

# UC San Diego

## UC San Diego Previously Published Works

### Title

Spike Train Coactivity Encodes Learned Natural Stimulus Invariances in Songbird Auditory Cortex.

### Permalink

<https://escholarship.org/uc/item/55h9c3km>

### Journal

Journal of Neuroscience, 41(1)

### ISSN

0270-6474

### Authors

Theilman, Brad  
Perks, Krista  
Gentner, Timothy Q

### Publication Date

2021-01-06

### DOI

10.1523/jneurosci.0248-20.2020

Peer reviewed

# Spike Train Coactivity Encodes Learned Natural Stimulus Invariances in Songbird Auditory Cortex

Brad Theilman,<sup>1</sup> Krista Perks,<sup>1</sup> and Timothy Q. Gentner<sup>1,2,3,4</sup>

<sup>1</sup>Neurosciences Graduate Program, University of California San Diego, La Jolla, California 92093, <sup>2</sup>Department of Psychology, University of California San Diego, La Jolla, California 92093, <sup>3</sup>Neurobiology Section, Division of Biological Sciences, University of California San Diego, La Jolla, California 92093, and <sup>4</sup>Kavli Institute for Brain and Mind, La Jolla, California 92093

The capacity for sensory systems to encode relevant information that is invariant to many stimulus changes is central to normal, real-world, cognitive function. This invariance is thought to be reflected in the complex spatiotemporal activity patterns of neural populations, but our understanding of population-level representational invariance remains coarse. Applied topology is a promising tool to discover invariant structure in large datasets. Here, we use topological techniques to characterize and compare the spatiotemporal pattern of coactive spiking within populations of simultaneously recorded neurons in the secondary auditory region caudal medial neostriatum of European starlings (*Sturnus vulgaris*). We show that the pattern of population spike train coactivity carries stimulus-specific structure that is not reducible to that of individual neurons. We then introduce a topology-based similarity measure for population coactivity that is sensitive to invariant stimulus structure and show that this measure captures invariant neural representations tied to the learned relationships between natural vocalizations. This demonstrates one mechanism whereby emergent stimulus properties can be encoded in population activity, and shows the potential of applied topology for understanding invariant representations in neural populations.

**Key words:** auditory; invariance; population; topology

## Significance Statement

Information in neural populations is carried by the temporal patterns of spikes. We applied novel mathematical tools from the field of algebraic topology to quantify the structure of these temporal patterns. We found that, in a secondary auditory region of a songbird, these patterns reflected invariant information about a learned stimulus relationship. These results demonstrate that topology provides a novel approach for characterizing neural responses that is sensitive to invariant relationships that are critical for the perception of natural stimuli.

## Introduction

Functional sensory systems must express a degree of representational invariance to convey relationships or abstract properties, such as category membership, between otherwise distinguishable stimuli. At the neural level, invariant representations may manifest as robustness to “noise,” or consistency across trials or learned stimuli in a single animal, or across animals. For high-dimensional natural signals in particular, whose representation is distributed, invariant representations likely involve the coordinated spiking activity of neuronal populations across time. The

nature of this coordination and how it supports invariance is an important open question in neuroscience.

Most current methods for quantifying neuronal population structure rely on converting spiking activity into a vector of instantaneous firing rates (Cunningham et al., 2009; Churchland et al., 2012; Gallego et al., 2018). Over time, these vectors trace out a path in a high-dimensional space. The dynamics of this path can be correlated with simultaneous information about the stimulus or behavior. However, these methods require averaging activity over time or space to overcome putative noise in single-neuron firing rates. This is problematic for populations that must represent intrinsically time-varying objects, such as vocal communication signals, on fast time scales. Auditory populations that do so often exhibit lifetime-sparse firing, sometimes complicating the estimation of instantaneous firing rates. Thus, while firing rate vectors may be decodable in some cases (Ince et al., 2013; Rigotti et al., 2013), they belie the near instantaneous and continuous nature of natural perception.

Alternatively, information may be carried in patterns of coincident spiking, or a “coactivation code” (Curto and Itskov, 2008;

Received Jan. 31, 2020; revised Oct. 30, 2020; accepted Oct. 31, 2020.

Author contributions: B.T., K.P., and T.Q.G. designed research; B.T. and K.P. performed research; B.T. analyzed data; B.T. wrote the first draft of the paper; B.T., K.P., and T.Q.G. edited the paper; B.T. and T.Q.G. wrote the paper.

This work was supported by National Institutes of Health DC0164081 and DC018055 to T.Q.G. We thank Charles Stevens for comments on an earlier draft.

The authors declare no competing financial interests.

Correspondence should be addressed to Timothy Q. Gentner at [tgentner@ucsd.edu](mailto:tgentner@ucsd.edu).

<https://doi.org/10.1523/JNEUROSCI.0248-20.2020>

Copyright © 2021 the authors

Brette, 2012). Curto and Itskov (2008) developed an algorithm for converting neural coactivity into a kind of topological space called a simplicial complex. The spatial structure of this complex is determined by the relative timing of spikes in a population. Their simulations showed that invariant topological measures of the spike-derived simplicial complex matched several measures of the space of stimuli driving the neural response. This powerful property allowed them to reconstruct some stimulus space properties solely from the population activity without having to correlate physical positions to population firing rates (i.e., compute receptive fields). To our knowledge, this topological approach has not been used to examine the invariant properties of the temporal coactivation patterns of *in vivo* sensory system responses.

Here, we apply the Curto-Itskov construction to extracellular responses from the caudal medial neostriatum (NCM), a secondary auditory cortical region, in European starlings, a species of songbird. Starlings adeptly learn complex acoustic signals and provide an important model for the neural basis of vocal learning and perception (Kiggins et al., 2012). The NCM, an auditory forebrain region analogous to secondary auditory cortex in mammals (Karten, 1967; Vates et al., 1996; Wang et al., 2010), is involved in processing complex behaviorally relevant acoustic signals (birdsongs) (Sen et al., 2001; Grace et al., 2003; Thompson and Gentner, 2010). NCM neurons respond selectively to subsets of conspecific songs (Bailey et al., 2002; De Groof et al., 2013), and single-neuron responses display complex receptive fields (Theunissen et al., 2000; Kozlov and Gentner, 2016). How the responses of single NCM neurons combine to represent invariant structure from natural vocal signals at the population level remains unknown. We hypothesized that invariant song representations in NCM are carried in the coactivation pattern of spikes in the population, and that these patterns manifest as a topologically invariant structure. Our results support this hypothesis. We show that the coactivity-based topology of NCM neural activity is nontrivial and carries stimulus-specific information. Using a novel mathematical approach to compare simplicial complexes, we find evidence that temporal coactivity patterns in NCM represent both stimulus identity and learned relationships between stimuli. We suggest that understanding the topology of sensory-driven neural population coactivity offers novel insight into the nature of how invariant representations are constructed from complex natural signals by neural populations.

## Materials and Methods

### Experimental model and subject details

All protocols were approved by the University of California San Diego Institutional Animal Care and Use Committee. This study did not generate new unique reagents. All of the custom code used for the topological analyses described here is available at <https://github.com/theilmbh/NeuralTDA>. Requests for resources, reagents, and further information is available from T.Q.G.

**Subjects.** Four birds were used in the study, identified as B1083, B1056, B1235, and B1075. All birds were wild-caught in southern California and housed communally in large flight aviaries before training and physiological testing. During behavioral training, birds were housed in isolation along with a custom operant apparatus in sound-attenuation chambers (Acoustic Systems) and maintained on a light schedule that followed local sunrise and sunset times. Birds had unrestricted access to water at all times. During behavioral training, birds received all food (chick chow) through the operant apparatus, contingent on task performance. All subjects were at least 1 year of age or older at the start of the experiment. We did not control for the sex of the subjects.

**Stimuli.** Starlings (both male and female) produce long, spectrotemporally rich, and individualized songs composed of repeating, shorter acoustic segments known as “motifs” that are learned over the bird’s lifespan. Motifs are on the order of 0.5–1.5 s in duration. In these experiments, a library of 16 6-s-long “pseudo-songs” were used as stimuli. Each pseudo-song was constructed from six 1-s-long motifs manually extracted from a larger library of natural starling song produced by several singers. The first motif of each pseudo-song was an introductory whistle motif, as commonly occurs in natural starling songs. The same whistle was used for all 16 pseudo-songs. After the introductory whistle, all motifs were unique to one pseudo-song (i.e., none of the other motifs occurred in more than one song). The amplitude of each pseudo-song was normalized to 65 dB SPL.

**Behavioral training.** We trained 4 birds on a two-alternative forced-choice task to recognize four naturalistic pseudo-song stimuli, following previously described operant conditioning procedures (Knudsen and Gentner, 2013). Briefly, the operant apparatus afforded access to three separate response ports. Subjects learned to peck at the center response port to trigger the playback of one of four pseudo-songs associated with two different responses: a peck to the left response port or a peck to the right response port. Pecking at the left response port following the presentation of two songs, or at the right response port following the other two songs, allowed for brief (1–2 s) access to a food hopper. Incorrect responses, pecking the response port not associated with a given song, resulted in a short timeout during which the house light was extinguished and the bird could not feed from the hopper. The 12 remaining pseudo-songs were never presented during operant training for that bird. Assignment of songs for behavioral training was counterbalanced across subjects. The accuracy of the birds over the last 500 trials before neural recording was as follows: B1083, 86.2%; B1056, 95.4%; B1235, 97.6%; B1075, 91.2%. All birds exceeded 90% accuracy during training. Acquisition of the behavior was fast; the number of 500 trial blocks required to first exceed 90% accuracy was as follows: B1083, 30; B1056, 14; B1235, 5; B1075, 13.

**Electrophysiology.** Once trained, we prepared the birds for physiological extracellular recording from caudo-medial nidopallium (NCM). We anesthetized birds (20% urethane, 7 ml/kg, i.m.), affixed a metal pin to their skull with dental cement, removed an ~1 mm<sup>2</sup> window from the top layer of skull, and placed a small craniotomy dorsal to NCM. We inserted a 16- or 32-channel silicon microelectrode (NeuroNexus Technologies) through the craniotomy into the NCM of the head-fixed bird positioned on a foam couch within a sound attenuation chamber (Acoustic Systems). Auditory stimuli were presented from a speaker mounted ~20 cm above the center of the bird’s head, at a mean level of 60 dB SPL measured at 20 cm. The microelectrode was lowered until all recording sites were within NCM, and then allowed to sit for ~15 min to allow any tissue compression that might have occurred during insertion to relax and stabilize before stimulus presentations began. Auditory activity was confirmed by exposing the birds to non-task-related sounds and observing sound-evoked responses in the raw voltage signal on one or more channels. The 16 pseudo-song stimuli in the block were presented pseudo-randomly until a total of 20 repetitions for each stimulus. For B1083, after the completion of the first block, the electrode was driven deeper into NCM by a distance greater than the electrode length, to obtain another block from a disjoint population. Raw voltages recorded from each microelectrode site were buffered and amplified (20× gain) through a headstage (TBSI), bandpass filtered between 300 Hz and 5 kHz, and amplified (AM Systems, model 3600) and sampled digitally at 20 kHz (16 bit; CED model 1401) via Spike2 software (CED). We saved the full waveforms for offline analysis.

### Experimental design and statistical analysis

**Spike sorting.** Extracellular waveforms were spike-sorted offline using the Phy spike sorting framework (Rossant et al., 2016). After automatic clustering, the sort was completed manually using the KwikGUI interface. Clusters with large signal-to-noise ratio and <1% refractory period violations (taken to be 1 ms) were labeled as “Good” clusters and included in the analysis. Because the analyses are meant to quantify population-level structure, sorting priority was given to extracting as many

neural signals as possible. Clusters identified automatically were mostly kept separate unless obvious duplicates were observed. Duplicate clusters were defined by overlapping distributions in PCA space, similar waveform shape, and cross-correlations that respected refractory periods. Despite these measures, the set of clusters labeled “Good” are likely to contain multiunit activity as well as single units.

**Data processing.** Data processing began by converting the neural population activity on each trial into an  $n \times m$  matrix comprising all the spikes across all the isolated clusters and binned into 10 ms windows with a 5 ms overlap between windows. We used the Perseus persistent homology software for computing Betti numbers from simplicial complexes (Mischaikow and Nanda, 2013). For computing samples from the simplicial configuration model, we used the code described by Young et al. (2017). All other computations were performed using custom-written Python and C code available at <http://github.com/theilmbh/NeuralTDA>.

