Title
Methods for Estimation and Control of Linear Systems Driven by Cauchy Noises

Author
Fernandez, Javier Huerta

Publication Date
2013

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Methods for Estimation and Control of Linear Systems Driven by Cauchy Noises

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mechanical Engineering

by

Javier Huerta Fernández

2013

Abstract of the Dissertation

Methods for Estimation and Control of Linear Systems Driven by Cauchy Noises

by

Javier Huerta Fernández

Doctor of Philosophy in Mechanical Engineering

University of California, Los Angeles, 2013

Professor Jason L. Speyer, Chair

An efficient recursive state estimator and an optimal model predictive controller are developed for two-state linear systems driven by Cauchy distributed process and measurement noises. For a general vector-state system, the estimator is based on recursively propagating the characteristic function (CF) of the conditional probability density function (cpdf), where the number of terms in the sum that expresses this CF grows with each measurement update. The proposed two-state estimator reduces substantially the number of terms needed to express the CF of the cpdf by taking advantage of relationships not yet developed in the general vector-state case. Further, by using a fixed sliding window of the most recent measurements, the improved efficiency of the proposed two state estimator allows an accurate approximation for real-time computation. For control of the general vector-state system, the conditional performance index is evaluated in the spectral domain using this CF. The expectation is of an objective function that is a product of functions resembling Cauchy pdfs. Using this method, the conditional performance index for a two-state system is obtained analytically in closed form by using Parseval's identity and integrating over the spectral variables. This forms a deterministic, non-convex function of the control signal and the measurement history that must be optimized numerically at each time step. Examples are given of both the estimator and the controller, to demonstrate their performance and expose their interesting robustness characteristics.

The dissertation of Javier Huerta Fernández is approved.

Robert M'Closkey

Steven Gibson

Thomas S. Ferguson

Jason L. Speyer, Committee Chair

University of California, Los Angeles

2013

To my parents, Randy and Emilia, and to my sister, Chantal. Without their love and support, this work would not have been possible.

To my fabulous family in LA, especially my aunt Vivian, for always reaching out to me from across the 405.

To Jason Speyer, from whose patience and guidance I have learned so much.

To Don Browne, for teaching me to teach and for helping me get through the middle years.

To the MAE staff, especially Angie, Abel, Lance, Samantha, Coral, Lili, Evgenia, David, and Heejin. Thank you for helping me navigate UCLA and complete the program, every step of the way.

And to my friends, you know who you are, wherever you are. Thank you for providing this wanderer with shelter when he showed up at your door.

The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way,
Where many paths and errands meet.
And whither then? I cannot say.
-B.B.

Table of Contents

List of Figures

Acknowledgments

This work was partially supported by Air Force Office of Scientific Research, Award No. FA9550-09-1-0374,

and by the United States - Israel Binational Science Foundation, Grant 2008040.

Vita

1985	Born, Los Angeles, California, USA.
2007	B.A. (Engineering) and A.B. (Spanish Language and Literature), Dartmouth College, Hanover, NH.
2008	B.S. (Engineering), Thayer School of Engineering at Dartmouth College, Hanover, NH.
2009–2011	Teaching Assistant, Department of Mechanical and Aerospace Engineering, UCLA. Taught sections of Mechanical Engineering 171A (classical control theory) and Engineering 183EW (engineering ethics and writing)
2009	M.S. (Mechanical Engineering), UCLA, Los Angeles, CA.
2011–2013	Research Assistant, Department of Mechanical and Aerospace Engineering, UCLA.

Publications

J.H. Fernández, J.L. Speyer, and M. Idan. "Linear Dynamic Systems with Additive Cauchy Noises Part 1: Stochastic Estimation for Two-State Systems." Submitted to *IEEE Transactions on Automatic Control*, 2013.

J.H. Fernández, J.L. Speyer, and M. Idan. "Linear Dynamic Systems with Additive Cauchy Noises Part 2: Stochastic Model Predictive Control." Submitted to *IEEE Transactions on Automatic Control*, 2013.

J.H. Fernández, J.L. Speyer, and M. Idan. "A Stochastic Controller for Vector Linear Systems with Additive Cauchy Noises." In *IEEE Conference on Decision and Control*, 2013.

J.L. Speyer, M. Idan, and J.H. Fernández. "A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise." *Automatica*, 2013.

J.L. Speyer, M. Idan, and J.H. Fernández. "The two-state estimator for linear systems with additive measurement and process Cauchy noise." In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pp. 4107-4114, Dec. 2012.

J.L. Speyer, M. Idan, and J.H. Fernández. "Mutli-step prediction optimal control for a scalar linear system with additive Cauchy noise." In *Decision and Control (CDC), 2010 IEEE 49st Annual Conference on*, pp. 4157-4164, Dec. 2010.

CHAPTER 1

Introduction

Gradually a hideous uncertainty came over me and I scrambled from the hole. One look at the newly made headpiece was enough. This was indeed my own grave ... but what fool had buried within it another corpse?

H.P. Lovecraft, "The Disinterment"

Dynamic processes involving uncertainty are frequently encountered in fields ranging from engineering and science to economics and finance. It is often assumed that the uncertainties are described by Gaussian probability distributions, mainly because modern methods and algorithms are able to handle such systems very efficiently [1]. However, in many applications the underlying random processes have an impulsive character producing deviations of high amplitude and small duration much more often than the Gaussian assumption permits [2]. Examples of such processes include radar and sonar noise [3] and disturbances due to air turbulence [4]. These uncertainties affect both the actual state of the system as well as the measurements that the controller depends on. For linear systems in modern stochastic optimal control, algorithms like the linear quadratic Gaussian (LQG) and the linear exponential Gaussian (LEG) assume linear dynamics and additive process and measurement noises described by the Gaussian pdf [1].

Impulsive uncertainties were shown to be better described by heavy-tailed distributions, such as the symmetric alpha-stable ($S\alpha S$) distributions [5]. These distributions are described not by their probability density functions (pdfs), but by their characteristic functions (CFs). They are of the form $\phi(\nu) = e^{-\sigma^{\alpha}|\nu|^{\alpha}+j\mu\nu}$, where σ is the scaling parameter, μ is the median, ν is the spectral variable, and the characteristic exponent α determines the type of distribution: $\alpha = 2$ implies a Gaussian distribution, and $\alpha = 1$ implies a Cauchy distribution.

Estimation assuming Cauchy distributed noises has shown improved performance over Gaussian estimators when faced with impulsive noises. For estimating the direction of arrival of a signal to a sensor array in [6], maximum likelihood estimators designed assuming Cauchy distributed noises were shown to exhibit performance very close to the Cramèr-Rao Bound against $S\alpha S$ noises with characteristic exponents $1 \leq \alpha \leq 2$. Similar performance was observed in various applications, including processing data in a multi-user communication network [7] and radar glint [8]; in particular, the α parameter for the in-phase component of a time series of sea clutter in radar in [6] was calculated to be $\alpha \approx 1.7$. A framework based on these stable distribution models was developed in [9] and shown to have significant improvements in performance against heavy tailed noises.

The apparent robustness and adaptability of the Cauchy probability model motivated the development of a sequential estimator for linear scalar systems driven by Cauchy distributed process and measurement noises [10, 11], as well as algorithms for optimal control of such systems in [12, 13]. There, the conditional performance index for model predictive control is determined by taking the conditional expectation of the objective function using the probability density given the measurement history. A dynamic programming algorithm is also developed in [13], where it is shown that the solution to the dynamic programming recursion is intractable because of the need to average over future measurements in determining the optimal return function. This cannot be done in closed form due to the complex dependency of the optimal return function on the measurement history, and hence we approximate the dynamic programming solution with model predictive control.

Subsequent work aimed to develop an estimator for general vector-state systems driven by Cauchy noise [14, 15]. For the vector state systems, the conditional pdf (cpdf) given the measurement history is not available. However, for these systems the CF of the cpdf can be recursively propagated [11, 14–17]. The estimator for general vector-state systems suffers from severe growth in numerical complexity, limiting its use to a small number of measurement updates and states. In Chapter 2 we present the derivation of the general measurement update and state propagation formulas using the CF approach of [14, 15], and we present the first measurement update for the general vector-state case. We also present theorems on the existence of moments of the conditional pdf for this general case.

The methodology in [14, 15] is based on finding the characteristic function of the conditional pdf of the state given the measurement history. The estimation algorithm presented in Chapter 3 follows that same procedure, and by exploiting certain relationships for the two-state structure, we can greatly reduce the complexity of the algorithm. That algorithm was derived inductively by working out the first three measurement updates, and then deducing the general update process [16, 17]. In Chapter 3, we present the first two measurement updates, followed by the general recursion, and then the third update is given as an example. The CF is expressed as a sum of terms, each of which has two components: a coefficient function denoted by $\mathcal{G}$ and an exponential function with argument $\mathcal{E}$. These functions are shown to have known structures that persist across measurement updates, and parameters that are contained in a set of fundamental arrays. The essence lies in deriving this structure and populating the arrays, which allows for a drastic reduction in the complexity of the algorithm. This result has been checked against the results in [14, 15].

In Chapter 4, the Cauchy optimal control algorithm for scalar systems [13] is extended to systems with a vector state [18, 19]. The significant contribution of that work is evaluating *in closed form* the conditional performance index using the cpdf's characteristic function instead of the cpdf itself, and thereby integrating over the spectral variables instead of the state variables [13]. The objective function is cast as a product of functions resembling Cauchy pdfs, which are easily transformed into functions of the spectral variables. The conditional performance index, found in a closed form, is a deterministic function of the control and measurement histories. Due to its complexity, the optimal control signal is determined by numerically optimizing this non-convex conditional performance index in a model predictive control setting. Finally, Chapter 5 presents concluding remarks as well as some thoughts on the future direction of this work.

1.1 Properties of $S\alpha S$ Distributions

The Gaussian distribution has many strong properties that have made it an extremely useful analytical tool for modeling uncertainty. How these properties change in the Cauchy distribution is the source of both the analytical power of the Cauchy model as well as the complexity it induces. Both the Gaussian and the Cauchy distributions belong to the class of probability distribution functions called the symmetric alpha-stable ($S\alpha S$) distributions [5]. As the name suggests, these distributions are symmetric, and also unimodal, and are stable in the sense that sums of random variables of a given $S\alpha S$ will be a random variable of the same type.

Distributions in the $S\alpha S$ class are described by their characteristic functions. In general the distribution functions for $S\alpha S$ distributions are unknown, making the Cauchy and Gaussian distributions exceptions. In this work we only refer to the subset of the $S\alpha S$ where $1 \leq \alpha \leq 2$, which contains both the Cauchy and the Gaussian distributions; in fact, the Cauchy and Gaussian distributions are the only members of this set that have known pdfs.

The characteristic function of a $S\alpha S$ distribution is given by

$$\phi(\nu) = e^{-\delta^{\alpha}|\nu|^{\alpha}} + jm\nu \tag{1.1}$$

These distributions are parameterized by a positive real number α; Cauchy distributions are those with $\alpha = 1$, and Gaussian distributions have $\alpha = 2$. The other two parameters, δ and m, refer to the scaling of the curvature and the median of the distribution, respectively. For the Gaussian distribution, then, the standard deviation is proportional to the parameter δ, and the mean is equal to m. The Cauchy pdf and its CF are given by

$$f_C(x) = \frac{\delta/\pi}{(x-m)^2 + \delta^2}, \qquad \phi_C(\nu) = \exp\left(-\delta\,|\nu| + jm\nu\right), \tag{1.2}$$

and the Gaussian pdf and its CF are given by

$$f_G(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \qquad \phi_G(\nu) = \exp\left(-\frac{1}{2}\sigma^2\nu^2 + jm\nu\right), \tag{1.3}$$

All members of this class of distributions are stable, meaning that linear combinations of random variables of a given α are also random variables of that type. The most important difference between the Cauchy and the Gaussian distributions is that the Cauchy has an infinite variance and an undefined mean, i.e., its first moment is undefined and its second moment is infinite; all the other distributions with $1 < \alpha < 2$ have an infinite variance and a finite mean. This lack of a finite second moment is what gives these distributions their impulsive character and differentiates them from the Gaussian distribution.

Another key difference is the lack of a unique general multivariate Cauchy pdf, as opposed to the Gaussian. The pdf of any n-dimensional Gaussian distributed random vector with variance $\Xi \in \mathbb{R}^{n\times n}$ and mean $m \in \mathbb{R}^n$ can be expressed as

$$f_G(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n|\Xi|}} \exp\left(\frac{1}{2}(x-m)^{\mathrm{T}}\Xi^{-1}(x-m)\right), \tag{1.4}$$

where $|\Xi|$ is the determinant of the variance Ξ. As in the scalar case, the distribution is determined by the variance and the mean, and retains its exponential structure. However, for the Cauchy case, there is no single, general multivariate structure [20]. That paper defines multivariate stable distributions as those where every one-dimensional marginal distribution. One such multivariate Cauchy distribution can be constructed as a product of scalar Cauchy distributions, where clearly every scalar marginal distribution is Cauchy. An advantage of using a multivariate Cauchy distribution of this form is that its CF can also be readily found.

$$f_C(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{\delta_i/\pi}{(x_i - m_i)^2 + \delta_i^2} \quad \Rightarrow \quad \phi_C(\nu) = \prod_{i=1}^{n} e^{-\delta_i\,|\nu_i| + jm_i\nu_i} \tag{1.5}$$

This assumption is fundamental to our approach, as the measurement update formula depends on this structure. Another structure for a multivariate Cauchy distribution is as an elliptical distribution, where the probability density is a function of a quadratic form [21]. Examples and properties of bivariate Cauchy distributions can be found in [20].

1.1.1 Generating Cauchy random variables

For simulation, it is necessary to generate Cauchy noise sequences. However, computer programs like MatLAB typically only generate Gaussian random variables. A Cauchy random variable can be generated from the quotient of two Gaussian random variables. This section will derive that relationship.

