaydi, *Discrete Competitive and Cooperative Models of LotkaVolterra Type*, Journal of Computational Analysis and Applications **3** (2001), no. 1, 53–73.

[Ush82] Shigehiro Ushiki, *Central difference scheme and chaos*, Physica D: Nonlinear Phenomena **4** (1982), no. 3, 407–424.