**Mathematical background.** Outside of neuroscience, invariance has been the subject of intense mathematical investigation. In particular, the mathematical field of topology is dedicated to studying the properties of arbitrary spaces that are invariant to different classes of transformations. This research has produced a wealth of techniques and methods for defining and quantifying invariant structures. Recently, these techniques have been increasingly applied to the problem of finding invariant structure in large datasets. Topology takes a global viewpoint, and the “quantities” of interest are entire spaces, rather than single numerical measures. By converting neural population activity into a topological space instead of reducing it to a series of numerical measures, we gain access to the library of techniques topology offers for characterizing invariant structure.

Here, we introduce the basics of the mathematics used for our analysis. More thorough introductions are available elsewhere (Hatcher, 2002), including for the many applications of algebraic topology (Ghrist, 2014). Our analyses work by constructing mathematical objects known as simplicial complexes. Simplicial complexes are topological spaces that are built from discrete building blocks known as simplexes. For each dimension, there is only one prototype simplex. For example, 0-dimensional simplexes are points, 1-dimensional simplexes are line segments, and 2-dimensional simplexes are triangles. Simplexes exist in every dimension, including dimensions  $>3$  where visualization is not possible. For every simplex of dimension  $d$ , there is a collection of  $d-1$  simplexes called faces. Consider the tetrahedron, which is the prototype 3-dimensional simplex (or 3-simplex). The faces of the tetrahedron are the four triangles, which themselves are 2-simplexes. The “faces” of the triangles are the 6 edges, which are 1-simplexes, and so on.

Simplicial complexes are built by “joining” simplexes along common faces. That information is carried through the boundary operators, which are linear maps encoding which  $d$ -simplexes are attached to which ( $d-1$ ) simplexes. Given a simplicial complex, we construct a series of vector spaces, one for each dimension of simplex in the complex. These vector spaces are called “chain groups.” The basis vectors of these vector spaces are taken to be, formally, the simplexes in the various dimensions. For example, a graph has a chain group in dimension 0 consisting of all formal linear combinations of vertices, and a chain group in dimension 1 consisting of all formal linear combinations of edges. The boundary maps are linear maps (matrices) between chain groups of adjacent dimensions, mapping the dimension  $d$  chain group to the  $d-1$  chain group. These matrices contain all the structure of the simplicial complex, which allows for all the computations described in this work to be performed using the standard, highly optimized numerical linear algebra routines available in scientific programming packages, such as SciPy.

**Constructing simplicial complexes for neural data.** To construct the simplicial complex associated to a population spike train, we follow methods similar to those developed by Curto and Itskov (2008). For each presentation of a stimulus, the neural response was divided into time bins of width  $dt$  and a percentage overlap between adjacent bins, both free parameters for the analysis. Unless otherwise noted, for all analyses,  $dt$  was 10 ms with 50% overlap, to preserve continuity of the spike trains. Each spike in the time period was assigned to its associated time bins. The result is an  $N_{\text{cells}} \times N_{\text{bins}} \times N_{\text{trials}}$  multidimensional array  $D$ , representing the population activity, where each element

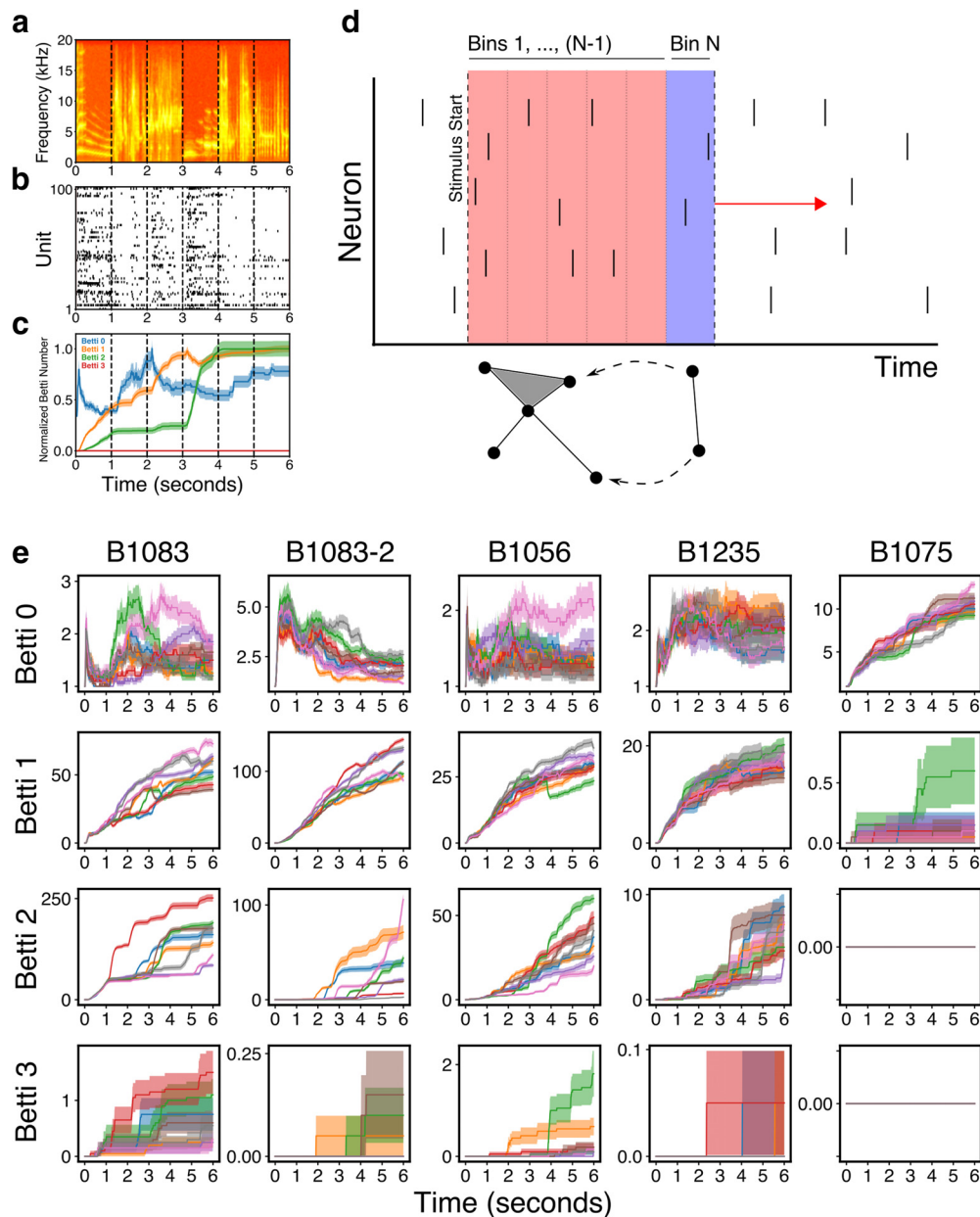
is a spike count in the time window  $dt$ . The spike counts were then divided by the window width to determine a “firing rate” for that cell in that time bin. The time average was taken to yield an  $N_{\text{cells}} \times N_{\text{trials}}$  matrix of average instantaneous firing rates for each cell for each trial. Then, the multidimensional array  $D$  was thresholded by determining which bins had a firing rate greater than some threshold value times the average firing rate. This yielded an  $N_{\text{cells}} \times N_{\text{bins}} \times N_{\text{trials}}$  binary array  $B$  that represents significantly active cells. Unless otherwise noted, the threshold value of 4 times the average firing rate of the unit was used. Since the average firing rate of the units was low, the threshold value did not make a significant impact on the results.

For each trial and for each time bin in  $B$ , the  $N_{\text{cell}}$ -long binary vector  $V$  was used to create a list of cells that were active in that time bin. This was repeated across all time bins to create a list of “cell groups” active during the trial, much like the cell groups defined by Curto and Itskov (2008). This list of cell groups was taken to define the maximal simplexes for the simplicial complex. Using algorithms derived from published work (Kaczynski et al., 2006), this list of cell groups was turned into a list of lists of basis vectors for the chain complex. The  $n$ -th sublist represents the chain complex basis vectors corresponding to the  $n$ -dimensional simplexes in the simplicial complex. We also computed the matrices that represent the boundary operators in the chain complex.

**Betti curves.** Homology is one of the most basic topological invariants one can use to describe a topological space. Loosely speaking, homology counts the number of  $n$ -dimensional holes in the space. These holes are described purely algebraically by computing how  $n$ -dimensional cycles (e.g., a circle) are not the boundary of an  $n+1$ -dimensional simplex. For each dimension  $n$ , the number of such holes is called the  $n$ -th Betti number. The reason that Betti numbers are a useful characterization of the topological spaces we construct is that any two topological spaces that differ in their Betti numbers cannot be topologically equivalent. Furthermore, topological holes interpreted neurally represent “gaps” in the coactivation pattern (which neurons do not fire together) and thus carry information about how the population activity is structured. We summarized the topology of neurally derived simplicial complexes by computing these Betti numbers across time to yield Betti curves.

Figure 1*d* illustrates the temporal filtration. From the start of the stimulus, cell groups from each time bin are added to the growing simplicial complex. The Betti numbers for the entire complex are computed to yield the value of the Betti curve for that bin. Then, the next bin is added, and the Betti numbers are recomputed. Simplicial complexes were fed into the Perseus persistent homology program (Mischaikow and Nanda, 2013; Nanda, 2013), which computed the sequence of Betti numbers in the evolving simplicial complex. In the language of applied topology, we computed the homology of the filtered simplicial complex using time as the filtration parameter (Giusti et al., 2016). Betti curves were computed for each dimension and for each trial of a given stimulus. Because the Betti numbers are discrete, we interpolated the output from Perseus using step functions. This allowed us to average the Betti curves across trials. We varied the time bin width  $dt$  between 5 and 250 ms and observed that the shape of the Betti curves was consistent for values of  $dt$  between 5 and 50 ms. For values of  $dt > 50$  ms, the Betti curves looked like low-pass-filtered versions of the small- $dt$  curves. Varying the percentage overlap from 0% to 50% did not appear to significantly alter the Betti curves.

The Betti curves reflect the accumulated effects of moment-to-moment spatiotemporal structure in the population activity. To serve as a control, we used various shuffling procedures to break this structure. We computed “fully shuffled” Betti curves by taking the neural response matrix for each trial and shuffling each row independently across columns. Since each row corresponds to a cell’s response, this approach shuffles each cell’s response independently across trials and across cells. The effect is to break the temporal correlations between cells while preserving the total number of spikes from a given cell in a given trial. We computed trial-shuffled Betti curves by using a similar shuffling procedure; but instead of shuffling a cell’s response across time independently for each trial and each cell, we shuffled each response across trials independently for each cell and each time bin.



**Figure 1.** NCM population activity produces nontrivial topological features. **a**, Spectrogram from a typical 6 s stimulus with motif boundaries marked with black dotted lines. **b**, Spike raster from 100 units from region NCM during one trial of the presentation of the stimulus in **a**. **c**, Normalized average Betti curves from the population in **b** for the stimulus in **a**. The values of the Betti numbers are normalized to their maximum during the stimulus presentation, to facilitate visualization of all dimensions on the same plot. The Betti curves are averaged over 20 presentations of the same stimulus. **d**, Schematic of the temporal filtration. Beginning with the start of the stimulus, the elementary simplex from each time bin is added to the growing simplicial complex. At each time bin, the Betti numbers for the entire complex up to that point in time are computed. Then, the next bin is added, and the values of the Betti numbers are recomputed. **e**, Betti curves for Betti numbers 0–3 for all stimuli for all populations. Each color represents a different stimulus, and identical colors across panels represent the same stimulus. Shaded regions represent SEM. Individual stimuli produce unique Betti curves. Betti curves for different stimuli overlap during the first 1 s of the stimulus, consistent with the first 1 s of all stimuli being identical.

Spatiotemporal structure in population responses to stimuli reflects two kinds of correlation. The first is correlation between individual cell spike trains and individual stimuli. We call this the stimulus specificity of single-unit responses. The other kind of correlation is the set of  $n$ -wise correlations between neurons on a single trial across time (sometimes referred to as the noise correlation). The fully shuffled condition destroys both these forms of correlation, by assigning different single-unit spike trains, and therefore a random  $n$ -wise correlation structure, to each neuron on each presentation of each stimulus. To dissociate these two correlation sources, we created an additional shuffle of the population data, termed the “within-stimulus mask shuffle,” in which each stimulus is associated with a unique time bin-to-time bin permutation mask. We then shuffle the spiking response of each neuron

individually and independently within a single trial for each stimulus, applying the same bin-to-bin permutation to all trials of that same stimulus. This creates a new, independent, stimulus-specific response for each neuron, and replaces the empirical  $n$ -wise population correlations with a random pattern of coactivity that repeats across repetitions of each stimulus.