Denote two Gaussian random variables x_1 and x_2 with zero medians and standard deviations of σ and 1. Denoting $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, the joint distribution is given by

$$f_X(x_1, x_2) = \frac{1}{2\pi \left|P\right|} \exp\left(-\frac{1}{2} x^{\mathrm{T}} P^{-1} x\right) \tag{1.6}$$

where $P = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 1 \end{bmatrix}$ and $|P|$ is the determinant of P. Denote the function $g(\cdot)$ as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \triangleq y = g(x) = \begin{bmatrix} \frac{x_1}{x_2} \\ x_2 \end{bmatrix} \tag{1.7}$$

Then, use [1, Prop 2.17] to express the joint distribution of y_1 and y_2 as

$$f_Y(y_1, y_2) = \frac{1}{2\pi\sigma} \exp\left(-\frac{1}{2} \frac{(y_1 y_2)}{\sigma^2} + y_2^2\right) |y_2| \tag{1.8}$$

Finally, the marginal distribution of y_1, the Cauchy random variable, is given by

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma} \exp\left(-\frac{1}{2} y_2^2 \left(\frac{y_1^2}{\sigma^2} + 1\right)\right) |y_2| \, dy_2 \\ &= \frac{\sigma/\pi}{y_1^2 + \sigma^2} \end{aligned} \tag{1.9}$$

Using (1.9), a Cauchy random variable with scaling paramter σ can be generated from a Gaussian random variable with standard deviation σ divided by another Gaussian random variable with a standard deviation of 1.

1.2 Comparing Cauchy and Gaussian distributions

In the development and analysis of our estimators and controllers, we have tried to find ways of comparing our performance with that of established Gaussian estimators and controllers. This has involved finding Gaussian and Cauchy distributions that are "closest" to each other in some meaningful sense. We present two results, (1.17) and (1.19), which are very similar despite their different approaches. When comparing pdfs, it is not clear what "closest" means in a stochastic sense. In our results, we have taken the view that when, under Gaussian noises, a Cauchy estimator matches closely the mean and error variance of its closest Kalman filter (that is, a Kalman filter whose design parameters are closest to the Cauchy parameters). Both of these methods attempt to define closeness within that framework.

1.2.1 Closest in the Least Squared Sense

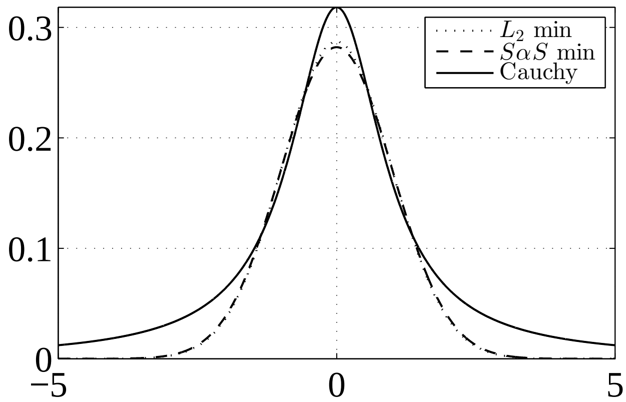


Figure 1.1: Comparison of a Cauchy pdf with $\delta = 1$ and its closest Gaussian pdfs, using both measures.

This view considers the pdfs as being in the space of square integrable functions. Minimizing the distance in L^2 should make the energies, so to speak, of the pdfs as close as possible. This was the comparison found first and thus is used in many of the papers. It produces an optimal ratio between the standard deviation σ of the Gaussian pdf and the scaling parameter δ of the Cauchy pdf. Denote the Gaussian and Cauchy pdfs as f_G and f_C, respectively,

$$f_G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \qquad\qquad f_C(x) = \frac{\delta/\pi}{x^2 + \delta^2} \tag{1.10}$$

Begin by evaluating the norm of the difference,

$$\begin{aligned} ||f_G - f_C||_2 = ||f_G - f_C||_2^2 = \int_{-\infty}^{\infty} (f_G - f_C)^2 \\ = \int_{-\infty}^{\infty} f_G^2 + \int_{-\infty}^{\infty} f_C^2 - 2\int_{-\infty}^{\infty} f_G f_C = ||f_G||_2^2 + ||f_C||_2^2 - 2\int_{-\infty}^{\infty} f_G f_C \end{aligned} \tag{1.11}$$

where the L^2 norms of the Gaussian and Cauchy pdfs are given by

$$||f_G||_2^2 = \frac{1}{2\sqrt{\pi}\sigma} \qquad\qquad ||f_C||_2^2 = \frac{1}{2\pi\delta} \tag{1.12}$$

In order to evaluate the last integral in (1.11), use Parseval's Identity and the CFs of the two pdfs as

$$\int_{-\infty}^{\infty} f_G f_C dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_G^* \phi_C d\nu = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2\nu^2}\cdot e^{-\delta|\nu|} d\nu \tag{1.13a}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{0} e^{-\frac{1}{2}\sigma^2\nu^2+\delta\nu} d\nu + \frac{1}{2\pi}\int_{0}^{\infty} e^{-\frac{1}{2}\sigma^2\nu^2-\delta\nu} d\nu \tag{1.13b}$$

$$= \frac{1}{\pi}\int_{-\infty}^{0} e^{-\frac{1}{2}\sigma^2\nu^2+\delta\nu} d\nu = \frac{e^{\frac{\delta^2}{2\sigma^2}}}{\pi}\int_{-\infty}^{0} e^{-\frac{1}{2}\frac{(\nu-\frac{\delta}{\sigma^2})^2}{\frac{1}{\sigma^2}}} d\nu \tag{1.13c}$$

where the last equality above comes from completing the square. This is as far as this can be taken analytically, since the Gaussian integral cannot be solved analytically for arbitrary bounds by using the error function $\mathrm{erf}\,(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$, which is what the computer provides. To do this, apply a change of variables $t = \frac{\nu-\frac{\delta}{\sigma^2}}{\frac{\sqrt{2}}{\sigma}}$ followed by some simplification, as

$$\int_{-\infty}^{\infty} f_G f_C = \frac{e^{\frac{\delta^2}{2\sigma^2}}}{\pi}\cdot\frac{\sqrt{2}}{\sigma}\int_{-\infty}^{\frac{-\delta}{\sqrt{2}\sigma}} e^{-t^2}dt = \frac{e^{\frac{\delta^2}{2\sigma^2}}}{\pi}\cdot\frac{\sqrt{2}}{\sigma}\frac{\sqrt{\pi}}{2}\left(1-\frac{2}{\sqrt{\pi}}\int_{0}^{\frac{\delta}{\sqrt{2}\sigma}} e^{-t^2}dt\right) \tag{1.14}$$

$$= \frac{e^{\frac{\delta^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\left(1-\frac{2}{\sqrt{\pi}}\int_{0}^{\frac{\delta}{\sqrt{2}\sigma}} e^{-t^2}dt\right) = \frac{e^{\frac{\delta^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\left(1-\mathrm{erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)\right) \tag{1.15}$$

Using (1.15) along with (1.12), take the partial derivative of (1.11) with respect to the Gaussian standard deviation σ,

$$\frac{\partial}{\partial\sigma}||f_G - f_C||_2^2 = \frac{\partial}{\partial\sigma}\left\{\frac{1}{2\sqrt{\pi}\sigma}+\frac{1}{2\pi\delta}-2\frac{e^{\frac{\delta^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\left(1-\mathrm{erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)\right)\right\}$$

$$= \frac{-1}{\sqrt{2\pi}\sigma^2}\left\{\frac{1}{\sqrt{2}}-2e^{\frac{\delta^2}{2\sigma^2}}\left(1-\mathrm{erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)\right)+\frac{4\delta}{\sqrt{2\pi}\sigma}-\frac{2\delta^2}{\sigma^2}e^{\frac{\delta^2}{2\sigma^2}}\left(1-\mathrm{erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)\right)\right\} \tag{1.16}$$

Then, (1.16) equals zero when the term in the brackets equals zero. Inside the brackets, the δ and σ parameters always appear in a ratio. Evaluating this numerically, the minimizing ratio is denoted κ and given by

$$\kappa := \min_{\sigma/\delta} ||f_G - f_C||_2 \approx 1.389801054561982 \tag{1.17}$$

1.2.2 Closest in the $S\alpha S$ Sense

This view considers the two pdfs as members of the space of $S\alpha S$ functions, which are parameterized by their medians m, scaling parameters δ, and alpha parameters α. Hence, in this view, the Gaussian pdf closes to a given Cauchy pdf is one with the same median and scaling parameter. We can find a constant that maps the Gaussian standard deviation to its $S\alpha S$ scaling parameter that can be equated to that of a Cauchy distribution. This constant is found by simply rewriting the Gaussian CF as

$$\phi_G(\nu) = \exp\left(-\frac{1}{2}\sigma^2\nu^2\right) = \exp\left(-\left(\frac{\sigma}{\sqrt{2}}\right)^2\nu^2\right) \tag{1.18}$$

Then, the ratio between a Cauchy distribution's scaling parameter δ and its closest Gaussian standard deviation σ is denoted κ and given by

$$\kappa := \frac{\sigma}{\delta} = \sqrt{2} \tag{1.19}$$

1.3 Problem Formulation

This work considers the discrete-time linear dynamic system with scalar measurement $z(k)$ and scalar inputs $u(k)$ and $w(k)$,

$$x(k+1) = \Phi x(k) + \Gamma w(k) + \Lambda u(k) \tag{1.20a}$$

$$z(k) = \mathrm{H}x(k) + v(k) \tag{1.20b}$$

with the state vector $x \in \mathbb{R}^n$ for a finite integer n, $\Phi \in \mathbb{R}^{n\times n}$, $\Gamma \in \mathbb{R}^{n\times 1}$, $\Lambda \in \mathbb{R}^{n\times 1}$, and $\mathrm{H} \in \mathbb{R}^{1\times n}$. The inputs $w(k)$ and $v(k)$ are random variables described by Cauchy pdfs with zero median and scaling parameters $\beta > 0$ and $\gamma > 0$, respectively. Their probability distribution functions (pdfs) and their characteristic fucntions (CFs) are denoted f and ϕ, respectively, and are given by

$$f_W(w(k)) = \frac{\beta/\pi}{w(k)^2 + \beta^2} \quad \Rightarrow \quad \phi_W(\sigma) = e^{-\beta\,|\sigma|}, \tag{1.21a}$$

$$f_V(v(k)) = \frac{\gamma/\pi}{v(k)^2 + \gamma^2} \quad \Rightarrow \quad \phi_V(\sigma) = e^{-\gamma\,|\sigma|}, \tag{1.21b}$$

where σ is a scalar spectral variable. The initial condition is also assumed to be Cauchy distributed by a multivariate Cauchy pdf that is constructed as a product of scalar Cauchy pdfs with medians $\bar{x}_i$ and scaling parameters $\alpha_i > 0$ as

$$f_{X_1}(x(1)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{(x_i(1) - \bar{x}_i(1))^2 + \alpha_i^2} \quad \Rightarrow \quad \phi_{X_1}(\nu) = \prod_{i=1}^{n} e^{-\alpha_i\,|\nu_i| \,+\, \jmath \bar{x}_i(1)\nu_i} \tag{1.21c}$$

where $\nu \in \mathbb{R}^n$. For a multivariate pdf to be a Cauchy pdf, all of its scalar marginal distributions must themselves be Cauchy [20], as in (1.5). Clearly, every scalar marginal distribution of (1.21c) is a Cauchy pdf. Moreover, the independence of the scalar pdfs implies that the CF of the multivariate Cauchy pdf is simply the product of the CFs of the scalar Cauchy pdfs.

1.3.1 The Conditional Performance Index

In posing an optimal control problem for the model in (1.20), commonly used objective functions like the quadratic or the exponential of a quadratic cannot be used because the expectations required to evaluate those objective functions are infinite when the system noise inputs have heavy-tailed Cauchy pdfs. Therefore, one has to introduce a new, computable objective function. In our work in [12, 13, 18, 19] we suggest an objective function that resembles in its form the Cauchy pdf, and which also allows an analytical derivation of the controller. This general objective function is reminiscent to the choice of the objective function for the LEG [1, Chp. 10], which was constructed as a product of functions resembling the Gaussian pdf. The original motivation for the LEG objective function [22] was to consider these exponential functions as membership functions in fuzzy set theory, where the objective function was constructed as a product of these Gaussian-shaped membership functions. Similarly, a objective function constructed from products of Cauchy-shaped membership functions is proposed here for systems with Cauchy noises. This similarity allows us to compare the optimal controllers for linear systems with analogous performance index and different types of noise.

Consequently, the membership functions that penalize the state and control are chosen as rational functions resembling Cauchy pdfs and are expressed as

$$M_x(x(k)) = \prod_{r=1}^{n} \frac{\eta_{k,r}/\pi}{x_r^2(k) + \eta_{k,r}^2}, \quad M_u(u(k)) = \frac{\zeta_k/\pi}{u^2(k) + \zeta_k^2}, \tag{1.22}$$

where the membership functions are all centered at the origin, and each is a function of a single, scalar variable. Smaller values of η and ζ induce heavier weighings on the respective variables. These particular functions are chosen because they make the expectation with respect to the conditional pdf generated by Cauchy noise of the resulting performance index analytic in the control and measurement history.

1.3.2 State Decomposition

In our derivation of the estimator, we assume that the initial condition is centered at the origin, and the computational efficiency of the estimator will depend in part on this assumption. However, the theory is more general than this because there always exists a transformation of the pdf's variables that centers it on the origin. The stochastic system (1.20) can be decomposed into two systems, one driven by the known input $u(k)$ and one by the unknown input $w(k)$, by exploiting the linearity of the system. Let $\bar{x}(k)$ and $\bar{z}(k)$ be the part of the system driven by the control $u(k)$ only, and $\tilde{x}(k)$ and $\tilde{z}(k)$ be the part of the system driven by the process noise $w(k)$ only and contains all the underlying random variables. Then,

$$x(k) = \bar{x}(k) + \tilde{x}(k) \tag{1.23a}$$

$$z(k) = \bar{z}(k) + \tilde{z}(k). \tag{1.23b}$$

The controlled part of the system is described by

$$\bar{x}(k+1) = \Phi\bar{x}(k) + \Lambda u(k) \tag{1.24a}$$

$$\bar{z}(k) = \mathrm{H}\bar{x}(k) \tag{1.24b}$$

with initial condition $\bar{x}(0)$. The process noise driven part is given by

$$\tilde{x}(k+1) = \Phi\tilde{x}(k) + \Gamma w(k) \tag{1.25a}$$

$$\tilde{z}(k) = \mathrm{H}\tilde{x}(k) + v(k). \tag{1.25b}$$

The process and measurements noise pdfs were defined in (1.21a) and (1.21b), and the initial condition is given by (1.21c).