**Stimulus specificity.** We analyzed the stimulus specificity of the empirical and shuffled Betti curves by first computing the distribution of the correlations between the Betti curves of pairs of trials. Then we compared the distributions of correlations from a pair of trials within a single stimulus to correlations from pairs of trials between different stimuli. The significance of these differences was computed using a linear mixed-effects model (Lindstrom and Bates, 1988). The model predicted the

correlation between Betti curves on pairs of trials treating the within-stimulus versus between-stimulus distinction as a random effect and treating the average firing rate and an interaction of firing rate and pair type as fixed effects. We included firing rate effects since the shuffles preserve average firing rate but destroy precise temporal correlations.

**Simplicial configuration model.** To test whether the topology of the neurally derived simplicial complexes is likely to be seen in a random simplicial complex, we used the simplicial configuration model (Young et al., 2017). Because of the computational constraints of the configuration model, for each stimulus, we averaged the neural activity over trials to yield an average population spike train. The simplicial complex for this spike train was computed and seeded the simplicial configuration model software (Young et al., 2017). This algorithm produces a Markov chain that yields samples from the uniform distribution over all simplicial complexes with local connectivity structure identical to the seed complex derived from neural activity. For each sample, the Betti numbers were computed using Perseus. The distribution of these Betti numbers serves as a null distribution against which Betti numbers derived from the empirical data can be compared. We interpret data having Betti numbers that fall well outside the null distribution as containing nonrandom topological structure. To quantify how far outside the null distribution an observed Betti number falls, we calculated empirical  $p$  values using the formula as follows (Davison and Hinkley, 1997):

$$p = \frac{1 + \sum (b_i > b_e)}{N + 1} \quad (1)$$

Where  $b_i$  represents the Betti number for the  $i$ -th sample from the simplicial configuration model,  $N$  is the number of samples from the simplicial configuration model, and  $b_e$  is the empirically observed Betti number.

**Simplicial Laplacian.** We generalized the methods of De Domenico and Biamonte (2016) to define information-theoretic quantities related to simplicial complexes. Equipped with the boundary operator matrices  $\partial_i$ , the simplicial Laplacians (Horak and Jost, 2013) were computed as follows:

$$L_i = \partial_i^* \partial_i + \partial_{i+1} \partial_{i+1}^* \quad (2)$$

where  $*$  indicates the adjoint, which in these analyses means the matrix transpose. Given a simplicial Laplacian  $L_n$ , the related density matrix was computed as follows:

$$\rho_n = \frac{e^{-\beta L_n}}{\text{Tr} e^{-\beta L_n}} \quad (3)$$

where  $\beta$  is a free parameter. Given a density matrix expressed in its eigen-basis, the eigenvalues represent the probability distribution over eigenstates with maximum entropy under the Hamiltonian given by the Laplacian. We can define the Kullback-Leibler (KL) divergence between any two such density matrices,  $\rho$  and  $\sigma$ , by first diagonalizing both  $\rho$  and  $\sigma$  and then sorting each vector of eigenvalues by magnitude. The KL divergence is then defined by the following:

$$D_{KL}(\rho, \sigma) = \sum_i \rho_i (\log \rho_i - \log \sigma_i) \quad (4)$$

Where  $\rho_i$  is the  $i$ -th eigenvalue. The Jensen-Shannon (JS) divergence was defined using the following:

$$M = \frac{\rho + \sigma}{2} \quad (5)$$

$$D_{JS}(\rho, \sigma) = \frac{1}{2} (D_{KL}(\rho, M) + D_{KL}(\sigma, M)) \quad (6)$$

The advantage of sorting the eigenvalues is that the divergence becomes invariant to the labels assigned to individual neurons. This is

desirable because these measures should depend only the neural activity and not the specific labels assigned to neurons, analogous to how different maps of the world describe the same geography with different coordinates. This definition of JS and KL divergence ignores the eigenvectors of  $\rho$  and  $\sigma$ . Indeed,  $\rho$  and  $\sigma$  may not be simultaneously diagonalizable (or even of the same dimension). Rectifying these issues requires a choice in the definition of the metrics, which like any other metric is justified by its ultimate use. We explain our choices in the following paragraph. Regardless, the eigenspectra of  $\rho$  and  $\sigma$  do form discrete probability distributions, and the definitions of KL and JS divergence here agree with the usual definitions of these divergences on discrete probability distributions.

One limitation in the naive generalization of the foregoing metrics to neural data are that they rely on the density matrices being square and of the same dimension. For real data, however, different stimuli or repetitions of the same stimulus often evoke different numbers of active neurons, giving rise to chain groups, boundary operators, and density matrices of different dimension. To address this, given two Laplacians of different dimensions, we expanded the dimension of the smaller Laplacian by padding with zeros. An alternative approach is to collapse the two simplicial complexes together along their common simplexes and derive “masked” boundary operators that only operate on the simplexes within the original, individual complexes, and then define the Laplacians using these masked boundary operators. Both the “zero-padding” and the “masking” approaches yielded similar results for our datasets. We report results from the computationally simpler method of zero-padding. In all analyses, the simplicial Laplacian was computed from the final simplicial complex constructed from the population response to an entire trial.

**Beta parameter.** The free parameter,  $\beta$ , that appears in the expression for the density matrix originates in statistical physics, where it is interpreted as “inverse temperature,” i.e.,  $\beta$  is proportional to  $1/\text{temperature}$ . The significance of  $\beta$  in the context of network theory is actively under study (Nicolini et al., 2018). Here we interpret  $\beta$  by noting that it acts to normalize the Laplacian matrix, and by extension, the eigenvalues of the Laplacian. In this sense,  $\beta$  acts as a “scale parameter” that sets the scale of the spectral features of interest. The eigenvalues of the Laplacian are always non-negative; and if  $\beta$  is large, then large eigenvalues are significantly damped by the negative exponential. Heuristically, increasing  $\beta$  decreases the proportion of larger eigenvalues that contribute to the quantities of interest. In our case, if  $\beta$  is too large, too little of the spectrum will be relevant for the JS or KL divergences, and we reasoned the results would be less interpretable. To confirm this exponential damping, we computed the proportion of eigenvalues above a threshold of  $1e-14$  as  $\beta$  varied from 0.1–100. Generally, this proportion started decreasing from 100% with  $\beta \cong 1$ , to below 20% when  $\beta > 10$ . We set  $\beta = 1$  in all of our analyses, but the main conclusions do not change as long as  $\beta$  was below the upper bound of 10.

**Simulated spiking populations.** We used synthetic spike trains to examine how well the proposed metrics reveal invariant properties between neural populations. For “target” spike trains, we simulated a population of 20 Poisson spiking neurons for 1000 time-steps with a rate parameter selected from the set (0.01, 0.02, 0.03, 0.04). We generated “test” spike trains by choosing 50 rate parameters evenly in the range of [0.001, 0.1] and simulating 25 population spike trains for each rate. We computed the simplicial KL divergence between each test spike train and the target spike train, then computed the average divergence between the target and the model at the given rate parameter.

We applied the same procedure to simulate heterogeneous Poisson populations, except that half of the neurons were simulated with rate parameter 0.02 and the other with 0.05. The assignment of neurons to rate-parameter subgroups was random. The above-described parameter sweep procedure was repeated over the two population rate parameters.

**Simulated environments.** We simulated physical environments similarly to the environments used in Curto and Itskov’s original work (Curto and Itskov, 2008). We randomly placed 0–4 holes of radius 0.3 m inside a  $2 \times 2$  m arena, ensuring hole centers did not overlap. We then randomly distributed 100 place fields of radius 0.2 m in the environment using an algorithm that ensured that the environment was fully tiled. To

model movement, we constructed a random walk trajectory simulating a 10 min traverse of the environment sampled at 10 samples per second to ensure the whole arena was explored. We constructed spike trains by simulating Poisson spike trains at a constant rate for each neuron while the random walk trajectory was within the cell's place field, and silent outside of it.

We constructed five different environments with each specified number of holes. For each environment, 10 random walks were simulated. The pairwise 1-JS divergence between pairs of individual trials was computed by taking the simulated spike trains' 1-Laplacians and computing the spectral JS divergence as defined above. We averaged the 1-JS divergences over all walks for each environment to yield the plot in Figure 5a. Then, the JS divergences for each of the 25 environment pairs with given numbers of holes were averaged to yield the plot in Figure 5b.

**In vivo JS divergences.** Data from trained, anesthetized birds was recorded and preprocessed to yield population spike trains as described earlier. The JS divergence analyses were restricted to the four learned stimuli and four other unfamiliar stimuli. For an arbitrary pair of responses from two separate stimulus presentations, the 1-Laplacian of each response was computed by first converting the neural data into a simplicial complex as described above, using a bin size of 10 ms with 5 ms overlaps. Then, the JS divergences between the 1-Laplacians were computed as described above. It is important to note that the JS divergences were computed using the final simplicial complex from each trial. As a result, it may "miss" homological features that do not persist through the whole trial. We do not consider this a limitation in the present study, since the final simplicial complex reflects the aggregate effects of temporal coactivations through the entire trial, and thus tells us something about the population response to the entire stimulus considered as a unit. Nevertheless, it will be valuable for future work to explore the properties of these features that may be missed by the current analysis.

**Correlation-based population response similarity.** As an alternative to the simplicial Laplacian spectral entropy (SLSE) measure of population similarity, we also compute a correlation-based measure to quantify the trial-to-trial similarity in the response of a population,  $C(A,B)$ , where  $A$  and  $B$  are the responses of the same population on different trials responding to either the same stimulus or to different stimuli. We first smooth the population spike trains with a Gaussian kernel with a width of 10 ms, then compute the correlation coefficients between the spike train of a neuron in response  $A$  and the spike train from the same neuron in response  $B$ , then average across all the correlation coefficients to yield a raw scalar estimate of similarity,  $C$ , between the two populations. We use  $1 - C$  in place of the raw correlation measure so that smaller values of both correlation and JS divergence indicated more similar spike trains, such that  $1 - C = 0$  when the response of every single neuron in the population for population response  $A$  is perfectly correlated with its corresponding response for population response  $B$ . Comparing this similarity measure with the SLSE measures allows us to test whether physically dissimilar responses (in terms of the actual spike trains) may evince similar topologies.

We also compute the distance between pairs of population responses using covariance matrices. For each trial, we calculated the pairwise covariance matrix between Gaussian-smoothed spike trains, again with a 10 ms window. This yields an  $N_{\text{cell}} \times N_{\text{cell}}$  matrix for each trial in which each  $i,j$  entry is the temporal covariance between the smoothed spike trains from neurons  $i$  and  $j$  on that trial. We obtained a distance between pairs of trials by computing the Frobenius norm of the difference between each trial's covariance matrix. The distance matrix was normalized by its maximum value over all pairs of trials to yield a normalized distance measure. This measure places two trials close to each other if the second-order (pairwise) correlation structure in the population for response  $A$  is similar to the second-order correlation structure for response  $B$ .

The correlation between population responses with relabeled neurons was obtained by shuffling the rows of the Gaussian smoothed population response matrix and repeating the population correlation computation described above on the shuffled matrices.

JS divergence and correlation measures were compared using the relative accuracy of logistic regressions. We fit logistic regressions to predict

whether a pair of responses to different stimuli came from stimuli that belonged to either the same or different behaviorally trained class based on the value of either the JS divergence or the specific correlation measure of interest between the responses. Regressions were fit to a random subset of 80% of the data and tested on the remaining 20%. A total of 240 separate regressions were performed for each condition.

We measured the similarity of topological structure across disjoint subpopulations by randomly splitting the full dataset from B1083 Population 1 into two disjoint subgroups each with 50 neurons. Then we conducted an identical analysis to the logistic regression described in the preceding paragraph but using the JS divergence between responses from the two disjoint populations. We repeated this 40 times with a different (randomly chosen) split of the original population. We assessed the difference between the means of the accuracies of the regressions fit to the JS divergence or correlations with a Gaussian GLM from the statsmodels python library with the type of measure (JS divergence/correlation) and the type of shuffle (shuffle/no shuffle) as regressors, and included an interaction term between the shuffle and measure.

**Linear mixed-effects model.** We used a linear mixed-effects model to examine the stimulus specificity of the empirically derived Betti curves associated with each stimulus, and those derived from the various shuffled versions of the spiking responses. The regression models predicted the correlation between Betti curves in a given dimension between pairs of trials, as a function of whether the pair came from two trials of a single stimulus or two trials of different stimuli. The model included population identity as a random effect. The model was implemented using MixedLM from the statsmodels python library. We included firing rate (and its interaction with stimuli) as a separate fixed effect in the model and exclude this from all reported effects on Betti curve correlations tied to stimulus specificity. Overall, the mean firing rate on a pair of trials tends to increase the strength of the correlation between Betti curves on those trials, in part because it tends to smooth over small moment-to-moment variations in a manner equivalent to computing simplicial complexes from spike trains quantized into larger time bins.

## Results

### Stimulus-specific topological structure in auditory neural populations

We first asked whether auditory stimulus-driven NCM population activity can be well described by its spike-based topological structure. We recorded song-evoked responses from five, distinct, large populations of neurons in the NCM of adult starlings (323 neurons total, distributed across 4 birds; B1083: 101, 95 neurons in two separate populations; B1056: 54; B1235: 40; B1075: 33). For each presentation of each song stimulus within each population, we rerepresented the full population spike train as a simplicial complex, then computed the Betti curves (see Materials and Methods) associated with each song stimulus. Each Betti curve describes structure in the population coactivity pattern across time. More specifically, each curve captures the evolution of the homology of the simplicial complex in a single dimension, as defined by the time-resolved spiking pattern across each neural population for a given stimulus (Fig. 1). Visual inspection of the trial-averaged Betti curves (Fig. 1e) suggests consistent stimulus-specific temporal dynamics in the homology of Dimensions 0, 1, and 2. Although the simplicial complexes sometimes display 3-dimensional homology (Fig. 1e, bottom row), these are not consistent across birds or stimuli. We note that the trajectories of the Betti curves in Dimensions 0–2 tended to overlap during the first motif for each stimulus, consistent with their stimulus specificity, as the first motif was the same for all songs.

Betti numbers provide one measure of topological properties present in the underlying neural population coactivity, and so can be interpreted as abstract proxies for structure. They also

provide insight into more concrete aspects of the underlying population activity. Betti 0 counts the number of connected components in the simplicial complex. Thus, for neural data, increasing Betti 0 suggests multiple independent subsets of neurons firing coincidentally within a group but not across groups. Higher dimensional Betti numbers, 1 to  $N$ , count the number of  $n$ -dimensional “holes” in the simplicial complex. For neural population coactivity patterns, the Betti  $n$  count corresponds loosely to the number of gaps (i.e., periods of simultaneous inactivity among neurons, or “missing” patterns of coactivity). As these gaps fill in, because of coincident activity, the corresponding Betti numbers may decrease.

To quantify stimulus specificity, we computed all the correlations between Betti curves from pairs of trials with the same stimulus and pairs of trials from different stimuli. We defined the population as being stimulus-specific if the distribution of within-stimulus correlations is significantly larger (i.e., closer to 1, more similar) than the distribution of between-stimulus correlations. By this measure, all Betti curves show significant stimulus specificity (linear mixed-effects model,  $Z \leq -4.083$   $p \leq 4.45e-5$  for all dimensions), with the coefficient significantly less than zero (Table 1), indicating that between-stimulus correlations are significantly lower than within-stimulus correlations. That is, the topology of auditory stimulus-driven population activity in NCM is stimulus-specific.

To further characterize the stimulus-specific dynamics of the Betti curves and examine the source of the stimulus specificity in the empirical population response, we compared the original Betti curves with those obtained after shuffling the spiking response of each neuron in time individually and independently, within each trial (see Materials and Methods). This shuffle, which we call the full-shuffle, preserves the total number of spikes and spike rate per trial, but destroys all the original spike-time coincidences between neurons in the population. Because the topological structure of the population depends on the spike-time coincidences, any stimulus specificity tied explicitly to the spike-time based topology (rather than the overall spike rate) should be abolished in the full-shuffled data. This in turn should yield Betti curves that are more similar across different stimuli and birds compared with the original Betti curves. Figure 2 shows the Betti curves that result from the full-shuffle (in orange) for a single stimulus from each of the populations. In each dimension, the fully shuffled curves show a degeneration to a stereotypical trajectory across populations. The same pattern is observed for all other stimuli. We quantified this explicitly by again comparing correlations between Betti curves from pairs of trials within and between stimuli. For Betti 0 and 1, the shuffled curves were not stimulus-specific (linear mixed-effects model,  $p > 0.202$ ). For Betti 2, some marginal stimulus specificity remained (linear mixed-effects model,  $Z = -2.993$ ,  $p = 0.003$ ), but the magnitude of the specificity was significantly smaller than that for the empirical data (linear mixed-effects model,  $Z = -21.24$ ,  $p < 1e-13$ ). Based on these results, we conclude that the topological structure of NCM population activity reveals stimulus-specific coincident patterns of neuronal firing.

How is it that the temporal coactivity pattern of the population response, as measured by the topology, carries stimulus-specific information? One possibility is that the population simply inherits specificity from single-unit responses. Individual NCM neurons have complex receptive fields (Kozlov and Gentner, 2016), and their responses are likely to be stimulus-specific (Thompson and Gentner, 2010). The collective responses of

**Table 1. Stimulus specificity of empirical and shuffled Betti curves<sup>a</sup>**

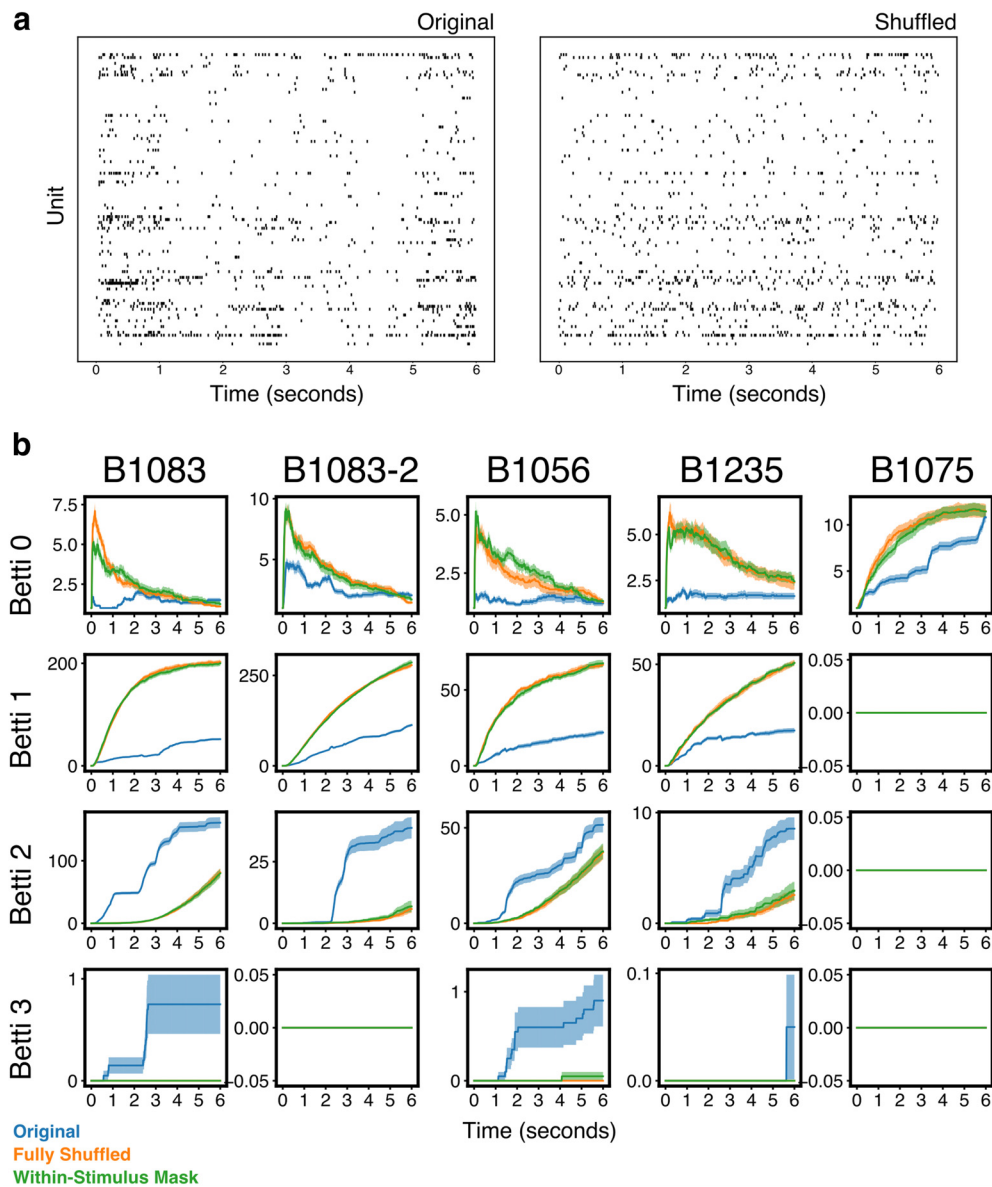
Dataset	Betti number	Coefficient	$z$	$p$	97.5% CI
Empirical	0	−0.031	−4.083	4.45e-05	[−0.046, −0.016]
Shuffled	0	−0.010	−1.275	0.202	[−0.024, 0.005]
Within-mask	0	−0.003	−0.33	0.741	[−0.017, 0.012]
Empirical	1	−0.032	−5.571	2.53e-08	[−0.043, −0.021]
Shuffled	1	0.000	0.303	0.762	[−0.003, 0.004]
Within-mask	1	0.002	1.432	0.152	[−0.001, 0.005]
Empirical	2	−0.149	−15.603	6.96e-55	[−0.167, −0.130]
Shuffled	2	−0.030	−2.993	0.003	[−0.050, −0.010]
Within-mask	2	−0.008	−0.857	0.391	[−0.027, 0.011]

<sup>a</sup>A linear mixed-effects model was used to compute the dependence of the Betti curve correlation on the type of trial pair (within stimulus or between stimuli). For all Betti numbers, the empirical curves show significant stimulus specificity, and this specificity is abolished by both fully shuffling the spikes and performing the within-stimulus mask shuffle (see Materials and Methods).

many such neurons are therefore also likely to be stimulus-specific, but the specific pattern of coactivations among individual neurons that defines the topology may or may not be determined by chance. That is, the topology may reflect random coactivity between independent spike trains driven by individual stimuli (i.e., the stimulus specificity of single-unit responses), and/or a unique nonrandom  $n$ -wise pattern of coactivity between neurons that is associated with a given stimulus, where  $n$  ranges from 2 to the number of recorded neurons. The full-shuffle destroys both these forms of coactivity by permuting single unit spike trains on each presentation of each stimulus, which in turn yields a unique  $n$ -wise coactivity structure on each trial. To further investigate these two possible sources of population coactivity structure, we created an additional shuffle of the original population data using a “shuffle mask” that describes how to permute the time bins for a given neuron on a single trial. We randomly generated a shuffle mask independently for each stimulus, and then applied the same mask to all trials from a single stimulus. This creates a new random correlation pattern between the spike trains of individual neurons and individual stimuli, and a random  $n$ -wise coactivity pattern between individual neurons, that are both preserved across separate trials from the same stimulus. If the observed stimulus-specific topologies emerge from a random alignment of spiking responses repeated across trials of the same stimulus, the mask-shuffled data should resemble an empirical response to a novel stimulus. If, however, the empirical  $n$ -wise correlation structure is privileged in some way, then the mask-shuffled data should resemble the full-shuffle responses.

We recomputed the Betti curves for the within-stimulus mask-shuffled data. Figure 2 shows the various shuffled Betti curves from an arbitrarily chosen stimulus for each population. Like the full-shuffle, the Betti curves for the mask-shuffled data do not appear to be stimulus-specific. To test this, we again measured stimulus specificity by computing the distributions of correlations between pairs of Betti curves from trials with either the same or different stimuli. The mask-shuffled Betti curves did not show significant stimulus specificity (linear mixed-effects model,  $p > 0.152$  for all Betti numbers). Table 1 gives the results of the linear mixed-effects model (see Materials and Methods) that assesses the significance of stimulus specificity in the Betti curves from the within-stimulus mask shuffle. Figure 3 presents histograms of the Betti curve correlations for within-stimulus and between-stimulus pairs of trials, for both the empirical and shuffled data, for population B1083. The lack of stimulus specificity in both the full- and mask-shuffled data supports the conclusion that stimulus-evoked coactivity in NCM carries stimulus-specific information that is not a trivial product of randomly aligned stimulus-specific single-neuron responses.