CHAPTER 2

Theory and the First Measurement Update

Yet now the sway of reason seemed irrefutably shaken, for this Cyclopean maze of squared, curved, and angled blocks had features which cut off all comfortable refuge. It was, very clearly, the blasphemous city of the mirage in stark, objective, and ineluctable reality.

H.P. Lovecraft, "In the Mountains of Madness"

This chapter derives the measurement update formula and the state propagation formula for the Cauchy estimator. In addition, we present the solution to the measurement update formula and the state propagation formula using CFs, and provide proofs of the existence of the first two conditional moments for the Cauchy estimator. In the process, we will derive the first measurement update for a general vector-state system, which is how we begin.

2.1 First Measurement Update

Begin with the first measurement update at $k = 1$ by taking a noisy measurement of the Cauchy distributed initial state as

$$z(1) = \mathrm{H}x(1) + v(1). \tag{2.1}$$

Here, $x(1)$ and $v(1)$ are the Cauchy random variables. For a scalar system, i.e. $x(1) \in \mathbb{R}^1$, the conditional mean estimator has been derived and is presented in [10]. For vector-state systems, an approach based on determining the *characteristic function* (CF) of the conditional pdf (cpdf) is used. Initial results for estimation of scalar systems using the characteristic function of the conditional pdf were developed and presented in [11, 23]. An algorithm for a multivariate Cauchy estimator is presented in [14, 15] and summarized here.

For $x(1) \in \mathbb{R}^n$, the CF of the initial state conditioned on the first measurement is given by the vector integral

$$\phi_{1|1}(\nu) = \int_{-\infty}^{\infty} f_{X_1|Z_1}(x(1)|z(1)) e^{j\nu^{\mathrm{T}}x(1)} dx(1). \tag{2.2}$$

The conditional pdf is computed from the joint distribution of $x(1)$ and $z(1)$ using Baye's Theorem [1] as

$$\begin{aligned} f_{X_1|Z_1}(x(1)|z(1)) &= \frac{f_{X_1,Z_1}(x(1), z(1))}{f_{Z_1}(z(1))} \\ &= \frac{f_{Z_1|X_1}(z(1)|x(1)) f_{X_1}(x(1))}{f_{Z_1}(z(1))} = \frac{f_V(z(1) - \mathrm{H}x(1)) f_{X_1}(x(1))}{f_{Z_1}(z(1))}. \end{aligned} \tag{2.3}$$

Then, (2.2) can be expressed as

$$\phi_{1|1}(\nu) = \frac{1}{f_{Z_1}(z(1))} \int_{-\infty}^{\infty} f_V(z(1) - \mathrm{H}x(1)) f_{X_1}(x(1))\, e^{j\nu^{\mathrm{T}}x(1)} dx(1) = \frac{\bar{\phi}_{1|1}(\nu)}{f_{Z_1}(z(1))}, \tag{2.4}$$

where $\bar{\phi}_{1|1}(\nu)$ is the characteristic function of the *unnormalized* cpdf (ucpdf). Note that, since $z(1)$ is known, $f_{Z_1}(z(1))$ is a constant; since $\phi_{1|1}|_{\nu=0} = 1$, then $\bar{\phi}_{1|1}|_{\nu=0} = f_{Z_1}(z(1))$.

Using the dual convolution property [24], $\bar{\phi}_{1|1}(\nu)$ in (2.4) can be expressed as n convolution integrals in the ν domain between the characteristic functions of $f_V(z(1) - \mathrm{H}x(1))$ and $f_{X_1}(x(1))$.

The CF of $f_{X_1}(x(1))$ is given in (1.21c). The CF of $f_V(z(1) - \mathrm{H}x(1))$ is denoted $\hat{\phi}_V(\nu)$ and is given by

$$\hat{\phi}_V(\nu) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_V(z(1) - \sum_{i=1}^{n} h_i x_i(1))\, e^{j \sum_{i=1}^{n} \nu_i x_i(1)} dx_1(1) \dots dx_n(1). \tag{2.5}$$

In order to proceed, we need some assumptions about the measurement vector $\mathrm{H} \triangleq [h_1 \dots h_n]$. The first is that at least one element of H is nonzero, i.e. there exists an i such that $h_i \neq 0$. This assumption is a prerequisite for observability of the state. The second assumption is that this nonzero element is h_n, which has no effect on generality [25].

However, note that if $h_i = 0$, then f_V loses dependence on the integration variable $x_i(1)$, and hence that part of f_{X_1} comes out of the integral. Then, the first two moments for that element of the state will not exist. This is an interesting effect of estimation where the unconditional distributions have undefined first and second moments. However, as long as the system is observable, then every element of the state will have a defined first and second moment in no more than n steps. For simplicity of the presentation of this algorithm, we will assume that every element of H is nonzero, i.e., $h_i \neq 0$ for all i.

To carry out the integration in (2.5), perform the change of variables

$$\xi = z(1) - \sum_{i=1}^{n} h_i x_i \tag{2.6}$$

in order to write

$$x_n(1) = \frac{1}{h_n}\left(z(1) - \xi - \sum_{i=1}^{n-1} h_i x_i(1)\right) \qquad \text{and} \qquad dx_n(1) = \frac{d\xi}{|h_n|}. \tag{2.7}$$

This allows us to manipulate (2.5) as

$$\begin{aligned}
\hat{\phi}_V(\nu) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_V(\xi)\, e^{j\left(\sum_{i=1}^{n-1} \nu_i x_i(1) + \nu_n \frac{z(1)-\xi-\sum_{i=1}^{n-1} h_i x_i(1)}{h_n}\right)} \frac{d\xi}{|h_n|} dx_1(1)\dots dx_{n-1}(1) \\
&= \frac{e^{j\frac{\nu_n}{h_n}z(1)}}{|h_n|}\left(\int_{-\infty}^{\infty} f_V(\xi)\, e^{-j\frac{\nu_n}{h_n}\xi} d\xi\right) \\
&\qquad \times \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{j\sum_{i=1}^{n-1}\left(\nu_i - \frac{h_1}{h_n}\nu_n\right)x_i(1)} dx_1(1)\dots dx_{n-1}(1)\right).
\end{aligned} \tag{2.8}$$

Here, the left parenthesis equals the CF of f_V if the spectral variable is $-\frac{\nu_n}{h_n}$. The right parenthesis is a product of Dirac delta functions, $\delta(\cdot)$, giving

$$\hat{\phi}_V(\nu) = \frac{e^{j\frac{\nu_n}{h_n}z(1)}}{|h_n|}\,\phi_V\left(-\frac{\nu_n}{h_n}\right)\prod_{i=1}^{n-1}\left[(2\pi)\delta\left(\nu_i - \frac{h_i}{h_n}\nu_n\right)\right]. \tag{2.9}$$

Using ϕ_{X_1} of (1.21c) and $\hat{\phi}_V$ of (2.9), we can express the first measurement update ucpdf's

CF using the dual convolution property [24] as

$$\bar{\phi}_{1|1}(\nu) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1}(x(1)) f_V(z(1) - \mathrm{H}x(1)) e^{j\nu^{\mathrm{T}}x(1)} dx$$

$$= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi_{X_1}(\nu - \sigma)\, \hat{\phi}_V(\sigma) d\sigma$$

$$= \frac{(2\pi)^{n-1}}{(2\pi)^n\, |h_n|} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi_{X_1}(\nu - \sigma)\, e^{j\frac{\sigma_n}{h_n} z(1)}\, \phi_V\left(-\frac{\sigma_n}{h_n}\right) \tag{2.10}$$

$$\times \prod_{i=1}^{n-1} \delta\left(\sigma_i - \frac{h_i}{h_n}\sigma_n\right) d\sigma_1 \ldots d\sigma_{n-1}\, d\sigma_n. \tag{2.11}$$

Integrating over $\sigma_1 \ldots \sigma_{n-1}$ is simple due to the delta functions in (2.11), and results in a single integral over the scalar σ_n:

$$\bar{\phi}_{1|1}(\nu) = \frac{1}{2\pi\, |h_n|} \int_{-\infty}^{\infty} \phi_{X_1}\left(\nu_1 - h_1 \tfrac{\sigma_n}{h_n}, \ldots, \nu_{n-1} - h_{n-1} \tfrac{\sigma_n}{h_n}, \nu_n - \sigma_n\right) e^{j\frac{\sigma_n}{h_n} z(1)} \phi_V\left(-\frac{\sigma_n}{h_n}\right) d\sigma_n$$

$$= \frac{1}{2\pi\, |h_n|} \int_{-\infty}^{\infty} \phi_{X_1}\left(\nu - \mathrm{H}^{\mathrm{T}} \frac{\sigma_n}{h_n}\right) e^{j\frac{\sigma_n}{h_n} z(1)} \phi_V\left(-\frac{\sigma_n}{h_n}\right) d\sigma_n. \tag{2.12}$$

Finally, we have a convolution integral involving the CF of the pdf of the state, and the CF of the measurement. This result indicates that the ucpdf's CF for a system of arbitrary (finite) order conditioned on a scalar measurement can be determined from this single, scalar convolution integral. A simple change of scalar variables $\sigma = \sigma_n / h_n$ and $d\sigma_n = d\sigma\, |h_n|$ gives the final form

$$\boxed{\bar{\phi}_{1|1}(\nu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{X_1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma)\, \phi_V(-\sigma)\, e^{jz(1)\sigma} d\sigma.} \tag{2.13}$$

Note that in the derivation of this results we did not use the fact that f_{X_1} and f_V were Cauchy, only that the element $h_n \neq 0$. This means that the update formula in (2.13) applies to any two $S\alpha S$ distributions, even different types of distribution (such as different α), using only their CFs. In subsequent updates, this formula is used by substituting the initial condition ϕ_{X_1} with the CF of the state propagated cpdf.

Applying (2.13) to our Cauchy estimator first update problem, we rewrite (2.13) using

(1.21b) and (1.21c) as

$$\bar{\phi}_{1|1}(\nu) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[-\left(\sum_{i=1}^{n+1} \rho_\ell \left|\mu_\ell - \sigma\right|\right) + jz(1)\sigma\right] d\sigma, \tag{2.14a}$$

where

$$\rho_\ell = \alpha_\ell \left|\epsilon_\ell \mathrm{H}^{\mathrm{T}}\right| \text{ for } \ell \in \{1,\dots,n\} \qquad \rho_{n+1} = \gamma, \tag{2.14b}$$

$$\mu_\ell = \frac{\epsilon_\ell \nu}{\epsilon_\ell \mathrm{H}^{\mathrm{T}}} \text{ for } \ell \in \{1,\dots,n\} \qquad \mu_{n+1} = 0, \tag{2.14c}$$

and ϵ_ℓ is the ℓ^{th} row of the n-dimensional identity matrix. Note that the μ_ℓ are scalars linear in ν, i.e. inner products of ν with given vectors. The next section will derive the solution to an integral of absolute values of the form in (2.14), which is also presented in [14, 15].

2.1.1 Solution to the First Measurement Update Convolution

This section will derive the solution to an integral of an exponential of absolute value terms, such as (2.14). The difficultly lies in that the absolute value function is not continuous at zero, e.g. whenever $\nu_\ell = h_\ell \sigma$ in (2.14). Since we can only use the fundamental theorem of calculus over domains where the function being integrated is continuous [26, Thm 6.21], we need to be able to divide the domain of integration into regions in which we can apply the theorem in order to evaluate the integral.

To do this, rewrite (2.14) as

$$\begin{aligned}\bar{\phi}_{1|1}(\nu) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[-\left(\sum_{i=1}^{n+1} \rho_\ell \left|\mu_\ell - \sigma\right|\right) + jz\sigma\right] d\sigma \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[-\left(\sum_{i=1}^{n+1} \rho_\ell \left(\mu_\ell - \sigma\right) \operatorname{sgn}\left(\mu_\ell - \sigma\right)\right) + jz\sigma\right] d\sigma \end{aligned} \tag{2.15}$$

In order to apply the fundamental theorem of calculus, we need to divide the domain up into sections in which the integrand is continuous [26]. This can be done by considering a fixed ν, and hence fixed μ_i, and defining constants s_i^ℓ as

$$\operatorname{sgn}\left(\mu_\ell - \sigma\right) \triangleq s_i^\ell = \begin{cases} \operatorname{sgn}\left(\mu_\ell - \mu_i\right) & \text{if } i \neq \ell \\ -1 & \text{if } i = \ell \end{cases} \tag{2.16}$$

From (2.16), for a given i, the discrete, two-index function s_i^ℓ is constant for all ℓ except for one switch at $\ell = i$, i.e. $s_i^\ell = -1\,\forall\,\ell \leq i$ and $s_i^\ell = +1\,\forall\,\ell > i$; this implies that $s_i^\ell = s_{i-1}^\ell\,\forall\,i \neq \ell$.