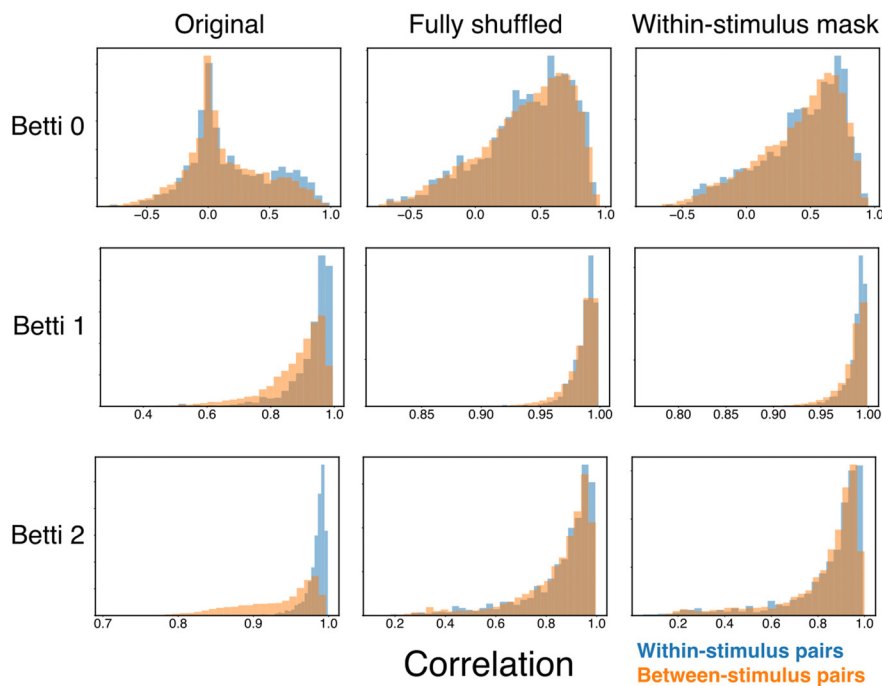




**Figure 2.** Shuffled Betti curves. *a*, Example empirical and shuffled population responses from an example trial. *b*, The average Betti curve for Dimensions 0–3 from an example stimulus is plotted for all populations. The empirically observed Betti curves (blue) are obtained from the raw neural data. Control Betti curves are computed from “fully shuffled” (orange) neural data in which the responses of individual cells are independently shuffled in time (see Materials and Methods), or the same neural data with the single-trial spike responses permuted by either a within-stimulus mask (green) or across trials (red). A different within-stimulus mask was generated for each stimulus, and the same mask was used for all trials of a given stimulus. The shuffled curves strongly overlap and are significantly different in shape and magnitude from the empirical Betti curves. Shaded regions represent SEM.

The differing shapes of the Betti curves for the empirical and both the full- and mask-shuffled spike trains suggest that the NCM population activity is structured nonrandomly at scales above the single neuron. To test this idea directly, we asked how likely it is that the observed topological structure in a given population might occur by chance. We constructed a null model for simplicial complexes that produces samples from the uniform distribution over all simplicial complexes that match the “local structure” of a given “seed” complex (Young et al., 2017), and used this model as the basis to compare the final Betti numbers for each stimulus (Fig. 4). For each bird, the majority of the stimuli showed responses with at least one Betti number significantly outside the null distribution (B1083: 8 of 8 stimuli; B1056: 8 of 8 stimuli; B1235: 8 of 8 stimuli; B1075: 7 of 8 stimuli). Thus, the stimulus-specific topology in NCM population spiking is not produced by chance spike coincidences between neurons.

As noted, the topological analyses we detail differ in fundamental ways from traditional correlation-based measures of population coactivity. Nonetheless, it is helpful to understand our results within the context of classically defined noise correlations (i.e., the stimulus-independent covariance in simultaneous firing rates between neurons). To do this, we again shuffled the empirical data, this time by permuting spikes across trials rather than time. The Betti curves for these trial-shuffled data show similar trajectories compared with those for the empirical data (Fig. 5), but with larger magnitudes. This suggests that the relative temporal dynamics of the coactivity pattern are governed largely by what would be referred to, classically, as the “signal correlation,” whereas the absolute number of “holes” in a given topological dimension is constrained by the “noise correlation.” The magnitude of Betti  $n$  corresponds roughly to the number gaps in the population responses that are bounded by coactive cell groups of



**Figure 3.** Example Betti curve correlation distributions. Stimulus selectivity of the Betti curves is assessed by computing the distributions of the pairwise correlations between Betti curves on pairs of trials. Shown are distributions from bird B1083. Blue distributions come from trial pairs that belonged to the same stimulus. Orange distributions come from pairs that belong to different stimuli. The original, empirical distributions show significant differences, indicating stimulus selectivity of the Betti curves. The distributions computed from shuffled data show little difference among within or between stimulus pairs. The significance of the differences in these distributions was assessed using a linear mixed-effects model. Table 1 lists the significance for the empirical, fully shuffled, and within-stimulus mask shuffled conditions.

order  $n$ ; and as these gaps fill in, Betti  $n$  decreases. In other words, noise correlations increase the numbers of coactive cell groups. Overall, both the signal and noise correlation contribute to the topological structure of coactivity on each trial. This is consistent with previous studies (Jeanne et al., 2013) that show NCM signal and noise correlations are not independent.

### Topological tools for comparing population spiking activity

Representing population spiking activity with a simplicial complex has the advantage of providing a single mathematical object that encodes the entirety of the spatiotemporal structure of the population response on a single trial. This entity is non-numerical, however; and so to facilitate numerical comparisons, we sought to define a numerical measure of similarity between simplicial complexes. To do this, we generalized recent advances in network theory to define information-theoretic measures of simplicial complex structure and similarity. A detailed description of these measures is provided in Materials and Methods; Table 2 gives the principle formulas in this analysis. We refer to this general approach of computing information theoretic quantities from simplicial complexes as SLSE.

### Fitting spiking models using SLSE

The KL divergence is often used as a cost function for statistical model fitting, in which one attempts to choose the parameters of a statistical model such that the model distribution is as close to the data distribution as possible. The KL divergence can be defined to allow for this same sort of model fitting, but for graphs (De Domenico and Biamonte, 2016). We extended the KL divergence to simplicial complexes (see Materials and Methods) with

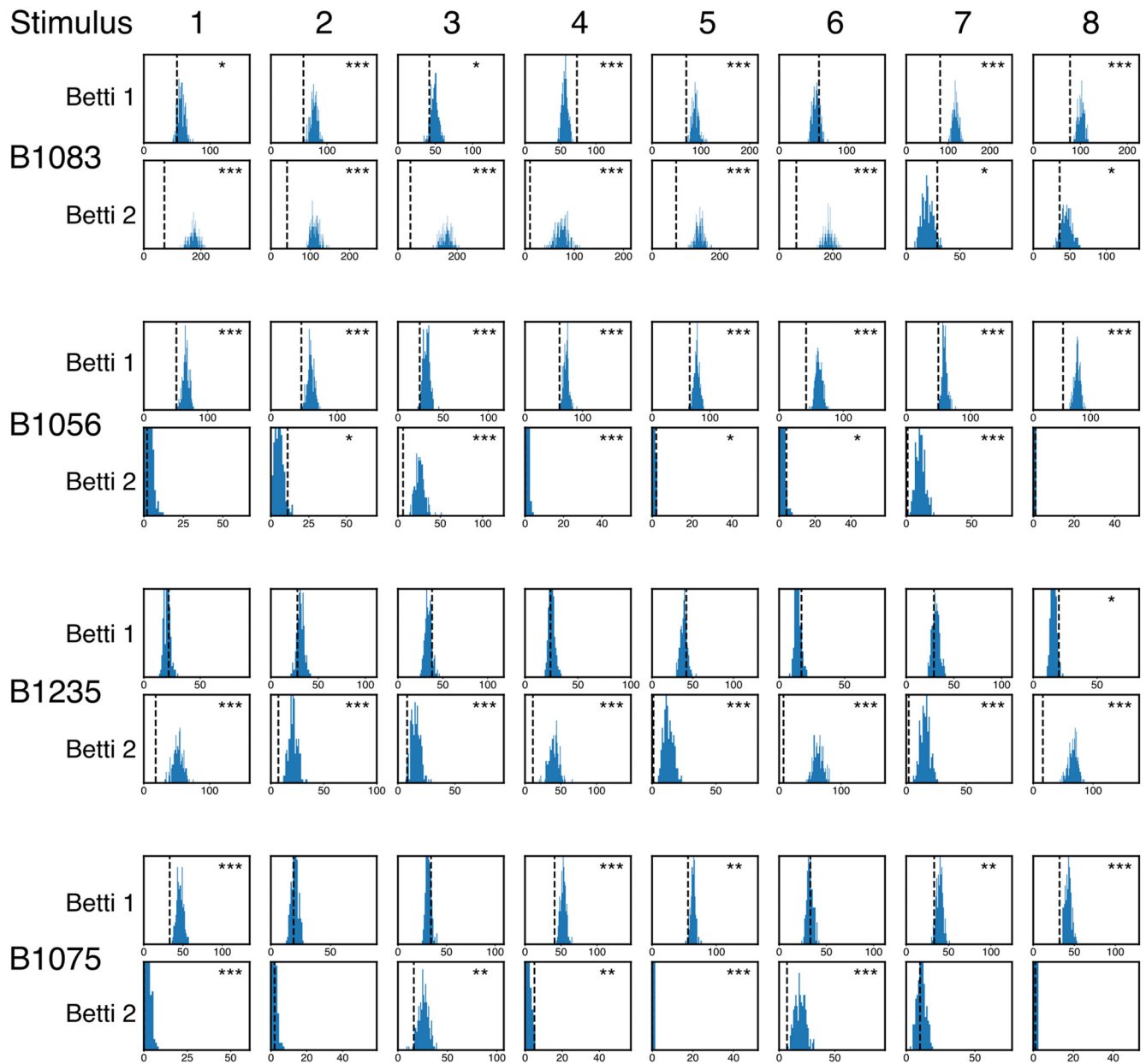
the reasoning that it could be used similarly to fit the parameters of spiking neural network models. To test this reasoning, we began with a proof of concept: fitting the rate parameter of a simulated population of Poisson-spiking neurons. The population consisted of 20 neurons each firing independently and with a rate parameter that was constant across the population. We simulated a single trial of 1000 time samples to serve as the “target” population spike train. We then performed a parameter sweep by simulating the population with a rate parameter that varied in some range. For each choice of rate parameter, 25 “test” trials from the model were simulated and the 1-KL divergence between the “target” population spike train and each “test” trial population spike train was computed. The 1-KL divergences from each of the 25 trials of the chosen parameter were averaged together. Figure 6a displays the 1-KL divergence as a function of the test rate parameter for four different models with different “target” rate parameters. The dotted lines indicate the true value of the rate parameter used to generate the target population spike train for each model, separated by color. For each model, the minimum of the KL divergence closely approximates the true value of the rate parameter.

We next tested a slightly more complicated model that had two distinct populations of neurons. Each population was again Poisson, but the two populations had differing rate parameters. Figure 6b shows a heat map plotting the 1-KL divergence as a function of the two rate parameters. The white dots indicate the true parameters. There is a degeneracy in this model in that it does not matter which subpopulation is labeled A and which is labeled B. Again, the minima of the 1-KL divergence closely approximate the true values of the parameters for the heterogeneous population.

Importantly, for all of the Poisson spike train simulations, the “test” and “target” spike trains never coincided. Instead, similarities quantified by the simplicial KL divergence rely on the global topological structure of the population spike trains, not on the match between specific spike trains of single neurons. The SLSE is sensitive to how the activity of an individual neuron relates to the simultaneous activity of all the other neurons in the population in which it is embedded.

### Reconstructing latent stimulus relationships using SLSEs

Our observation that NCM populations produce nonrandom, stimulus-specific, topologies implies that Poisson spike trains are not a good source of biologically relevant topologies, and so the above evaluation of these metrics is limited. As a more biologically relevant validation of the SLSE measures, we turned to the place cell system used in the original Curto and Itskov (2008) study of neural population-derived simplicial complexes and began by replicating these earlier results. We simulated the activity of hippocampal place cells during free exploration of a

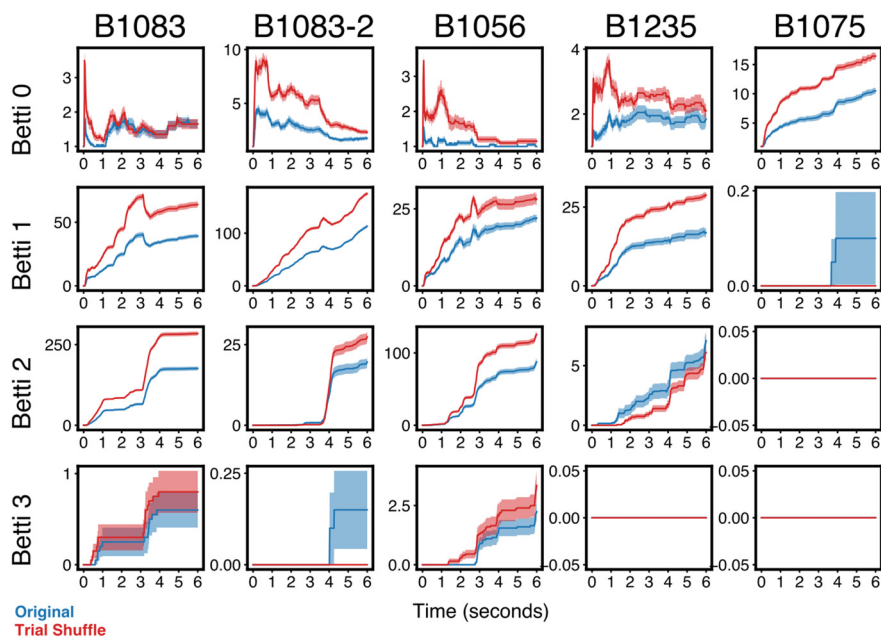


**Figure 4.** Most Betti numbers lie outside their null distributions. Each plot represents the null distribution of values of a given Betti number for a given stimulus and neural population generated by the simplicial configuration model (see Materials and Methods). Black dashed lines indicate the observed value of the Betti numbers in dimensions 1 and 2 at the end of the trial-averaged response to the indicated stimulus. \*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ ; the degree to which the dashed lines lie outside of the null distributions, interpreted as an empirical  $p$  value.

2-dimensional rectangular arena, within which we randomly placed different numbers of circular “holes.” A virtual animal explored the space through a random walk trajectory that was excluded from the holes in the simulated environment. Population spike trains were generated by covering the simulated environment with place fields and, for each field, generating spikes from simulated place cells as a function of the random walk trajectory. Environments differed in the number and position of the holes in each space, and the position of the simulated place fields.

The functional utility of any neural representation rests on its ability to encode the relationships between stimuli. Consider, for example, a pair of our simulated environments that happen to share the same number of holes. This invariance in the stimulus is independent of how different sets of place fields might cover either environment, or the paths that different animals might

take when navigating through either environment. If the neural topology mirrors the topology of physical stimulus structure, as is thought to be the case (Curto and Itskov, 2008), then the relationships between our simulated environments should be reflected in some of the relationships between the associated neural topologies. Thus, we hypothesize that the simplicial JS divergence on simulated population spike trains from these environments should capture the invariant relationships between environments. That is, population spike trains from an environment with  $N$  holes should be most similar to those from other environments with  $N$  holes, and the simplicial JS divergence should increase as the difference between the number of holes in the environments increases. The results of our simulations (Fig. 6*c–e*) support this hypothesis and show that the simplicial JS divergence can detect similarities in the neural representation



**Figure 5.** Trial-shuffled Betti curves. Red curves result from shuffling spikes across trials rather than across time. Betti curves show similar trajectories but different magnitudes. A single stimulus from each population is shown, and this stimulus is different from that displayed in Figure 2.

induced by relationships between the stimuli. As the difference in the number of holes between a pair of environments grows, the 1-JS divergence between spike trains from each environment grows. Importantly, these similarities persist across trials with unique paths through the arena, across populations that sample from different receptive fields, and across environments with different arrangements (but similar numbers) of holes. In each case, the SLSE technique detects similar representations of similar environments. Although two spike trains may manifest wildly different spatiotemporal spike patterns, the structure encoded in their coactivity patterns can encode invariant relationships between the stimuli they represent. Because stimuli have no intrinsic value to the neural population, we argue that these relationships define the stimuli themselves.

### Spectral entropy-based divergences between *in vivo* neural population activities

Having demonstrated that SLSE-based divergences are useful in quantifying invariant representations in simulated neural populations, we next applied these techniques to activities recorded from the brains of anesthetized birds trained to perform a two-alternative choice task (see Materials and Methods). Briefly, birds were trained using established techniques to peck at assigned locations on an operant panel in response to different acoustic stimuli, responding “left” when presented with two pseudo-songs and to respond “right” when presented with another two. From this training, the birds learned an arbitrary categorical structure (invariance) independent of any similarities in physical acoustics of the stimuli. We hypothesized that since the simplicial JS and KL divergences are designed to capture latent invariant structure in population activity, these divergences should reveal the learned categorical structure imposed on the stimuli by the behavior and presumably reflected in the population-level representation of the stimuli. Figure 7 shows the JS divergence between 1-Laplacians computed from the single-trial responses to both learned and unlearned songs in B1083, the bird with the

most units recorded. Figure 7*a, b* shows the divergences for learned stimuli in two disjoint populations in the same bird. Population 2 was located deeper than Population 1 by more than one probe length, in this case 600  $\mu\text{m}$ . The stimuli are organized so that the two left-signaling stimuli are adjacent, followed by the two right-signaling stimuli. In both populations, the block structure of the JS divergence matrix shows that temporal coactivity patterns elicited by stimuli in different behavioral classes are “farther apart” (less similar) than the coactivity patterns elicited by different stimuli in the same behavioral class. This allows for the reconstruction of the learned latent categorical structure of the stimuli, supporting our hypothesis. Shuffling individual neurons’ spike trains in time abolishes the categorical structure (Fig. 7*c, d*). As a control, computing the JS divergences between trials for four unfamiliar stimuli shows no consistent categorization of the stimuli among the two populations (Fig. 7*e–h*).

To see whether learned behavioral categories could be encoded by the physical similarity of the population spike trains, we computed the correlations (see Materials and Methods) between the population spike trains on single trials. Figure 7*i–l* show the dissimilarity ( $1 - \text{Correlation}$ ) for each pair of trials. Under this measure, physically similar population responses result in lower values, like the JS divergence. Not surprisingly, the single-trial population responses to individual stimuli are physically similar (Fig. 7*i–l*), as the population correlation reflects the trial-to-trial repeatability of the stimulus-locked timing of spikes in individual neurons. In contrast to the JS divergence, which captures the relative timing of spikes across neurons, the learned behavioral categories are not observable in the correlation (Fig. 7*i, k*). Thus, behaviorally relevant invariances captured by the coactivity spiking pattern are not observable in a measure of trial-to-trial repeatability that considers each neuron as independent.

To quantify the foregoing observation, we trained logistic regressions, using either the JS divergence scores or the raw population correlations, to predict whether the population responses to two different stimuli belonged to either the same or different learned classes, then tested the accuracy of each decoder on held-out data (see Materials and Methods). Figure 7*m* shows the accuracy of each decoder. For both populations, the accuracy of the decoder using JS divergence was significantly higher than the decoders using correlations, in both populations (*t* test,  $t = 95.36, 34.6$ ;  $p = 3.1\text{e-}313, 2.96\text{e-}132$ ). Both population response measures outperformed decoders trained on shuffled datasets (JS divergence: *t* test;  $t = 122.4, 44.49$ ;  $p < 1\text{e-}313, = 4.62\text{e-}172$ ; correlation: *t* test;  $t = 12.42, 29.62$ ;  $p = 6.36\text{e-}31, 2.87\text{e-}110$ ), indicating that even the raw correlations contain some information about the learned invariances.

If neurons are not independent, then trial-to-trial repeatability may not manifest in the stimulus-locked spiking of individual neurons but may instead be present in their statistical

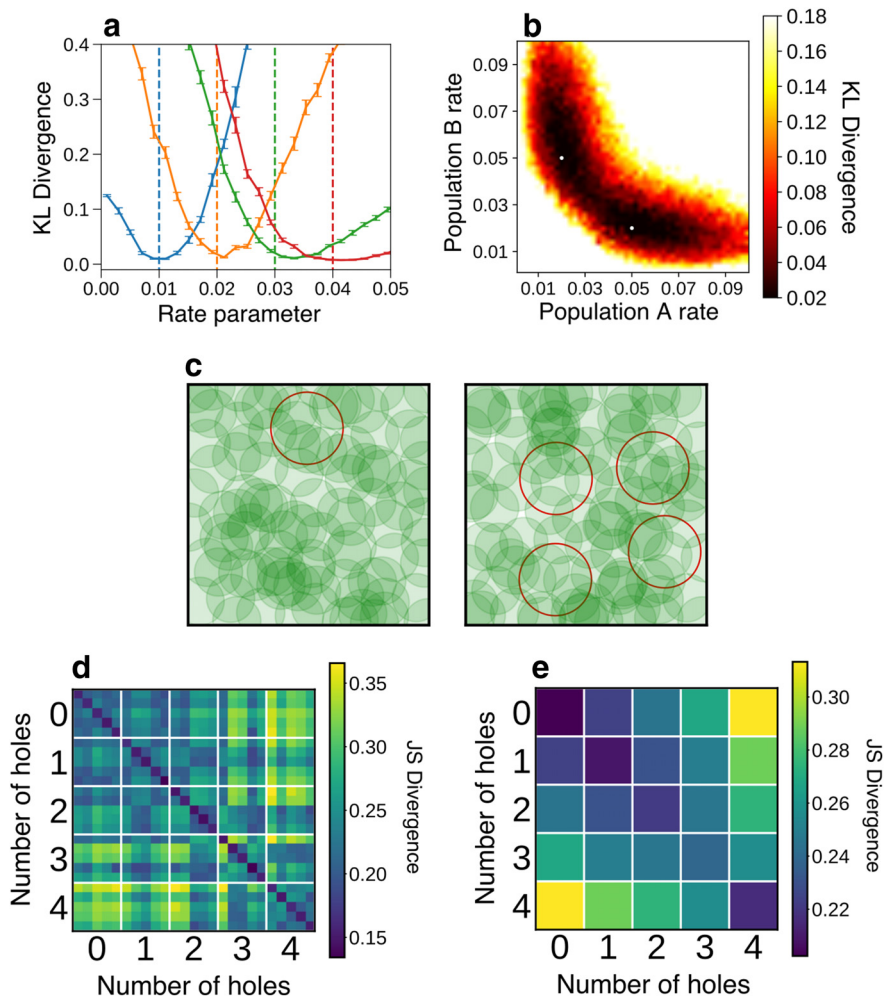
**Table 2. Summary of the main formulas for SLSE<sup>a</sup>**

Quantity	Formula	Major property
Simplicial Laplacian	$L_i = \partial_i^* \partial_i + \partial_{i+1} \partial_{i+1}^*$	Dimension of kernel is <i>i</i> -th Betti number
Density matrix	$\rho = \frac{e^{-\beta L}}{\text{Tr} e^{-\beta L}}$	Eigenvalues form discrete probability distribution
KL divergence	$D_{KL}(\rho, \sigma) = \sum_i \rho_i (\log \rho_i - \log \sigma_i)$	Measure of distinction between density matrices
JS divergence	$M = \frac{\rho + \sigma}{2}, D_{JS}(\rho, \sigma) = \frac{1}{2} (D_{KL}(\rho, M) + D_{KL}(\sigma, M))$	Symmetrized version of KL divergence

<sup>a</sup>Included are the names, formulas, and the most important properties of each quantity. The formulas should be taken as definitions.

covariances. To test this, we computed the neuron-to-neuron pairwise covariance matrix for each trial (see Materials and Methods), then found the normalized distance between these covariance matrices for pairs of trials (see Materials and Methods). Figure 8c shows the distances between these covariance matrices. As in the independent neuron correlations (Fig. 7i–l), we did not observe a strong signature of learned behavioral categories in the similarity between pairwise covariance matrices. We again quantified this observation by comparing the accuracy of logistic regression-based decoders trained to detect the learned behavioral class from the distance between covariance matrices on pairs of trials (Fig. 8e) to that for the JS divergence. Accuracy using the JS divergence was significantly higher than that for the covariance matrix distance (*t* test,  $t = 102.43, 89.62; p < 1e-313, = 5.0e-301$ ; Fig. 7e).