Substituting s_i^ℓ into (2.15) allows us to divide the integral into a sum of $n+1$ solvable integrals,

$$\begin{aligned}\bar{\phi}_{1|1}(\nu) &= \frac{1}{2\pi}\sum_{i=0}^{n+1}\left\{\int\limits_{\mu_i}^{\mu_{i+1}} \exp\left[-\sum_{\ell=1}^{n+1}\rho_\ell\left(\mu_\ell-\sigma\right)s_i^\ell + jz\sigma\right]d\sigma\right\}\\ &= \frac{1}{2\pi}\sum_{i=0}^{n}\left\{\frac{\exp\left[-\sum\limits_{\ell=1}^{n+1}\rho_\ell\left(\mu_\ell-\mu_{i+1}\right)s_i^\ell + jz\mu_{i+1}\right]}{jz+\sum\limits_{\ell=1}^{n+1}\rho_\ell s_i^\ell}\right\}\\ &\qquad\qquad -\frac{1}{2\pi}\sum_{i=0}^{n}\left\{\frac{\exp\left[-\sum\limits_{\ell=1}^{n+1}\rho_\ell\left(\mu_\ell-\mu_i\right)s_i^\ell + jz\mu_i\right]}{jz+\sum\limits_{\ell=1}^{n+1}\rho_\ell s_i^\ell}\right\}\end{aligned} \tag{2.17}$$

where $\mu_0 = -\infty$ and $\mu_{n+2} = \infty$. The first sum in (2.17) can be manipualted as follows

$$
\begin{aligned}
&\sum_{i=0}^{n+1} \left\{ \frac{\exp\left[-\sum\limits_{\ell=1}^{n+1} \rho_\ell \left(\mu_\ell - \mu_{i+1}\right) s_i^\ell + jz\mu_{i+1}\right]}{jz + \sum\limits_{\ell=1}^{n+1} \rho_\ell s_i^\ell} \right\} \\
&\qquad = \sum_{i=0}^{n} \left\{ \frac{\exp\left[-\sum\limits_{\ell=1}^{n+1} \rho_\ell \left(\mu_\ell - \mu_{i+1}\right) s_i^\ell + jz\mu_{i+1}\right]}{jz + \sum\limits_{\ell=1}^{n+1} \rho_\ell s_i^\ell} \right\} \\
&\qquad\qquad + \underbrace{\frac{\exp\left[-\sum\limits_{\ell=1}^{n+1} \rho_\ell (\mu_\ell - \mu_{n+2}) s_{n+1}^\ell + jz\mu_{n+2}\right]}{jz + \sum\limits_{\ell=1}^{n+1} \rho_\ell s_{n+1}^\ell}}_{=0 \text{ because } s_n^\ell = -1\, \forall\, \ell \leq n+1 \text{ and } \mu_{n+2} = \infty} \\
&\qquad = \sum_{i=1}^{n+1} \left\{ \frac{\exp\left[-\sum\limits_{\ell=1}^{n+1} \rho_\ell \left(\mu_\ell - \mu_i\right) s_{i-1}^\ell + jz\mu_i\right]}{jz + \sum\limits_{\ell=1}^{n+1} \rho_\ell s_{i-1}^\ell} \right\} \\
&\qquad = \sum_{i=1}^{n+1} \left\{ \frac{\exp\left[-\sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n+1} \rho_\ell \left(\mu_\ell - \mu_i\right) s_i^\ell + jz\mu_i\right]}{jz + \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n+1} \rho_\ell s_i^\ell} \right\},
\end{aligned}
\tag{2.18}
$$

where the last equality is because $s_{i-1}^i = 1$ and $s_{i-1}^\ell = s_i^\ell \ \forall i \neq \ell$, and the zero term $\rho_i(\mu_i - \mu_i)s_{i-1}^i = 0$ was dropped from the exponent.

The second sum in (2.17) can be manipualted in a similar fashion, as

$$\sum_{i=0}^{n+1}\left\{\frac{\exp\left[-\sum\limits_{\ell=1}^{n+1}\rho_\ell\left(\mu_\ell-\mu_i\right)s_i^\ell+jz\mu_i\right]}{jz+\sum\limits_{\ell=1}^{n+1}\rho_\ell s_i^\ell}\right\}$$

$$=\sum_{i=1}^{n+1}\left\{\frac{\exp\left[-\sum\limits_{\ell=1}^{n+1}\rho_\ell\left(\mu_\ell-\mu_i\right)s_i^\ell+jz\mu_i\right]}{jz+\sum\limits_{\ell=1}^{n+1}\rho_\ell s_i^\ell}\right\}+\underbrace{\frac{\exp\left[-\sum\limits_{\ell=1}^{n+1}\rho_\ell(\mu_\ell-\mu_0)s_0^\ell+jz\mu_0\right]}{jz+\sum\limits_{\ell=1}^{n+1}\rho_\ell s_0^\ell}}_{=0\text{ because } s_0^\ell=-1\,\forall\,\ell\geq1\text{ and }\mu_0=-\infty}$$

$$=\sum_{i=1}^{n+1}\left\{\frac{\exp\left[-\sum\limits_{\substack{\ell=1\\ \ell\neq i}}^{n+1}\rho_\ell\left(\mu_\ell-\mu_i\right)s_i^\ell+jz\mu_i\right]}{jz-\rho_i+\sum\limits_{\substack{\ell=1\\ \ell\neq i}}^{n+1}\rho_\ell s_i^\ell}\right\}. \tag{2.19}$$

where the last equality is true because $s_i^i=-1$ and the zero term $\rho_i(\mu_i-\mu_i)s_{i-1}^i=0$ was dropped from the exponent. Now, combine (2.18) and (2.19), as well as (2.16), to restate the solution to (2.14a) as

$$\bar{\phi}_{1|1}(\nu)=\frac{1}{2\pi}\sum_{i=1}^{n+1}\mathcal{G}_{1|1}^i(\nu)\cdot e^{\mathcal{E}_{1|1}^i(\nu)}=\frac{1}{2\pi}\sum_{i=1}^{n+1}\exp\left[-\sum_{\substack{\ell=1\\ \ell\neq i}}^{n+1}\rho_\ell\left|\mu_\ell-\mu_i\right|+jz(1)\mu_i\right]$$
$$\times\left\{\left[jz(1)+\rho_i+\sum_{\substack{\ell=1\\ \ell\neq i}}^{n+1}\rho_\ell\operatorname{sgn}\left(\mu_\ell-\mu_i\right)\right]^{-1}-\left[jz(1)-\rho_i+\sum_{\substack{\ell=1\\ \ell\neq i}}^{n+1}\rho_\ell\operatorname{sgn}\left(\mu_\ell-\mu_i\right)\right]^{-1}\right\}. \tag{2.20}$$

The beauty of this result is that, although we fixed ν in this integration, this procedure works for any $\nu\in\mathbb{R}^n$ such that $\mu_\ell-\mu_m\neq0$ (so that the sign functions are well defined), meaning that the solution is valid for any valid ν. More precisely, no matter what mapping is used between the sets $\{\mu_i\}_{i=1}^{n+1}$ and $\{\nu_i\}_{i=1}^{n+1}$, there exists a ν such that the desired ordering is valid, and upon finishing and substituting the sign functions back in for s_i^ℓ, the solution becomes valid for any valid ν.

The mean and variance for the first measurement were derived in closed form directly from (2.20) after much algebraic manipulation. It involves finding the normalizing constant $f_{Z_1}(z(1))$ and the first two partial derivatives of $\bar{\phi}_{1|1}(\nu)$. The normalizing constant is given by

$$f_{Z_1}(z(1)) = \frac{1}{\pi} \cdot \frac{\sum\limits_{\ell=1}^{n} \alpha_\ell \left|h_\ell\right| + \gamma}{z(1)^2 + \left(\sum\limits_{\ell=1}^{n} \alpha_\ell \left|h_\ell\right| + \gamma\right)^2}. \tag{2.21a}$$

Then, the minimum conditional-variance estimate is given by

$$\hat{x}(1) = z(1) \cdot \frac{\begin{bmatrix}\alpha_1 \operatorname{sgn}(h_1) & \cdots & \alpha_n \operatorname{sgn}(h_n)\end{bmatrix}^{\mathrm{T}}}{\sum\limits_{\ell=1}^{n} \alpha_\ell \left|h_\ell\right| + \gamma} \tag{2.21b}$$

and its conditional error variance is given by

$$\begin{aligned}\Xi(1) = &\left(1 + \frac{z(1)^2}{\left(\sum_{\ell=1}^{n} \alpha_\ell \left|h_\ell\right| + \gamma\right)^2}\right) \\ &\times \begin{bmatrix} \frac{\alpha_1}{\left|h_1\right|}\left(\sum\limits_{\substack{\ell=1\\ \ell\neq 1}}^{n} \alpha_\ell \left|h_\ell\right| + \gamma\right) & \cdots & -\alpha_1\alpha_n \operatorname{sgn}(h_1) \operatorname{sgn}(h_n) \\ \vdots & & \vdots \\ -\alpha_1\alpha_n \operatorname{sgn}(h_1) \operatorname{sgn}(h_n) & \cdots & \frac{\alpha_n}{\left|h_n\right|}\left(\sum\limits_{\substack{\ell=1\\ \ell\neq n}}^{n} \alpha_\ell \left|h_\ell\right| + \gamma\right)\end{bmatrix}\end{aligned} \tag{2.21c}$$

2.2 State Propagation with Process Noise

This section will derive the characteristic function of the cpdf propagated by process noise $w(k)$ as $x(k+1) = \Phi x(k) + \Gamma w(k)$, following a method that uses Proposition 2.17 in [1]. Assume that at an arbitrary time step k the state's pdf is given by $f_{X_k}(x(k))$, its CF is given by $\phi_{X_k}(\nu)$, and that the process noise CF is given as in (1.21a). Note that we are not assuming anything about f_{X_k} except that it has a well defined CF. The pdf of $x(k+1)$ is determined using the linear transformation of the joint distribution $f_{X_k,W}(x_k, w_k) = f_{X_k}(x_k) f_W(w_k)$

$$\begin{Bmatrix} x(k+1) \\ w(k) \end{Bmatrix} = \begin{bmatrix} \Phi & \Gamma \\ 0 & I \end{bmatrix} \begin{Bmatrix} x(k) \\ w(k) \end{Bmatrix} \Leftrightarrow \begin{Bmatrix} x(k) \\ w(k) \end{Bmatrix} = \underbrace{\begin{bmatrix} \Phi^{-1} & -\Phi^{-1}\Gamma \\ 0 & I \end{bmatrix}}_{\triangleq J} \begin{Bmatrix} x(k+1) \\ w(k) \end{Bmatrix} \tag{2.22}$$

Using Proposition 2.17 in [1], the joint distribution $f_{X_{k+1}}(x(k+1))$ is given by

$$f_{X_{k+1}}(x(k+1)) = |J|\, f_{X_k}\left(\Phi^{-1}x(k+1) - \Phi^{-1}\Gamma w(k)\right) f_W(w(k)) \tag{2.23}$$

where $|J| = |\Phi^{-1}|$ is the determinant of J from (2.22). Since f_{Xk+1} is the marginal distribution of $f_{X_{k+1}}(x(k+1))$, integrate over $w(k)$ as

$$f_{X_{k+1}}(x(k+1)) = \left|\Phi^{-1}\right| \int_{-\infty}^{\infty} f_{X_k}\left(\Phi^{-1}x(k+1) - \Phi^{-1}\Gamma w(k)\right) f_W(w(k)) dw(k). \tag{2.24}$$

Since what we seek is the CF of $f_{X_{k+1}}(x(k+1))$, what we actually want to evaluate is

$$\phi_{X_{k+1}}(\nu) = \\ \left|\Phi^{-1}\right| \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f_{X_k}\left(\Phi^{-1}x(k+1) - \Phi^{-1}\Gamma w(k)\right) f_W(w(k)) dw(k) \right] e^{j\nu^{\mathrm{T}}x(k+1)} dx(k+1) \tag{2.25}$$

Interchange the order of integration and use the change of variables $x(k+1) = \Phi x(k) + \Gamma w(k) \Rightarrow dx(k+1) = |\Phi|\, dx(k)$ to write

$$\phi_{X_{k+1}}(\nu) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f_{X_k}(x(k)) e^{j\nu^{\mathrm{T}}(\Phi x(k) + \Gamma w(k))} dx(k) \right] f_W(w(k)) dw(k) \tag{2.26}$$

$$= \left[\int_{-\infty}^{\infty} f_{X_k}(x(k)) e^{j\nu^{\mathrm{T}}\Phi x(k)} dx(k) \right] \cdot \left[\int_{-\infty}^{\infty} f_W(w(k)) e^{j\nu^{\mathrm{T}}\Gamma w(k)} dw(k) \right] \tag{2.27}$$

The contents of each of the terms in square brackets above are the known characteristic functions of the state, propagated by the dynamics, and the process noise so that the characteristic function of the propagated cpdf is given by

$$\boxed{\phi_{X_{k+1}}(\nu) = \phi_{X_k}(\Phi^{\mathrm{T}}\nu)\phi_W(\Gamma^{\mathrm{T}}\nu)} \tag{2.28}$$

Therefore, a time propagation with process noise is simply the product of the process noise CF, with scalar parameter $\Gamma\nu$, and the state pdf's CF, with the spectral variable transformed linearly by Φ. Note that we assumed that Φ is invertible, which is a safe assumption for state transition matrices obtained from discretizing continuous time linear system dynamics [25].

2.3 Existence of the Moments of the cpdf

This section will prove the existence of the first and second moments of a multivariate α-stable random vector (such that $1 \leq \alpha \leq 2$), conditioned on a scalar α-stable random variable. This means establishing the continuity of the cpdf's CF and its first two derivatives. Then we can take limits as ν goes to zero, thereby obtaining the desired moments [1].

First, we present conditions for the continuity of the cpdf's CF in Theorem 1, and then appy it to our Cauchy estimation problem in Corollary 1.

Theorem 1. *Let $x \in \mathbb{R}^n$, $z \in \mathbb{R}^1$ is a known constant, $\mathrm{H} \in \mathbb{R}^{1\times n}$ is a given row vector with nonzero elements, $f_X(x) : \mathbb{R}^n \to \mathbb{R}^1$ and $f_V(z - \mathrm{H}x) : \mathbb{R}^n \to \mathbb{R}^1$. Moreover, $f_X(x) \in L^1(\mathbb{R}^n)$ and $f_V \in L^1(\mathbb{R}^n)$, and both are continuous and bounded. Assume also that for all i $x_i f_X(x_i)$ and $x_i^2 f_X(x_i)$ are bounded in x_i by M_1 and M_2, respectively. Define the Fourier transform of the product of these functions as*

$$\bar{\phi}(\nu) = \int_{-\infty}^{\infty} f_X(x)\, f_V(z - \mathrm{H}x)\, e^{j x^{\mathrm{T}} \nu} dx \tag{2.29}$$

Then, $\bar{\phi}(\nu)$ and its first two derivatives with respect to ν are all continuous and bounded, and all three are $L^1(\mathbb{R}^n)$.

Proof. Since a product of two continuous L^1 functions is a continuous L^1 function [27, Thm 5.15a], and the Fourier transform of an L^1 function is uniformly continuous and bounded [24, Thm 8.1] [27, Thm 5.1], then $\bar{\phi}(\nu)$ is continuous and bounded. Also, since f_X and f_V are bounded, their product is bounded and thus their Fourier transform $\bar{\phi}(\nu)$ is in L^1 [24, Thm 8.10].

These theorems cited above are for one-dimensional Fourier transforms; however, since a multivariate Fourier transform is simply n scalar integrals over each variable, the continuity exists in all directions, and hence the whole multivariate transform is continuous [26, Thm 6.20]. This requires that both f_X and f_V depend on all the elements of x, which is why the elements of H must *all* be nonzero. This is the approach taken throughout the rest of the proof.

Consider the derivative of $\bar{\phi}(\nu)$ in (2.29). The derivative and integration operators can be interchanged by [26, Thm 9.42], so that we get

$$\frac{\partial}{\partial \nu}\bar{\phi}(\nu) = \int_{-\infty}^{\infty} \frac{\partial}{\partial \nu} f_X(x)\, f_V(z - \mathrm{H}x)\, e^{jx^{\mathrm{T}}\nu} dx$$
$$= \int_{-\infty}^{\infty} jx^{\mathrm{T}}\, f_X(x)\, f_V(z - \mathrm{H}x)\, e^{jx^{\mathrm{T}}\nu} dx \tag{2.30}$$

An integral of a vector is a vector of integrals [26, Def 6.23], so consider the i^{th} element's integrals:

$$\frac{\partial}{\partial \nu_i}\bar{\phi}(\nu) = j \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i\, f_X(x)\, f_V(z - \mathrm{H}x)\, e^{jx^{\mathrm{T}}\nu} dx_1 \ldots dx_n \tag{2.31}$$

Clearly, for transformations over x_j, $j \neq i$, the function $\frac{\partial}{\partial \nu_i}\bar{\phi}(\nu)$ is continuous because x_i comes out of the integral and what is left is the same as (2.29). For transformations over x_i, $\frac{\partial}{\partial \nu_i}\bar{\phi}(\nu)$ is continuous because $x_i\, f_X(x)\, f_V(z - \mathrm{H}x) \in L^1$. To see this, compute the L^1 norm, using the boundedness of $x_i f_X$, to get

$$\int_{-\infty}^{\infty} |x_i\, f_X(x)\, f_V(z - \mathrm{H}x)|\, dx \leq M_1 \int_{-\infty}^{\infty} f_V(z - \mathrm{H}x) dx < \infty \tag{2.32}$$

Since i was arbitrary, this is true for all i and so $\frac{\partial}{\partial \nu}\bar{\phi}(\nu) \in L^1$ is continuous and bounded.

Now, consider the second derivative and look at the element $\frac{\partial^2}{\partial \nu_i \nu_j}\bar{\phi}(\nu)$ for $i \neq j$,

$$\frac{\partial^2}{\partial \nu_i \nu_j}\bar{\phi}(\nu) = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i x_j\, f_X(x)\, f_V(z - \mathrm{H}x)\, e^{jx^{\mathrm{T}}\nu} dx \tag{2.33}$$

As before, since $i \neq j$, the function $\frac{\partial^2}{\partial \nu_i \nu_j}\bar{\phi}(\nu)$ is continuous because x_i comes out of the integrals and what is left is the same as (2.31). For transformations over x_i (i.e. $i = j$), the function $\frac{\partial^2}{\partial \nu_i^2}\bar{\phi}(\nu)$ is continuous and bounded because $x_i^2\, f_X(x) f_V(z - \mathrm{H}x) \in L^1$, by the same argument as before using the boundedness of $x_i^2\, f_X(x)$. Moreover, since f_V is also bounded, $\frac{\partial}{\partial \nu}\bar{\phi}(\nu)$ is in L^1.

Since i was arbitrary, this is true for all i and so $\frac{\partial}{\partial \nu}\left(\frac{\partial}{\partial \nu}\bar{\phi}(\nu)\right) \in L^1$ is continuous and bounded. □

Corollary 1. *For the n-state initial condition f_{X_1} defined in* (1.21c)*, and the scalar measurement noise measurement noise f_V defined in* (1.21b)*, the characteristic function of f_{X_0} conditioned on the measurement, denoted by $\phi_{1|1}(\nu)$ and defined as* (2.13) *divided by the constant $f_{Z_0}(z_0)$, is twice continuously differentiable. Moreover, the characteristic function and its first two derivatives are all $L^1(\mathbb{R}^n)$ and bounded.*

Proof. Since f_{X_1} and f_V are pdfs, they are both in L^1 with norm equal to 1. Since f_{X_1} is constructed as a product of independent, scalar Cauchy random variables as $f_{X_1}(x) = \prod_{i=1}^{n} f_{X_1,i}(x_i)$, we can check boundedness for any of these terms and it will be true for all of them. The boundedness of $x_i f_{X_1,i}(x_i)$ is due to the fact that it is continuous and vanishes at the origin and at $\pm\infty$; similarly, boundedness of $x_i^2 f_{X_1,i}(x_i)$ is due to the fact that it is continuous and vanishes at the origin and goes to one at $\pm\infty$ [26, Thms 2.41,4.15]. Hence, f_{X_1} and f_V satisfy the conditions of Theorem 1. Since the unnormalized cpdf's CF $\bar{\phi}_{1|1}$ is proportional to $\phi_{1|1}$ by a constant, the twice continuous differentiability of $\bar{\phi}_{1|1}$ implies the same for $\phi_{1|1}$. $\square$

Theorem 2. *Let $\bar{\phi}_{k|k-1}(\nu) \in L^1(\mathbb{R}^n)$ be twice continuously differentiable and bounded. Also, let the functions $\phi_W(\Gamma^{\mathrm{T}}\nu) \in L^1(\mathbb{R}^n)$, and $\phi_V(\sigma) \in L^1(\mathbb{R})$ be bounded and continuous. Moreover, $\frac{\partial}{\partial\nu}\phi_W(\Gamma^{\mathrm{T}}\nu) \in L^1(\mathbb{R}^n)$ and $\frac{\partial}{\partial\sigma}\phi_V(\sigma) \in L^1(\mathbb{R})$ are both bounded. Define $\bar{\phi}_{X_k|Z_k}$ as*

$$\bar{\phi}_{k|k}(\nu) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\nu - \Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\sigma)\,\phi_W(\Gamma^{\mathrm{T}}\nu)\,\phi_V(-\sigma)e^{jz(k)\sigma}d\sigma \tag{2.34}$$

where $\Phi \in \mathbb{R}^{n\times n}$ is an invertible matrix, $z_k \in \mathbb{R}$ is a known scalar constant, and $\mathrm{H} \in \mathbb{R}^{1\times n}$ and $\Gamma \in \mathbb{R}^n$ are vectors such that $\mathrm{H}\Gamma \neq 0$. Then, $\bar{\phi}_{k|k}(\nu)$ and its first two derivatives with respect to ν are all continuous and bounded, and all three are L^1.

Proof. Continuity of $\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\nu - \Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\sigma)$ with respect to ν is due to Φ being full rank.

Therefore, we can take the derivative with respect to the vector ν to get

$$\frac{\partial}{\partial\nu}\bar{\phi}_{k|k-1}(\nu) =$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\nu}\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\nu-\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\sigma)\right]\phi_W(\Gamma^{\mathrm{T}}\nu-\mathrm{H}\Gamma\sigma)\,\phi_V(-\sigma)\,e^{jz_k\sigma}d\sigma$$

$$+\frac{1}{2\pi}\int_{-\infty}^{\infty}\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\nu-\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\sigma)\left[\frac{\partial}{\partial\nu}\phi_W(\Gamma^{\mathrm{T}}\nu-\mathrm{H}\Gamma\sigma)\right]\phi_V(-\sigma)\,e^{jz_k\sigma}d\sigma \quad (2.35)$$

If $\mathrm{H}\Gamma = 0$ then the term in square brackets in the second line, $\frac{\partial}{\partial\nu}\phi_W(\Gamma^{\mathrm{T}}\nu-\mathrm{H}\Gamma\sigma)$, would lose dependence on σ and come out of the convolution integral. This would be problematic, because it is the very term that contains the discontinuities that need to be smoothed out in the convolution.

However, we have assumed that $\mathrm{H}\Gamma \neq 0$. There are two facts that will be used in this proof: the convolution of an L^1 function with an L^∞ function is continuous and bounded [27, Thm 5.13c] [28, Prp 8.8]; and the convolution of an L^1 function with another L^1 function is also L^1 [27, Thm 5.13a] [28, Prp 8.7]. Since ϕ_W and $\frac{\partial}{\partial\nu}\bar{\phi}_{k|k-1}$ are bounded, and ϕ_V is L^1, then the first line in (2.35) is continuous. Moreover, since ϕ_W and $\frac{\partial}{\partial\nu}\bar{\phi}_{X_k|Z_{k-1}}$ are also L^1, then the first line is L^1.

Then, since $\frac{\partial}{\partial\nu}\phi_W$ and $\bar{\phi}_{k|k-1}$ are L^1 and ϕ_V is bounded, the second line in (2.35) is also continuous. Moreover, since ϕ_V is L^1, then the second line is L^1. Since sums of continuous functions are continuous and sums of L^1 functions are L^1, then $\frac{\partial}{\partial\nu}\bar{\phi}_{k|k}(\nu) \in L^1$ is continuous and bounded.

Substitute into (2.35) the following change of variables

$$\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}-\sigma=\eta \quad \Rightarrow \quad \sigma=\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}-\eta\ ,\ d\sigma=-d\eta \quad (2.36)$$

to get

$$\frac{\partial}{\partial \nu}\bar{\phi}_{k|k}(\nu) =$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial \nu}\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\eta)\right]\phi_W(\mathrm{H}\Gamma\eta)\,\phi_V\left(-\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}+\eta\right)\,e^{jz_k\left(\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}-\eta\right)}d\eta$$

$$+\frac{1}{2\pi}\int_{-\infty}^{\infty}\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\eta)\left[\frac{\partial}{\partial \nu}\phi_W(\mathrm{H}\Gamma\eta)\right]\phi_V\left(-\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}+\eta\right)\,e^{jz_k\left(\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}-\eta\right)}d\eta \quad (2.37)$$

Take the second derivative, and interchange it with the integration operator to get

$$\frac{\partial}{\partial \nu}\left(\frac{\partial}{\partial \nu}\bar{\phi}_{k|k}(\nu)\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial \nu}\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\eta)\right]\phi_W(\mathrm{H}\Gamma\eta)$$

$$\frac{\partial}{\partial \nu}\left\{\phi_V\left(-\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}+\eta\right)\,e^{jz_k\left(\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}-\eta\right)}\right\}d\eta$$

$$+\frac{1}{2\pi}\int_{-\infty}^{\infty}\bar{\phi}_{k|k-1}(\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}\eta)\left[\frac{\partial}{\partial \nu}\phi_W(\mathrm{H}\Gamma\eta)\right]$$

$$\frac{\partial}{\partial \nu}\left\{\phi_V\left(-\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}+\eta\right)\,e^{jz_k\left(\frac{\Gamma^{\mathrm{T}}\nu}{\mathrm{H}\Gamma}-\eta\right)}\right\}d\eta \quad (2.38)$$

Here, the terms in brackets are bounded and the terms not in brackets are L^1, so $\frac{\partial}{\partial \nu}\left(\frac{\partial}{\partial \nu}\bar{\phi}_{k|k}(\nu)\right)$ is continuous and bounded. Moreover, since the terms in brackets are also L^1, then $\frac{\partial}{\partial \nu}\left(\frac{\partial}{\partial \nu}\bar{\phi}_{k|k}(\nu)\right) \in L^1$ is continuous and bounded. □

Corollary 2. *Let $\bar{\phi}_{k|k-1}(\nu) \in L^1$ be twice continuously differentiable and bounded, and let f_W and f_V be defined as in* (1.21a) *and* (1.21b)*, respectively. Then, $\bar{\phi}_{k|k}(\nu)$ defined as in* (2.34) *is twice continuously differentiable. Moreover, $\bar{\phi}_{k|k}(\nu)$ and its first two partial derivatives are L^1 and bounded.*

Proof. As before, ϕ_W and ϕ_V, and their first derivatives, are all bounded and in L^1. Moreover, ϕ_W and ϕ_V are themselves continuous. Since $\bar{\phi}_{k|k-1}(\nu)$ is assumed from Theorem 1 to have the properties required in the statement of Theorem 2, then $\bar{\phi}_{k|k}(\nu)$ defined as in (2.34) has the stated properties. □

In summary, the measurement update formula for the first update requires that at least one element of H be nonzero. This is a minimal requirement for observability in general. However, it is possible that only one state variable will have a defined mean and finite error variance. If the system is observable, then the entire state will acquire a defined mean and finite error variance in no more than n updates. For subsequent updates that occur after a state propagation step to have a defined mean and finite error variance, it is necessary that $\mathrm{H}\Gamma \neq 0$. Moreover, the state propagation step itself requires that Φ be invertible, which is a natural assumption if this discrete time transition matrix came from a continuous time dynamical system. The rest of this work will assume that all elements of H are nonzero, that $\mathrm{H}\Gamma \neq 0$, and that Φ is invertible.

CHAPTER 3

Two State Estimator

As I had expected, the canvas was warped, mouldy, and scabrous from dampness and neglect; but for all that I could trace the monstrous hints of evil cosmic outsideness that lurked all through the nameless scene's morbid content and perverted geometry.

H.P. Lovecraft, "Medusa's Coil"

In the previous section we considered a system with a state vector of general order n, and an algorithm for the general vector-state system is presented in [14, 15]. However, that algorithm suffers from a very aggressive growth in computational complexity with each new measurement update. For a second order system there are certain patterns and algebraic relationships that allow for significant reductions in numerical complexity and allow the estimator to run effectively over a large number of measurement updates [17]. There are two main aspects to this simplification: a way of expressing and indexing vectors that multiply (as inner products) the spectral vector ν in the exponential argument in (2.20); and a set of algebraic relations that can be used to simplify the coefficients of the exponential functions. Both of these aspects are addressed here for the first measurement update in (2.20) presented above. In Section 3.3 the second update is presented, indicating by induction the general measurement update and time propagation recursions given in Section 3.4. Sections 3.5 and 3.6 presents the second state propagation and the third measurement update, respectively, that were found in order to determine the general recursion.