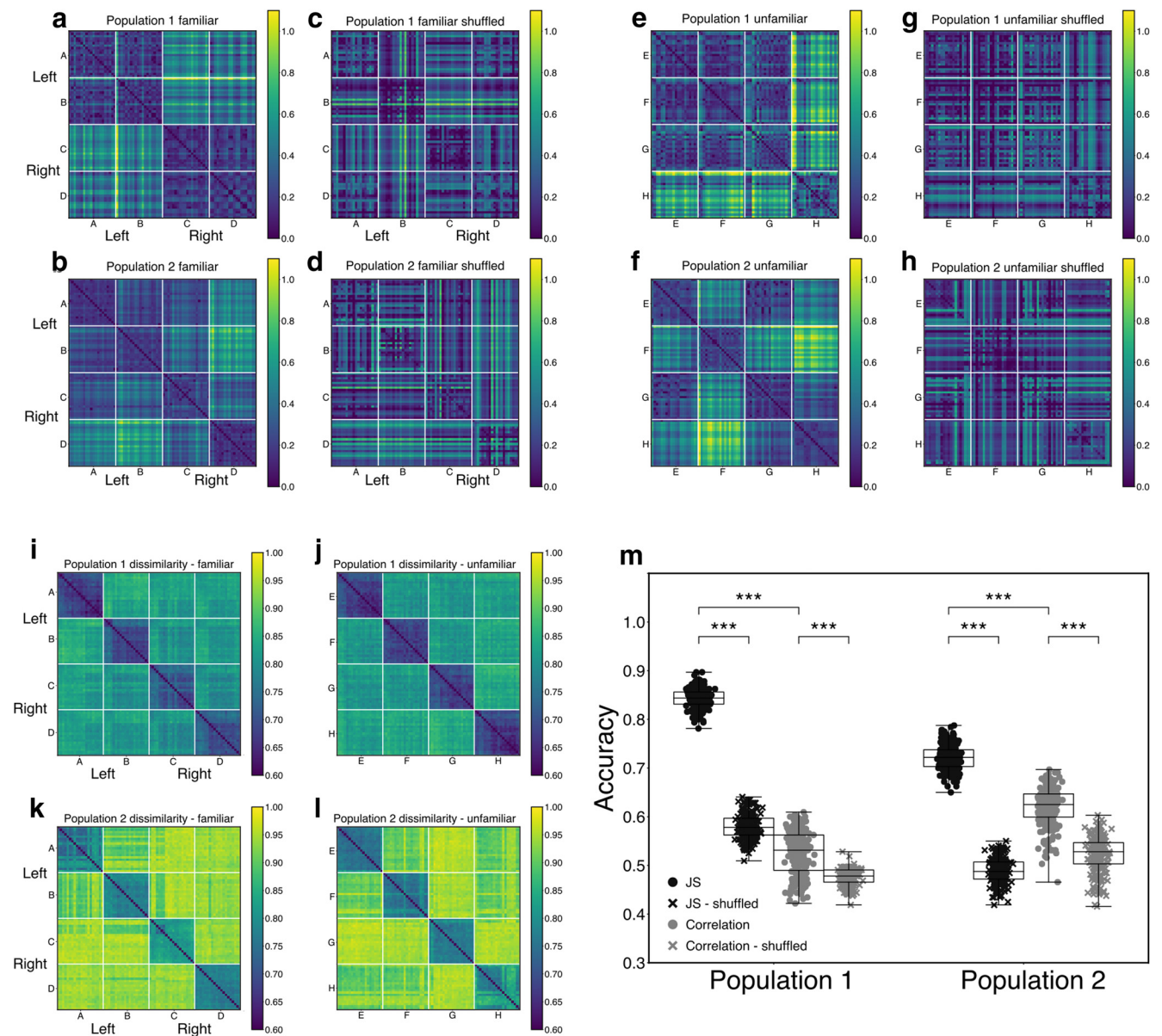
To understand why the JS divergence might capture information in the population representation that measures of correlation do not, it is helpful to note that statistical measures of coactivity (e.g., correlation) assume coactivity is tied to fixed collections of neurons. Correlations at the population level (as in Fig. 7) assume that each neuron is independent, and simply track response reliability within each neuron across trials. Pairwise (or higher-order) correlations (as in Fig. 8) require one to sample a defined pair (or larger set) of neurons multiple times, either across time within a trial or across trials to estimate variance. Moreover, because most pairwise responses are not correlated, these effects can wash out as population sizes increase. The JS divergence measure is, by definition, invariant to the arbitrary labeling of neurons, can be computed over any interval, and is immune to addition of noncoactive neurons. Thus, the population coactivity pattern is different from a correlation. To highlight this, we computed the quantity 1-correlation between population responses on each trial, having relabeled the neurons independently for each trial. The resulting distance matrix is shown in Figure 8d, again demonstrating a lack of segregation by behavioral class. Comparing



**Figure 6.** SLSE captures invariant population activity structure. **a**, KL divergence as a function of the rate parameter for four different models (represented by color) show that the minimum aligns with the target parameter (dotted vertical line) in each model. Error bars indicate SEM. **b**, Heat map showing KL divergence as a function of rate for a model consisting of two independent Poisson subnetworks with different rate parameters. White dots indicate the true values for each network. **c**, Two example simulated environments with 1 and 4 holes (red circles) tiled by place fields (filled green circles). **d**, Average JS divergence between simplicial complexes derived from simulated place cell activity in environments with the indicated number of holes. **e**, Same as in **d**, but averaged over all similar environments, showing that the JS divergence grows with the differences between environments.

this measure of correlation to the JS divergence using logistic regression (Fig. 8e) again revealed a significant difference between the two measures (*t* test,  $t = 144.62, 51.06; p < 1e-313, = 1.12e-195$ ).

We performed the same logistic regression analysis between the 1-JS and correlation measures of response similarity on the neural populations obtain from the remaining birds. B1056 showed a statistically significant difference between JS and correlation (*t* test;  $t = 6.09, p = 2.3e-09$ ) and between JS and covariance

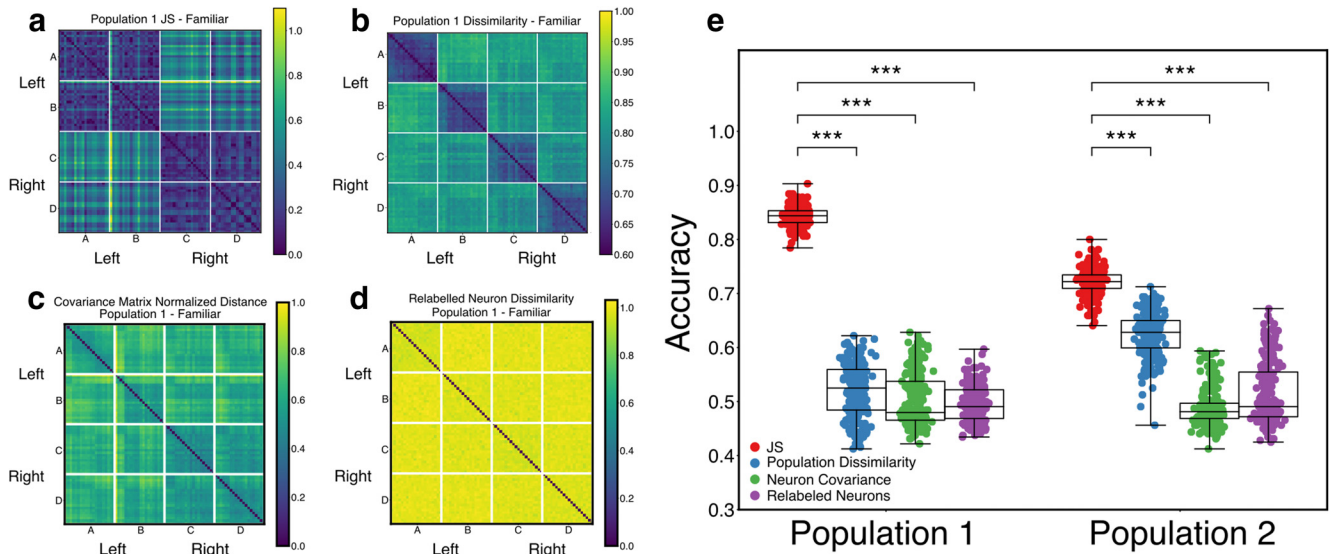


**Figure 7.** JS divergence between *in vivo* neural populations. **a–h**, Each pixel represents a JS divergence between two trials. White lines indicate trials belonging to different stimuli, identified by letters A–H. **a, b**, JS divergence between trials for familiar stimuli for Populations 1 and 2, respectively, on which the bird was trained to peck either left or right as indicated. Both populations show categorical structure reflecting the learned behavioral categories. **c, d**, Same data as in **a, b**, but each neuron's spike train has been shuffled in time, breaking correlations but keeping absolute firing rates, and abolishing categorical structure. **e, f**, JS divergence between trials for untrained stimuli. No overarching categorical pattern is observed. **g, h**, Results from shuffling the data in **e** and **f**, respectively. **i–l**, Correlation dissimilarity between population spike trains (see Materials and Methods), where 0 or 1 indicates perfect or no correlation, respectively, and behavioral categories are not observable for either population. **m**, Accuracy of predicting same/different behavioral classes from the two different populations using either the JS divergence between simplicial complexes (black) or spike train correlation (gray). \*\*\* $p < 0.001$ .

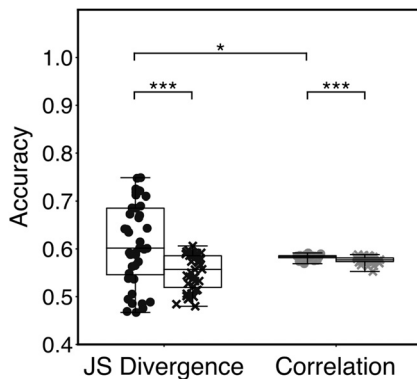
matrix distance ( $t$  test;  $t = 7.06$ ,  $p = 5.7e-12$ ). B1235 also showed a statistically significant difference between JS and correlation ( $t$  test;  $t = 11.8$ ,  $p = 1.4e-28$ ) and between JS and covariance matrix distance ( $t$  test;  $t = 5.85$ ,  $p = 9.25e-9$ ), but the regressions on the JS from fully shuffled data slightly outperformed the JS on empirical data ( $t$  test;  $t = -8.36$ ;  $p = 6.5e-16$ ). B1075 did not have enough neurons in the population to robustly compute simplicial Laplacians on individual trials. In total, the JS divergence outperformed correlations in all of the five populations for which the JS divergence was computable.

The results so far suggest that the topological structure within a population is specific to behavioral classes. We next asked whether this topological structure was similar across disjoint but simultaneously recorded populations. To test this, we took

Population 1 from B1083 and split it into two disjoint populations containing equal numbers of neurons. Then, we computed both the JS divergence and correlation between individual trials both within and between these subpopulations. Finally, we computed the accuracy of a logistic regression model to predict whether trials from physically distinct stimuli nevertheless belonged to the same behavioral class. We repeated this measurement for 40 random, independent splits of the original population. Figure 9 shows that the JS divergence outperforms correlation (glm;  $z = 2.501$ ;  $p = 0.012$ ), indicating there is some similarity between the topological structures produced by simultaneous but independent populations that reflects the learned behavioral class. Furthermore, both the JS divergence and correlation outperform the shuffled versions of their own measures



**Figure 8.** JS Divergence outperforms correlational measures of coactivity. *a, b*, Distance matrices determined using the JS divergence and population dissimilarity (as in Fig. 6). *c*, Normalized distance between neuron–neuron covariance matrices on pairs of trials, showing no segregation into behavioral classes. *d*, Correlation between trials with neuron identities randomly relabeled on each trial, showing no segregation by behavioral class. *e*, Accuracies of a logistic decoder trained to determine whether a pair of trials belongs to the same or different behavioral classes for each measure of population similarity. The JS divergence significantly outperforms all other measures tested, for both populations.  $***p < 0.001$ .



**Figure 9.** JS Divergence shows behavioral selectivity across disjoint populations. Accuracy of logistic regression for predicting whether two trials from different stimuli belong to the same behavioral class, for either JS divergence (black) or correlation (gray). Each dot represents a random split of the population into two disjoint subpopulations. Xs correspond to shuffled data. Included in the regression is the JS divergence or correlation between the responses from the joint population, indicating that simultaneous disjoint populations share some topological similarity.  $*p < 0.05$ .  $***p < 0.001$ .

(*t* test;  $t = 3.876, 4.61$ ;  $p = 0.0003, 1.57e-5$ ). The large variance in the JS divergence we attribute to the fact that the split populations have half as many neurons as in Figure 7, leading to a “lower resolution” reconstruction of the topology.

## Discussion

We have shown that the population activity in a secondary auditory region of a songbird contains nontrivial structure in the temporal relationships between the spiking activities of individual neurons. The population coactivity structure can be captured directly using the Curto-Itskov construction. We demonstrated that the population coactivity, as captured by its topology, is not due solely to the aggregation of independent, stimulus-specific responses from individual neurons. Instead, the structure emerges intrinsically at the population level. Using novel mathematical tools to compare simulated topological population

structures, we revealed an invariant structure common to distinct neural populations, which corresponds to learned invariances between sensory stimuli. Using these same tools to examine *in vivo* neural population activity, we showed that learned categorical relationships between natural stimuli yield invariant relationships in the topological population structure of avian auditory cortex, despite low pairwise correlations between single-trial spike trains. From these results, we conclude that algebraic topology offers a valuable tool for understanding invariant representations in neural populations.

One concern with computing Betti curves from spike coactivations aggregated across time (temporal filtration) is that “noise” in the responses may lead to spurious coactivations that fill in (and thus destroy) the holes that the Betti curves measure. While valid, we did not consider it a severe limitation in this case because declaring any particular spikes to be noise amounts to an assumption of what constitutes “signal,” an assumption that we did not want to make. Instead, we intended for the actual observed topology to define the signal. We wanted to see how individual trials related to each other through their topology, agnostic to the precise origins of the coactivations in the population. This informed our decision to use the simplicial complex that results from coactivations aggregated across the whole trial, rather than the most persistent homological features. This also motivated our use of time as a parameter in the Betti curves, to characterize how the simplicial complex evolved to this final state. Nevertheless, the results presented necessarily include contributions from possible “spurious” spikes, and exploration and comparison of other filtration strategies on similar datasets would be an important addition to the literature.

Our analysis differs from other approaches to population activity in that we do not begin by separating stimulus-induced structure (signal correlations) from within-trial covariation (noise correlations). Our approach does not consider the population responses as variations around a mean response from individual neurons but rather seeks to understand how a single event is represented by the population on a single “trial.” There is increasing evidence that there exists stimulus-specific noise

correlations (Gu et al., 2011; Jeanne et al., 2013; Ramalingam et al., 2013; Bondy et al., 2018), and separating these sources of variance may be an artificial convenience of the way responses have traditionally been measured. Separating these two forms of correlation at the outset of an analysis is liable to miss behaviorally relevant structure in the population response. Our results suggest that structures in the population response do carry information about behaviorally relevant stimuli. Furthermore, these structures are not reducible to collections of individual neuron responses. Thus, these coincidence patterns carry meaning of their own.

This approach is a conceptual alternative to neural representations centered on the concept of receptive fields. By definition, computing a receptive field requires access to both neural activity and the ground-truth stimulus. Since neural networks do not have access to the stimulus outside of their own stimulus-induced activity, neural networks cannot, in principle, perform this computation. In other words, they do not have access to their own receptive fields (Curto and Itskov, 2008). However, because natural stimuli are not random, structure in the stimulus will induce structure in the firing responses of neurons (Brette, 2015). One way this structure manifests is as temporal structure of spike coactivations in populations of neurons. Neural networks can in principle access this structure because it requires only the ability to detect temporal coincidences among spikes within the neural population itself, and no information from outside the animal. Thus, the approach avoids some of the conceptual difficulties with receptive fields. At the same time, the approach is consistent with the idea of receptive fields. Indeed, a key insight of Curto and Itskov (2008) is that relationships between spikes captured in the topology of the population response are sufficient to reconstruct some features of the physical environment without knowledge of explicit place fields. In other words, while receptive fields can be computed, they are not required.

Receptive field-based analyses also require strong *a priori* assumptions about the basis feature space for external stimulus representation. The Curto-Itskov construction allows for an alternative definition of stimulus space, as the topological space constructed from neural activity. This alternative definition offers advantages in higher-order sensory systems. Like other natural signals, the relevant stimulus space for birdsong does not easily lend itself to simple (low-dimensional) parameterizations, such as, for example, the orientation angle of drifting gratings that one might use to parameterize simple stimuli for the visual system. While useful in quantifying neural responses, these experimenter-imposed parameterizations are not guaranteed to align with the dimensional axes for neural “tuning.” In contrast, the intrinsic relationships between neuronal responses, regardless of their receptive field tuning, are precisely the response properties that are parameterization invariant. The topological constructions used here directly capture some of these relationships. As for most natural stimuli, we do not have access to the “correct” parameterized stimulus space for bird song. Using the topological methods described here allows us to avoid this problem altogether, measure invariant properties of the neural representation, and reconstruct a neurally defined stimulus space intrinsically defined by secondary auditory neurons in songbirds.

Understanding invariance is central to understanding perception, and the general notion features prominently in a longstanding debate regarding the nature of sensory systems and neural computation. The classical view, identified most strongly with Marr (1982), conceives of sensory systems as computational

pipelines, taking discrete stimuli (or stimulus features) as inputs and implementing algorithms to produce more complex representations as output. Accordingly, a class of algorithms may be specified to extract similarities from multiple stimulus representations and output an invariant representation (Riesenhuber and Poggio, 1999). A contrasting view, articulated by Gibson (1986), posits that sensory systems encode invariances directly, through a sensitivity to the patterns of information in the environment that define the relationships (i.e., changes or the lack thereof) between stimuli. Neuroscience has the potential to inform these deep debates on the nature of perception; but to do so, we must improve the currently limited methods for quantifying invariant representations.

Searching for invariants in neural activity, however, by correlating aggregate measures of neural activity with stimulus properties is problematic. As noted, the animal cannot perform the same computation (Curto and Itskov, 2008) because it does not have direct access to the stimulus. Gibson used this “closed” property of sensory systems to infer that invariant representations must be encoded directly from patterns of information in the environment (Gibson, 1986). We have demonstrated how to extract potential sensory representational invariants using only the information directly available to the nervous system, the spatiotemporal relationships between spikes. These techniques provide a route to understanding representational invariances and their foundational role in perception.

The precise computational mechanisms for how neural populations might make use of this topological structure remain a topic for future research. However, previous studies illuminate possible avenues of investigation. In networks dynamically balanced excitation-inhibition, neurons can be highly sensitive to these temporal coincidences, with a sensitivity surpassing the limits of experimentally measurable statistical correlations between spike trains (Rossant et al., 2011). A state of excitation-inhibition balance is commonly observed throughout the cortex of mammals and has been observed in songbird NCM neurons (Rossant et al., 2011; Perks and Gentner, 2015), and could support the use of temporal coincidences as a substrate for computation (Brette, 2012). As a general phenomenon, coincidence detection is a nearly ubiquitous capability in neurophysiological systems.

A growing body of literature investigates neural systems from the viewpoint of topology. Recent demonstrations that driven and spontaneous activity produces characteristic dynamics in the functional topology (Reimann et al., 2017) agree qualitatively with our findings. Likewise, persistent homology applied to spontaneous and evoked activity in cortical area V1 of monkeys is able to reconstruct the spherical topology of orientation-selective cells (Singh et al., 2008), and theoretical work demonstrates how topological structure in the hippocampus can persist in the face of noise (Babichev and Dabaghian, 2017). Our work represents the first use algebraic topology to examine representational invariance in a sensory system. Combined with the earlier theoretical and empirical studies, this work helps to establish neural topology as a useful tool to investigate neural activity.

## References

- Babichev A, Dabaghian Y (2017) Transient cell assembly networks encode stable spatial memories. *Sci Rep* 7:3959.
- Bailey DJ, Rosebush JC, Wade J (2002) The hippocampus and caudomedial neostriatum show selective responsiveness to conspecific song in the female zebra finch. *J Neurobiol* 52:43–51.
- Bondy AG, Haefner RM, Cumming BG (2018) Feedback determines the structure of correlated variability in primary visual cortex. *Nat Neurosci* 21:598–606.



- Brette R (2012) Computing with neural synchrony. *PLoS Comput Biol* 8:e1002561.
- Brette R (2015) Philosophy of the spike: rate-based vs. spike-based theories of the brain. *Front Syst Neurosci* 9:151.
- Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV (2012) Neural population dynamics during reaching. *Nature* 487:51–56.
- Cunningham JP, Gilja V, Ryu SI, Shenoy KV (2009) Methods for estimating neural firing rates, and their application to brain-machine interfaces. *Neural Netw* 22:1235–1246.
- Curto C, Itskov V (2008) Cell groups reveal structure of stimulus space. *PLoS Comput Biol* 4:e1000205.
- Davison AC, Hinkley DV (1997) *Bootstrap methods and their application*. Cambridge, MA: Cambridge UP.
- De Domenico M, Biamonte J (2016) Spectral entropies as information-theoretic tools for complex network comparison. *Phys Rev X* 6:041062.
- De Groof G, Poirier C, George I, Hausberger M, Van der Linden A (2013) Functional changes between seasons in the male songbird auditory forebrain. *Front Behav Neurosci* 7:196.
- Gallego JA, Perich MG, Naufel SN, Ethier C, Solla SA, Miller LE (2018) Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nat Commun* 9:4233.
- Ghrist R (2014) *Elementary applied topology*. Scotts Valley, CA: Createspace.
- Gibson JJ (1986) *The ecological approach to visual perception*. London: Psychology.
- Giusti C, Ghrist R, Bassett DS (2016) Two's company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data. *J Comput Neurosci* 41:1–14.
- Grace JA, Amin N, Singh NC, Theunissen FE (2003) Selectivity for conspecific song in the zebra finch auditory forebrain. *J Neurophysiol* 89:472–487.
- Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, Sunkara A, DeAngelis GC, Angelaki DE (2011) Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* 71:750–761.
- Hatcher A (2002) *Algebraic topology, illustrated*, reprint. Cambridge, MA: Cambridge UP.
- Horak D, Jost J (2013) Spectra of combinatorial Laplace operators on simplicial complexes. *Adv Math (NY)* 244:303–336.
- Ince RA, Panzeri S, Kayser C (2013) Neural codes formed by small and temporally precise populations in auditory cortex. *J Neurosci* 33:18277–18287.
- Jeanne JM, Sharpee TO, Gentner TQ (2013) Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* 78:352–363.
- Kaczynski T, Mischaikow K, Mrozek M (2006) *Computational homology, illustrated*. New York: Springer.
- Karten HJ (1967) The organization of the ascending auditory pathway in the pigeon (*Columba livia*): I. Diencephalic projections of the inferior colliculus (nucleus mesencephali lateralis, pars dorsalis). *Brain Res* 6:409–427.
- Kiggins JT, Comins JA, Gentner TQ (2012) Targets for a comparative neurobiology of language. *Front Evol Neurosci* 4:6.
- Knudsen DP, Gentner TQ (2013) Active recognition enhances the representation of behaviorally relevant information in single auditory forebrain neurons. *J Neurophysiol* 109:1690–1703.
- Kozlov AS, Gentner TQ (2016) Central auditory neurons have composite receptive fields. *Proc Natl Acad Sci USA* 113:1441–1446.
- Lindstrom MJ, Bates DM (1988) Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *J Am Stat Assoc* 83:1014.
- Marr D (1982) *Vision*. New York: Holt.
- Mischaikow K, Nanda V (2013) Morse theory for filtrations and efficient computation of persistent homology. *Discrete Comput Geom* 50:330–353.
- Nanda V (2013) Perseus, the persistent homology software. Available at <http://www.sas.upenn.edu/~vnanda/perseus>.
- Nicolini C, Vlasov V, Bifone A (2018) Thermodynamics of network model fitting with spectral entropies. *Phys Rev E* 98:022322.
- Perks KE, Gentner TQ (2015) Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex. *Eur J Neurosci* 41:725–733.
- Ramalingam N, McManus JN, Li W, Gilbert CD (2013) Top-down modulation of lateral interactions in visual cortex. *J Neurosci* 33:1773–1789.
- Reimann MW, Nolte M, Scolamiero M, Turner K, Perin R, Chindemi G, Dłotko P, Levi R, Hess K, Markram H (2017) Cliques of neurons bound into cavities provide a missing link between structure and function. *Front Comput Neurosci* 11:48.
- Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2:1019–1025.
- Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–590.
- Rossant C, Kadir SN, Goodman DF, Schulman J, Hunter ML, Saleem AB, Grosmark A, Belluscio M, Denfield GH, Ecker AS, Tolias AS, Solomon S, Buzsaki G, Carandini M, Harris KD (2016) Spike sorting for large, dense electrode arrays. *Nat Neurosci* 19:634–641.
- Rossant C, Leijon S, Magnusson AK, Brette R (2011) Sensitivity of noisy neurons to coincident inputs. *J Neurosci* 31:17193–17206.
- Sen K, Theunissen FE, Doupe AJ (2001) Feature analysis of natural sounds in the songbird auditory forebrain. *J Neurophysiol* 86:1445–1458.
- Singh G, Memoli F, Ishkhanov T, Sapiro G, Carlsson G, Ringach DL (2008) Topological analysis of population activity in visual cortex. *J Vis* 8:11.1–18.
- Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20:2315–2331.
- Thompson JV, Gentner TQ (2010) Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. *J Neurophysiol* 103:1785–1797.
- Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *J Comp Neurol* 366:613–642.
- Wang Y, Brzozowska-Prechtel A, Karten HJ (2010) Laminar and columnar auditory cortex in avian brain. *Proc Natl Acad Sci USA* 107:12676–12681.
- Young JG, Petri G, Vaccarino F, Patania A (2017) Construction of an efficient sampling from the simplicial configuration model. *Phys Rev E* 96:032312.