3.1 First Update

Consider the arguments of the absolute value terms in (2.20). The μ_ℓ scalars are defined in (2.14c) as scaled inner products of ν with vectors we call *fundamental directions.* For the first measurement update, these fundamental directions are the rows of the $n \times n$ identity matrix. For the two-state system, the set of fundamental directions from the initial condition is $\mathcal{B}_{1|0} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. In the subscript of $\mathcal{B}_{1|0}$, the first element denotes the time step, and the second element denotes the number of measurements that have been processed.

Inner products are linear operations, and hence a difference of inner products with a given vector is also an inner product. This new inner product introduces a new fundamental direction. This new direction is the same for any real, 2×2 $\mathcal{B}_{1|0}$. Using a superscript on $\mathcal{B}^\ell_{1|0}$ to denote its ℓ^{th} row, apply this notation to the definitions in (2.14c) to get

$$\begin{aligned} \mu_\ell - \mu_m &= \frac{\mathcal{B}^\ell_{1|0}\nu}{\mathcal{B}^\ell_{1|0}\mathrm{H}^{\mathrm{T}}} - \frac{\mathcal{B}^m_{1|0}\nu}{\mathcal{B}^m_{1|0}\mathrm{H}^{\mathrm{T}}} = \frac{\mathrm{H}\mathcal{B}^{m\mathrm{T}}_{1|0}\mathcal{B}^\ell_{1|0}\nu - \mathrm{H}\mathcal{B}^{\ell\mathrm{T}}_{1|0}\mathcal{B}^m_{1|0}\nu}{\mathrm{H}\mathcal{B}^{m\mathrm{T}}_{1|0}\mathcal{B}^\ell_{1|0}\mathrm{H}^{\mathrm{T}}} \\ &= \frac{\mathrm{H}\left(\mathcal{B}^{m\mathrm{T}}_{1|0}\mathcal{B}^\ell_{1|0} - \mathcal{B}^{\ell\mathrm{T}}_{1|0}\mathcal{B}^m_{1|0}\right)\nu}{(\mathcal{B}^m_{1|0}\mathrm{H}^{\mathrm{T}})(\mathcal{B}^\ell_{1|0}\mathrm{H}^{\mathrm{T}})}. \end{aligned} \tag{3.1}$$

The key here is to recognize that the term in parenthesis in the numerator of (3.1) is a matrix minus its own transpose, which implies that it is antisymmetric, i.e., that $\mathcal{B}^{m\mathrm{T}}_{1|0}\mathcal{B}^\ell_0 - \mathcal{B}^{\ell\mathrm{T}}_{1|0}\mathcal{B}^m_{1|0} = \begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix}$ for some $c \in \mathbb{R}$. This constant c can be computed and pulled out of the matrix, which allows us to express any two-dimensional antisymmetric matrix as $c\mathrm{A}$ where $\mathrm{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and c can be verified to be

$$c = -\mathcal{B}^\ell_{1|0}\mathrm{A}\mathcal{B}^{m\mathrm{T}}_{1|0}. \tag{3.2}$$

Hence, we can write (3.1) as

$$\mu_\ell - \mu_m = \left(\frac{-\mathcal{B}^\ell_{1|0}\mathrm{A}\mathcal{B}^{m\mathrm{T}}_{1|0}}{(\mathcal{B}^m_{1|0}\mathrm{H}^{\mathrm{T}})(\mathcal{B}^\ell_{1|0}\mathrm{H}^{\mathrm{T}})}\right) \cdot \mathrm{HA}\nu. \tag{3.3}$$

This produces the new fundamental direction is HA, scaled by the term in parenthesis. For the relationship in (3.3) to hold, $\mu_\ell \neq \mu_m$ and neither μ_ℓ nor μ_m can equal zero. In (2.14c) we also defined an extra constant $\mu_3 = 0$. This implies that the old fundamental directions are retained in the measurement updated cpdf's CF for terms in the exponent of (2.20) that involve μ_3. Therefore, the set of fundamental directions for the first measurement update, denoted by $\mathcal{B}_{1|1}$, is given by

$$\mathcal{B}_{1|1} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \mathrm{HA} \end{bmatrix}. \tag{3.4}$$

In (2.20), we use (3.3) to express $\mu_2 - \mu_1$ as

$$\mu_2 - \mu_1 = \frac{\epsilon_2 \nu}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} - \frac{\epsilon_1 \nu}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} = \frac{-(\epsilon_2 \mathrm{A} \epsilon_1^{\mathrm{T}}) \mathrm{HA} \nu}{(\epsilon_1 \mathrm{H}^{\mathrm{T}})(\epsilon_2 \mathrm{H}^{\mathrm{T}})} = \frac{\mathrm{HA}\nu}{(\epsilon_1 \mathrm{H}^{\mathrm{T}})(\epsilon_2 \mathrm{H}^{\mathrm{T}})}. \tag{3.5}$$

Using (3.5) and the definitions for ρ_ℓ from (2.14b) in (2.20) yields

$$\begin{aligned}
\bar{\phi}_{1|1}(\nu) &= \mathcal{G}^1_{1|1}(\nu) \exp\left(-\alpha_2 \left|\epsilon_2 \mathrm{H}^{\mathrm{T}}\right| \left| \frac{\mathrm{HA}\nu}{(\epsilon_1 \mathrm{H}^{\mathrm{T}})(\epsilon_2 \mathrm{H}^{\mathrm{T}})} \right| - \gamma \left| \frac{\epsilon_1 \nu}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \right| + jz(1) \frac{\epsilon_1 \nu}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \right) \\
&\quad + \mathcal{G}^2_{1|1}(\nu) \exp\left(-\alpha_1 \left|\epsilon_1 \mathrm{H}^{\mathrm{T}}\right| \left| \frac{\mathrm{HA}\nu}{(\epsilon_1 \mathrm{H}^{\mathrm{T}})(\epsilon_2 \mathrm{H}^{\mathrm{T}})} \right| - \gamma \left| \frac{\epsilon_2 \nu}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \right| + jz(1) \frac{\epsilon_2 \nu}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \right) \\
&\quad + \mathcal{G}^3_{1|1}(\nu) \exp\left(-\alpha_1 \left|\epsilon_1 \mathrm{H}^{\mathrm{T}}\right| \left| \frac{\epsilon_1 \nu}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \right| - \alpha_2 \left|\epsilon_2 \mathrm{H}^{\mathrm{T}}\right| \left| \frac{\epsilon_2 \nu}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \right| \right) \\
&= \mathcal{G}^1_{1|1}(\nu) \exp\left(-\frac{\gamma}{\left|\epsilon_1 \mathrm{H}^{\mathrm{T}}\right|} \left|\epsilon_1 \nu\right| - \frac{\alpha_2}{\left|\epsilon_1 \mathrm{H}^{\mathrm{T}}\right|} \left|\mathrm{HA}\nu\right| + j \frac{z(1)}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \epsilon_1 \nu \right) \\
&\quad + \mathcal{G}^2_{1|1}(\nu) \exp\left(-\frac{\gamma}{\left|\epsilon_2 \mathrm{H}^{\mathrm{T}}\right|} \left|\epsilon_2 \nu\right| - \frac{\alpha_1}{\left|\epsilon_2 \mathrm{H}^{\mathrm{T}}\right|} \left|\mathrm{HA}\nu\right| + j \frac{z(1)}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \epsilon_2 \nu \right) \\
&\quad + \mathcal{G}^3_{1|1}(\nu) \exp\left(-\alpha_1 \left|\epsilon_1 \nu\right| - \alpha_2 \left|\epsilon_2 \nu\right| \right).
\end{aligned} \tag{3.6}$$

Notice that each of the three terms involves only two of the three fundamental directions.

The efficiency of the proposed two-state estimator is achieved by both keeping track of which directions are used in each term, as well as the scalar coefficients that multiply the absolute value functions and the scalar coefficients in the imaginary part of the argument of the exponential. The most important of these is an array of integers where the elements of each row correspond to the rows of $\mathcal{B}_{1|1}$ that appear in the exponential argument. This array is denoted by $M(1|1)$.

The exponential argument of the term corresponding to $i = 3$ in (3.6) is exactly the same as the initial condition. In fact, the only difference between the initial condition and this term is the coefficient $\mathcal{G}^3_{1|1}$. This is due to the cancellations that occur in general in (3.6), so that one term produced from the convolution will always have the same exponential argument with a different coefficient function. This term is referred to as the *old term*. It will be shown later that the old exponential arguments persist across measurement updates, and therefore it is useful to order the terms with the old fundamental directions first, so that the last term in (3.6) moves to the top. Consequently, $M(1|1)$ is constructed as

$$M(1|1) = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix}. \tag{3.7}$$

The other two terms, corresponding to $i = 1$ and $i = 2$ in (3.6), are called the *intermediate new terms*; we say *intermediate* because in all subsequent measurement updates, as will be shown, they will combine with other terms with the same exponential argument, and *new* because they involve the new fundamental directions just generated during the measurement update.

Following this new ordering, we define two additional arrays $P(1|1)$ and $Z(1|1)$ whose elements correspond to the coefficients of the absolute value functions in the exponents and the coefficients in the imaginary part, respectively; hence, the i^{th} rows of $P(1|1)$ and $Z(1|1)$ are related to the i^{th} term in $\bar{\phi}_{1|1}$. For the measurement updated ordering of terms used to define $M(1|1)$ in (3.7), these arrays are given by

$$P(1|1) = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \dfrac{\gamma}{|\epsilon_1 \mathrm{H}^{\mathrm{T}}|} & \dfrac{\alpha_2}{|\epsilon_1 \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\epsilon_2 \mathrm{H}^{\mathrm{T}}|} & \dfrac{\alpha_1}{|\epsilon_2 \mathrm{H}^{\mathrm{T}}|} \end{bmatrix}, \quad Z(1|1) = \begin{bmatrix} 0 & 0 \\ \dfrac{z(1)}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} & 0 \\ \dfrac{z(1)}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} & 0 \end{bmatrix}. \tag{3.8}$$

They have the same dimension as $M(1|1)$, so that an element of $P(1|1)$ or $Z(1|1)$ goes with the fundamental direction indexed by the corresponding element of $M(1|1)$.

Finally, define a vector array of integers $L_{1|1}$ with as many elements as rows of $M(1|1)$. Each element of $L_{1|1}$ indicates the number of fundamental directions in that term, i.e., the width of the corresponding row of $M(1|1)$. For the first measurement update, $L_{1|1}$ is given by

$$L_{1|1} = \begin{bmatrix} 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}}. \tag{3.9}$$

This final array is unnecessary in the first measurement update, but will become essential later when different terms involve different numbers of fundamental directions. Finally, define the number of terms as $N_{1|1}$. Clearly, $N_{1|1} = 3$, the same as the number of rows of $M(1|1)$, $P(1|1)$, $Z(1|1)$, and $L_{1|1}$.

Using (3.4) (3.7) (3.8) (3.9) and the definition for $N_{1|1} = 3$, the exponential argument for the first measurement update ucpdf's CF given in (2.20) can be expressed as

$$\mathcal{E}^i_{1|1}(\nu) = -\sum_{\ell=1}^{L^i_{1|1}} P^\ell_i \left| \mathcal{B}^{M^\ell_i}_{1|1} \nu \right| + j \left(\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{1|1} \right) \nu \quad , i \in \{1, \ldots, N_{1|1}\}. \tag{3.10}$$

3.1.1 Polynomial Coefficient

Using (3.5), (2.14b), and (2.14c), the bracketed terms $\mathcal{G}^i_{1|1}$ in (2.20) are of the same form, and each involves two sign functions, denoted s_1 and s_2, with the same arguments as the absolute values in the corresponding exponential parts. Functions of this form can be reduced to a four parameter polynomial of these sign functions as $a_i + b_i s_1 s_2 + j c_i s_1 + j d_i s_2$. These relationships are given in Appendix A; for $k = 1$ in particular, Result 3 in Appendix A can be used to obtain the four parameter polynomial by assuming that $a^m_i = 1$ and $b^m_i = c^m_i = d^m_i = 0$. The sign functions involve the same fundamental directions as the exponential argument, so denote a final array $G(1|1)$ with the same number of rows as $M(1|1)$ and width four, and let each row contain the parameters for the polynomial, as

$$G(1|1) = \begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{bmatrix}. \tag{3.11}$$

Note that the same reordering used to form $M(1|1)$ must be used here, so that the top row contains the parameters for the old term's coefficient.

Using the parameters in (3.11), along with (3.4) and (3.7), the coefficient function $\mathcal{G}^i_{1|1}$ for the first update ucpdf's CF given in (2.20) can be expressed as a polynomial

$$\mathcal{G}^i_{1|1}(\nu) = a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{1|1}\nu\right) \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{1|1}\nu\right) + j c_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{1|1}\nu\right) + j d_i \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{1|1}\nu\right). \tag{3.12}$$

Remark 1. *In the notation, for P, Z, and M the subscripts denote the row (i.e. which term in the sum) and the superscripts denote the column (i.e., where in the exponential argument it appears). Only $\mathcal{B}$, L, and N retain the time dependence in their subscripts. The superscript for $\mathcal{B}$ denotes the row, i.e., the fundamental direction, that is indexed by the integer M_i^ℓ. It is assumed that P , Z, and M inherit their time index from the associated fundamental directions from $\mathcal{B}$.*

3.2 First Time Propagation

In this section the CF of the cpdf is propagated using the dynamics given in (1.20a). Applying the time propagation equation (2.28) to the characteristic function of the first measurement update given in (2.20) yields

$$\phi_{2|1}(\nu) = \phi_{1|1}(\Phi^{\mathrm{T}}\nu)\phi_W(\Gamma^{\mathrm{T}}\nu) = \frac{\bar{\phi}_{1|1}(\Phi^{\mathrm{T}}\nu)\phi_W(\Gamma^{\mathrm{T}}\nu)}{f_{Z_1}(z(1))} = \frac{\bar{\phi}_{2|1}(\nu)}{f_{Z_1}(z(1))}. \tag{3.13}$$

Then, the ucpdf's characteristic function for the propagated state is given by

$$\begin{aligned}\bar{\phi}_{2|1}(\nu) = \bar{\phi}_{1|1}(\Phi^{\mathrm{T}}\nu)\phi_W(\Gamma^{\mathrm{T}}\nu) = \bar{\phi}_{1|1}(\Phi^{\mathrm{T}}\nu)\cdot e^{-\beta\left|\Gamma^{\mathrm{T}}\nu\right|}&\\ = \sum_{i=1}^{N_{1|1}} \mathcal{G}_{1|1}^i(\Phi^{\mathrm{T}}\nu) e^{\mathcal{E}_{1|1}^i(\Phi^{\mathrm{T}}\nu) - \beta\left|\Gamma^{\mathrm{T}}\nu\right|}&\\ = \sum_{i=1}^{N_{1|1}} \mathcal{G}_{1|1}^i(\Phi^{\mathrm{T}}\nu)\cdot \exp\left(-\sum_{\ell=1}^{L_{1|1}^i} P_i^\ell \left|\mathcal{B}_{1|1}^{M_i^\ell}\Phi^{\mathrm{T}}\nu\right| - \beta\left|\Gamma^{\mathrm{T}}\nu\right| + j\left(\sum_{\ell=1}^{2} Z_i^\ell \mathcal{B}_{1|1}^{M_i^\ell}\right)\Phi^{\mathrm{T}}\nu\right).\end{aligned} \tag{3.14}$$

The changes in $\bar{\phi}_{2|1}(\nu)$ from $\bar{\phi}_{1|1}(\nu)$ are a new element in the sum of the absolute value terms, and a linear transformation on ν. The time propagation step adds no new terms to the sum, hence $N_{2|1} = N_{1|1}$. Moreover, since the process noise has a zero median, the time propagation has no effect on the complex part of the exponential argument, so that $Z(2|1) = Z(1|1)$. Finally, there is no effect on the parameters of the coefficient functions $\mathcal{G}$, so that $G(2|1) = G(1|1)$. The time propagated ucpdf's characteristic function can be restated in the same form as $\bar{\phi}_{1|1}$, i.e.,

$$\bar{\phi}_{2|1}(\nu) = \sum_{i=1}^{N_{2|1}} \mathcal{G}_{2|1}^i(\nu)\cdot e^{\mathcal{E}_{2|1}^i(\nu)}, \tag{3.15a}$$

where

$$\mathcal{G}^i_{2|1}(\nu) = a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1}\nu\right) \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{2|1}\nu\right) + jc_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1}\nu\right) + jd_i \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{2|1}\nu\right), \quad (3.15\text{b})$$

$$\mathcal{E}^i_{2|1}(\nu) = -\sum_{\ell=1}^{L^i_{2|1}} P^\ell_i \left|\mathcal{B}^{M^\ell_i}_{2|1}\nu\right| + j\left(\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{2|1}\right)\nu, \quad (3.15\text{c})$$

$Z(2|1) = Z(1|1)$,

$$P(2|1) = \begin{bmatrix} & \beta \\ P(1|1) & \beta \\ & \beta \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \beta \\ \frac{\gamma}{|\epsilon_1 \mathrm{H}^\mathrm{T}|} & \frac{\alpha_2}{|\epsilon_1 \mathrm{H}^\mathrm{T}|} & \beta \\ \frac{\gamma}{|\epsilon_2 \mathrm{H}^\mathrm{T}|} & \frac{\alpha_1}{|\epsilon_2 \mathrm{H}^\mathrm{T}|} & \beta \end{bmatrix}, \quad \mathcal{B}_{2|1} = \begin{bmatrix} \mathcal{B}_{1|1}\Phi^\mathrm{T} \\ \Gamma^\mathrm{T} \end{bmatrix} = \begin{bmatrix} \epsilon_1\Phi^\mathrm{T} \\ \epsilon_2\Phi^\mathrm{T} \\ \mathrm{HA}\Phi^\mathrm{T} \\ \Gamma^\mathrm{T} \end{bmatrix}, \quad (3.15\text{d})$$

$$M(2|1) = \begin{bmatrix} & 4 \\ M(1|1) & 4 \\ & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}, \quad L_{2|1} = L_{1|1} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

The process noise introduces a third absolute value term, with a new fundamental direction, into all of the CF's exponential arguments. The next section deals with performing the next measurement update to this propagated structure.

3.3 Second Measurement Update

The second measurement update process involves finding the characteristic function of the unnormalized conditional pdf, i.e., $\bar{\phi}_{2|2}$. The formula used for this second measurement update is the same as the first update, given in (2.13), and is applied to (3.15):

$$\begin{aligned} \bar{\phi}_{2|2}(\nu) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\phi}_{2|1}(\nu - \mathrm{H}^\mathrm{T}\sigma)\phi_V(-\sigma)e^{j\sigma z(2)}d\sigma \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\phi}_{2|1}(\nu - \mathrm{H}^\mathrm{T}\sigma)e^{-\gamma|\sigma|}e^{j\sigma z(2)}d\sigma. \end{aligned} \quad (3.16)$$

Since $\bar{\phi}_{2|1}$ is a sum of $N_{2|1} = 3$ terms, the measurement update process is to solve the convolution integral $N_{2|1}$ times, once for each term. Substituting (3.15a) into (3.16), we get

$$\bar{\phi}_{2|2}(\nu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{i=1}^{N_{2|1}} \mathcal{G}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) \cdot \exp\left[\mathcal{E}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) - \gamma\,|-\sigma| + j\sigma z(2)\right] d\sigma. \tag{3.17}$$

Interchanging the integration and summation operations produces $N_{2|1}$ convolution integrals. As with the first measurement update, each convolution will produce an old term as well as new terms, called *intermediate terms*. Since many of these terms have the same exponential arguments, they can be combined to reduce the total number of terms in the CF's sum.

Therefore, we begin with an arbitrary i^{th} convolution in (3.17). Using (3.15c) we define

$$\begin{aligned} I_i &= \int_{-\infty}^{\infty} \mathcal{G}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) \cdot \exp\left[\mathcal{E}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) - \gamma\,|\sigma| + j\sigma z(2)\right] \\ &= \int_{-\infty}^{\infty} \mathcal{G}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) \cdot \exp\Bigg[-\sum_{\ell=1}^{L^i_{2|1}} P^\ell_i \left|\mathcal{B}^{M^\ell_i}_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma)\right| - \gamma\,|\sigma| \\ &\qquad + j\left(\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{2|1}\right)(\nu - \mathrm{H}^{\mathrm{T}}\sigma) + j\sigma z(2)\Bigg] d\sigma. \end{aligned} \tag{3.18}$$

In order to solve (3.18), we need to rewrite it in a manner similar to (2.15) using ρ_ℓ and μ_ℓ substitutions. Begin as in the first measurement update by defining constants μ_ℓ obtained by expressing

$$\mathcal{B}^{M^\ell_i}_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) = \mathcal{B}^{M^\ell_i}_{2|1}\mathrm{H}^{\mathrm{T}}(\mu_\ell - \sigma), \tag{3.19a}$$

and thus

$$\mu_\ell = \frac{\mathcal{B}^{M^\ell_i}_{2|1}\nu}{\mathcal{B}^{M^\ell_i}_{2|1}\mathrm{H}^{\mathrm{T}}}. \tag{3.19b}$$

Next, we rewrite the argument of the exponential of (3.18) in terms of these μ_ℓs and σ as

$$\begin{aligned} &-\sum_{\ell=1}^{L^i_{2|1}} P^\ell_i \left|\mathcal{B}^{M^\ell_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right| \cdot |\mu_\ell - \sigma| - \gamma\,|-\sigma| \\ &\qquad + j\left(Z^1_i \mathcal{B}^{M^1_i}_{2|1} + Z^2_i \mathcal{B}^{M^2_i}_{2|1}\right)\nu + j\left(z(2) - Z^1_i \mathcal{B}^{M^1_i}_{2|1}\mathrm{H}^{\mathrm{T}} - Z^2_i \mathcal{B}^{M^2_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right)\sigma \\ &= j\left(Z^1_i \mathcal{B}^{M^1_i}_{2|1} + Z^2_i \mathcal{B}^{M^2_i}_{2|1}\right)\nu - \sum_{\ell=1}^{L^i_{2|1}+1=4} \rho_\ell\,|\mu_\ell - \sigma| + j\theta^i_2\sigma, \end{aligned} \tag{3.20a}$$

where the complex part multiplying ν does not depend on σ and comes out of the convolution. The parameter definitions are $\mu_{L^i_{2|1}+1} = \mu_4 = 0$,

$$\theta^i_2 = z(2) - Z^1_i \mathcal{B}^{M^1_i}_{2|1} \mathrm{H}^{\mathrm{T}} - Z^2_i \mathcal{B}^{M^2_i}_{2|1} \mathrm{H}^{\mathrm{T}}, \tag{3.20b}$$

and the $L^i_{2|1} + 1 = 3 + 1 = 4$ constants called ρ_ℓ are given by

$$\rho_\ell = \begin{cases} P^\ell_i \left| \mathcal{B}^{M^\ell_i}_{2|1} \mathrm{H}^{\mathrm{T}} \right| & \ell \in \{1, \ldots, L^i_{2|1} = 3\}, i \in \{1, \ldots, N_{2|1} = 3\} \\ \gamma & \ell = L^i_{2|1} + 1 = 4. \end{cases} \tag{3.20c}$$

This is the same procedure used in the first measurement update except that, due to the new absolute value term in the exponential argument introduced in the time propagation step, there are now four ρ_ℓ constants instead of three. Moreover, there are three fundamental directions in the exponential argument, instead of just two.

Let's turn our attention to the coefficients of the exponents. The coefficient functions for the three terms in $\bar{\phi}_{2|1}$ are all of the form

$$\mathcal{G}^i_{2|1}(\nu) = a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1} \nu\right) \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{2|1} \nu\right) + jc_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1} \nu\right) + jd_i \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{2|1} \nu\right). \tag{3.21}$$

Hence the same manipulation used in (3.19b) to form the μs can be used here to get

$$\begin{aligned} &\mathcal{G}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) \\ &\quad = a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1} \mathrm{H}^{\mathrm{T}} \cdot \mathcal{B}^{M^2_i}_{2|1} \mathrm{H}^{\mathrm{T}}\right) \left(\frac{1}{2}\Big(\operatorname{sgn}(\mu_1 - \sigma) + \operatorname{sgn}(\mu_2 - \sigma)\Big)^2 - 1\right) \\ &\qquad + jc_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1} \mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}(\mu_1 - \sigma) + jd_i \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{2|1} \mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}(\mu_2 - \sigma), \end{aligned} \tag{3.22}$$

where in the real part we use the identity

$$\operatorname{sgn}(\mu_1 - \sigma) \operatorname{sgn}(\mu_2 - \sigma) = \frac{1}{2}\left(\operatorname{sgn}(\mu_1 - \sigma) + \operatorname{sgn}(\mu_2 - \sigma)\right)^2 - 1. \tag{3.23}$$

In order to use the same integration method as in the first measurement update, it is necessary for the coefficient of the exponential term to be constant within each subdomain of integration $(\mu_\ell, \mu_{\ell+1})$. Since these coefficients are polynomials of sums of sign functions, they are clearly constant in these regions. Now define the following constants,

$$\begin{aligned} \bar{a}_i &= a_i, & \bar{b}_i &= b_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1} \mathrm{H}^{\mathrm{T}} \cdot \mathcal{B}^{M^2_i}_{2|1} \mathrm{H}^{\mathrm{T}}\right), \\ \bar{\rho}_1 &= c_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{2|1} \mathrm{H}^{\mathrm{T}}\right), & \bar{\rho}_2 &= d_i \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{2|1} \mathrm{H}^{\mathrm{T}}\right), \\ \tilde{\rho}_1 &= 1, & \tilde{\rho}_2 &= 1. \end{aligned} \tag{3.24}$$

Our integration method also requires that the coefficient involve sums over the same sign functions as the exponential argument, which has two more absolute value functions than the coefficient has sign functions. To do this, introduce two more sign functions into the sum in (3.22) by multiplying them by constants $\tilde{\rho}_3 = \tilde{\rho}_4 = 0$ and $\bar{\rho}_3 = \bar{\rho}_4 = 0$. Using these, the coefficient in the convolution can be written with sums over the same μ_ℓ as in (3.20a):

$$\begin{aligned} \mathcal{G}^i_{2|1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) &= \bar{a}_i + \bar{b}_i \left(\frac{1}{2} \left(\sum_{\ell=1}^{4} \tilde{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right)^2 - 1 \right) + j \left(\sum_{\ell=1}^{4} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right) \\ &= \mathcal{G}^i_{2|1} \left(\sum_{\ell=1}^{4} \tilde{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \sum_{\ell=1}^{4} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right) := \mathcal{G}^i_{2|1} \left(\sum_{\ell=1}^{4} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right). \end{aligned} \tag{3.25}$$

Since the only difference between the summations in (3.25) are the scalar $\tilde{\rho}_\ell$ and $\bar{\rho}_\ell$ constants, we can simplify the notation and use a shorthand $\check{\rho}_\ell$ in (3.25) to represent all the summations. This notation will become more useful in subsequent measurement updates as the coefficient polynomials become more complex.

Using these substitutions, each of the $i \in \{1, \ldots, N_{2|1}\}$ convolution integrals in (3.17) can be written as

$$I_i = \int_{-\infty}^{\infty} \mathcal{G}^i_{2|1} \left(\sum_{\ell=1}^{4} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right) \cdot \exp\left(- \sum_{\ell=1}^{4} \rho_\ell |\mu_\ell - \sigma| + j\theta_2^i \sigma \right) d\sigma. \tag{3.26}$$

The solution to the i^{th} convolution integral for the second measurement update from [14,15] is

$$\begin{aligned} I_i = \sum_{m=1}^{L^i_{2|1}+1} \exp\left(- \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{2|1}+1} \rho_\ell |\mu_\ell - \mu_m| + j\theta_2^i \mu_m \right) \\ \times \left\{ \frac{\mathcal{G}^i_{2|1}\left(+\check{\rho}_m + \sum\limits_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{2|1}+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m) \right)}{j\theta_2^i + \rho_m + \sum\limits_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{2|1}+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)} - \frac{\mathcal{G}^i_{2|1}\left(-\check{\rho}_m + \sum\limits_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{2|1}+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m) \right)}{j\theta_2^i - \rho_m + \sum\limits_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{2|1}+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)} \right\}. \end{aligned} \tag{3.27}$$

The convolution produces $L^i_{2|1} + 1 = 4$ terms, indexed here by m, for each of the i terms in the sum in $\bar{\phi}_{2|1}$. The first three terms are called the *intermediate new terms* because some of the terms from different convolutions have the same exponential arguments and can be combined by summing their coefficient functions together, as will be shown. The last term produced by the convolution that corresponds to $m = 4$ is the *old term* because, since $\mu_4 = 0$, it has the same exponential argument as the i^{th} term in $\bar{\phi}_{2|1}$ that was convolved. Simplifications of these terms and their coefficient functions are addressed next.

3.3.1 Second Measurement Update - New Intermediate Terms

The terms corresponding to $m \in \{1, 2, 3\}$ are called the *intermediate new terms*. These always involve only two fundamental directions: one is from $\mathcal{B}_{2|1}$ and the other is the new direction HA as in the first measurement update. The formation of the new fundamental direction is what recovers the structure involving two fundamental directions, as will be shown here.

Until the intermediate new terms are combined into the final structure for the second measurement update, it is useful to denote the exponential argument in terms of both indices i and m. Hence, for the intermediate terms, denote the coefficient and exponential arguments as $\mathcal{G}_{2|2}^{i,m}$ and $\mathcal{E}_{2|2}^{i,m}$, respectively.

For the i^{th} term, use the definitions in (3.19b), (3.20b), and (3.20c) to rewrite the exponential argument of (3.27) as

$$\begin{aligned}
\mathcal{E}_{2|2}^{i,m} = -\gamma \left| \frac{\mathcal{B}_{2|1}^{M_i^m} \nu}{\mathcal{B}_{2|1}^{M_i^m} \mathrm{H}^{\mathrm{T}}} \right| &- \left(\sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{2|1}^i + 1} P_i^\ell \left| \mathcal{B}_{2|1}^{M_i^\ell} \mathrm{H}^{\mathrm{T}} \right| \left| \frac{-\mathcal{B}_{2|1}^{M_i^\ell} \mathrm{A} \mathcal{B}_{2|1}^{M_i^m \mathrm{T}}}{\mathcal{B}_{2|1}^{M_i^\ell} \mathrm{H}^{\mathrm{T}} \cdot \mathcal{B}_{2|1}^{M_i^m} \mathrm{H}^{\mathrm{T}}} \right| \right) |\mathrm{HA}\nu| \\
&+ j \frac{z(2)}{\mathcal{B}_{2|1}^{M_i^m} \mathrm{H}^{\mathrm{T}}} \mathcal{B}_{2|1}^{M_i^m} \nu + j \left(\sum_{\ell=1}^{2} Z_i^\ell \left(\mathcal{B}_{2|1}^{M_i^\ell} \mathrm{H}^{\mathrm{T}} \right) \frac{-\mathcal{B}_{2|1}^{M_i^\ell} \mathrm{A} \mathcal{B}_{2|1}^{M_i^m \mathrm{T}}}{\mathcal{B}_{2|1}^{M_i^\ell} \mathrm{H}^{\mathrm{T}} \cdot \mathcal{B}_{2|1}^{M_i^m} \mathrm{H}^{\mathrm{T}}} \right) \mathrm{HA}\nu \\
&= \frac{-\gamma \left| \mathcal{B}_{2|1}^{M_i^m} \nu \right| - \sum\limits_{\substack{\ell=1 \\ \ell \neq m}}^{L_{2|1}^i + 1} P_i^\ell \left| \mathcal{B}_{2|1}^{M_i^\ell} \mathrm{A} \mathcal{B}_{2|1}^{M_i^m \mathrm{T}} \right| |\mathrm{HA}\nu|}{\mathcal{B}_{2|1}^{M_i^m} \mathrm{H}^{\mathrm{T}}} \\
&\qquad + j \frac{z(2) \mathcal{B}_{2|1}^{M_i^m} \nu - \sum\limits_{\ell=1}^{2} Z_i^\ell \left(\mathcal{B}_{2|1}^{M_i^\ell} \mathrm{A} \mathcal{B}_{2|1}^{M_i^m \mathrm{T}} \right) \mathrm{HA}\nu}{\mathcal{B}_{2|1}^{M_i^m} \mathrm{H}^{\mathrm{T}}}. \qquad (3.28)
\end{aligned}$$

This final form for the new intermediate term's exponential argument involves only two fundamental directions: one that is a row of $\mathcal{B}_{2|1}$ that corresponds to the M_i^m integers for $m \in \{1, \ldots, L^i_{2|1} = 3\}$, and the new vector HA. Moreover, the denominators of both terms are equal. Therefore, we define the set of updated fundamental directions as $\mathcal{B}_{2|2}$ and construct it by appending HA to the bottom of $\mathcal{B}_{2|1}$ as

$$\mathcal{B}_{2|2} = \begin{bmatrix} \mathcal{B}_{2|1} \\ \mathrm{HA} \end{bmatrix} = \begin{bmatrix} \epsilon_1 \Phi^{\mathrm{T}} \\ \epsilon_2 \Phi^{\mathrm{T}} \\ \mathrm{HA}\Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \\ \mathrm{HA} \end{bmatrix}. \tag{3.29}$$

3.3.2 Second Measurement Update - Recursive Structure of the Arrays

Denote an array of integers $M(2|2)$ where the elements of a given row index the rows of $\mathcal{B}_{2|2}$ that appear in the corresponding term. The rows of the old terms will be unchanged, and a set of new rows will be appended to the bottom, each of width two. Since each convolution produces terms involving one of the old term's directions, the new rows of $M(2|2)$ will contain all possible combinations of rows of $\mathcal{B}_{2|1}$ with the new row in $\mathcal{B}_{2|2}$ as

$$M(2|2) = \begin{bmatrix} M(2|1) \\ 1 \ \ 5 \\ 2 \ \ 5 \\ 3 \ \ 5 \\ 4 \ \ 5 \\ 4 \ \ 5 \\ 4 \ \ 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 4 \\ 2 & 3 & 4 \\ 1 & 5 & \\ 2 & 5 & \\ 3 & 5 & \\ 4 & 5 & \\ 4 & 5 & \\ 4 & 5 & \end{bmatrix}. \tag{3.30}$$

Each of the three convolutions in the second measurement update produced four terms, the old term and three new intermediate ones, for a total of nine new terms. However, from (3.30) it is clear that we only have six distinct new terms with different pairs of fundamental directions. This is because three pairs of terms, from different convolutions, have the same exponential arguments and can be combined into one term by summing their polynomial coefficients. These three combined pairs produce the first three new terms, and thus they introduce the first three new rows of $M(2|2)$. The last three rows of $M(2|2)$ are due to the time propagation step and involves the Γ^{T} direction. Although these last three terms have the same fundamental directions, the coefficients for the absolute value functions are different, as will be shown next.

Now we construct an array of coefficients for the absolute value functions in the exponents. From (3.28), a pattern emerges for constructing these coefficients. They are stored in the array $P(2|2)$ that is given by

$$P(2|2) = \begin{bmatrix} & P(2|1) \\ \dfrac{\gamma}{|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\alpha_2 \left|\det \Phi\right| + \beta \left|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right|}{|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\alpha_1 \left|\det \Phi\right| + \beta \left|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right|}{|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\mathrm{H} \mathrm{A} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\gamma \left|\det \Phi\right| + \beta \left|\mathrm{H} \mathrm{A} \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right|}{|\mathrm{H} \mathrm{A} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\mathrm{H}\Gamma|} & \dfrac{\alpha_1 \left|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right| + \alpha_2 \left|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right|}{|\mathrm{H}\Gamma|} \\ \dfrac{\gamma}{|\mathrm{H}\Gamma|} & \dfrac{\frac{\gamma}{|\epsilon_1 \mathrm{H}^{\mathrm{T}}|} \left|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right| + \frac{\alpha_2}{|\epsilon_1 \mathrm{H}^{\mathrm{T}}|} \left|\mathrm{H} \mathrm{A} \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right|}{|\mathrm{H}\Gamma|} \\ \dfrac{\gamma}{|\mathrm{H}\Gamma|} & \dfrac{\frac{\gamma}{|\epsilon_2 \mathrm{H}^{\mathrm{T}}|} \left|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right| + \frac{\alpha_1}{|\epsilon_2 \mathrm{H}^{\mathrm{T}}|} \left|\mathrm{H} \mathrm{A} \Phi^{\mathrm{T}} \mathrm{A} \Gamma\right|}{|\mathrm{H}\Gamma|} \end{bmatrix}. \tag{3.31}$$

Next, denote the new array of coefficients for the imaginary part of the exponential argument as $Z(2|2)$. Since the time propagation step has no effect on the imaginary part of the exponential argument, this array always has width two, involving only the original two directions, and it has the same number of rows as $M(2|2)$. Based on the same manipulations used to obtain the elements of $P(2|2)$, $Z(2|2)$ is given by

$$Z(2|2) = \begin{bmatrix} & Z(2|1) & \\ \dfrac{z(2)}{\epsilon_1 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & & 0 \\ \dfrac{z(2)}{\epsilon_2 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & & 0 \\ \dfrac{z(2)}{\mathrm{H A} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & & -\dfrac{z(1) \det \Phi}{\mathrm{H A} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} \\ \dfrac{z(2)}{\mathrm{H}\Gamma} & & 0 \\ \dfrac{z(2)}{\mathrm{H}\Gamma} & & -\dfrac{\dfrac{z(1)}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_1 \Phi^{\mathrm{T}} \mathrm{A}\Gamma}{\mathrm{H}\Gamma} \\ \dfrac{z(2)}{\mathrm{H}\Gamma} & & -\dfrac{\dfrac{z(1)}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_2 \Phi^{\mathrm{T}} \mathrm{A}\Gamma}{\mathrm{H}\Gamma} \end{bmatrix}. \tag{3.32}$$

The usefulness of $L_{2|2}$ is more apparent in this measurement update, since older terms involve more fundamental directions. It is formed by simply appending an array of 2s of length six to $L_{2|1}$ as

$$L_{2|2} = \begin{bmatrix} L_{2|1}^{\mathrm{T}} & 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 3 & 3 & 3 & 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}}. \tag{3.33}$$

3.3.3 Second Measurement Update - New Coefficients $\mathcal{G}$

Consider now the new coefficient functions for the new intermediate terms produced by the i^{th} term of $\bar{\phi}_{2|1}$. For $m = 1$ and $m = 2$, the numerators of (3.27) are not equal and hence cannot come out of the bracket term. They are of a form compatible with Result 3 in Appendix A. Denote the coefficient of the m^{th} intermediate term as $\mathcal{G}_{2|1}^{i,m}$. Then, the numerators for $m = 1$ are given by

$$
\begin{aligned}
&\mathcal{G}_{2|1}^{i,1}\left(\pm\check{\rho}_1 + \sum_{\ell=2}^{4}\check{\rho}_\ell\ \mathrm{sgn}\left(\mu_\ell - \mu_1\right)\right) \\
&\quad = \bar{a}_i + \bar{b}_i\left(\frac{1}{2}\left(\pm\tilde{\rho}_1 + \sum_{\ell=2}^{4}\tilde{\rho}_\ell\,\mathrm{sgn}\left(\mu_\ell - \mu_1\right)\right)^2 - 1\right) + j\left(\pm\bar{\rho}_1 + \sum_{\ell=2}^{4}\bar{\rho}_\ell\,\mathrm{sgn}\left(\mu_\ell - \mu_1\right)\right) \\
&\qquad = \bar{a}_i + \bar{b}_i\left(\frac{1}{2}\Big(\pm\tilde{\rho}_1 + \tilde{\rho}_2\,\mathrm{sgn}\left(\mu_2 - \mu_1\right)\Big)^2 - 1\right) + j\Big(\pm\bar{\rho}_1 + \bar{\rho}_2\,\mathrm{sgn}\left(\mu_2 - \mu_1\right)\Big) \\
&\qquad\quad = a_i \pm b_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^1}\mathrm{H}^{\mathrm{T}}\cdot\mathcal{B}_{2|1}^{M_i^2}\mathrm{H}^{\mathrm{T}}\right)\ \mathrm{sgn}\left(\mu_2 - \mu_1\right) \\
&\qquad\qquad\qquad \pm jc_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^1}\mathrm{H}^{\mathrm{T}}\right) + jd_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^2}\mathrm{H}^{\mathrm{T}}\right)\ \mathrm{sgn}\left(\mu_2 - \mu_1\right), \\
&\qquad\quad = a_i \pm b_i\,\mathrm{sgn}\left(-\mathcal{B}_{2|1}^{M_i^2}\mathrm{A}\mathcal{B}_{2|1}^{M_i^1\mathrm{T}}\right)\ \mathrm{sgn}\left(\mathrm{HA}\nu\right) \\
&\qquad\qquad \pm jc_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^1}\mathrm{H}^{\mathrm{T}}\right) + jd_i\,\mathrm{sgn}\left(-\mathcal{B}_{2|1}^{M_i^2}\mathrm{A}\mathcal{B}_{2|1}^{M_i^1\mathrm{T}}\cdot\mathcal{B}_{2|1}^{M_i^1}\mathrm{H}^{\mathrm{T}}\right)\ \mathrm{sgn}\left(\mathrm{HA}\nu\right).
\end{aligned} \tag{3.34}
$$

The same manipulations can be done for the numerators for $m = 2$ by interchanging μ_1 and c_i with μ_2 and d_i, respectively. Recalling that $\check{\rho}_3 = 0$, the new term obtained for $m = 3$ corresponds to the coefficient

$$
\begin{aligned}
&\mathcal{G}_{2|1}^{i,3}\left(0 + \sum_{\substack{\ell=1\\ \ell\neq 3}}^{4}\check{\rho}_\ell\ \mathrm{sgn}\left(\mu_\ell - \mu_3\right)\right) \\
&\qquad = a_i + b_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^1}\mathrm{H}^{\mathrm{T}}\cdot\mathcal{B}_{2|1}^{M_i^2}\mathrm{H}^{\mathrm{T}}\right)\ \mathrm{sgn}\left(\mu_1 - \mu_3\right)\ \mathrm{sgn}\left(\mu_2 - \mu_3\right) \\
&\qquad\quad + jc_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^1}\mathrm{H}^{\mathrm{T}}\right)\ \mathrm{sgn}\left(\mu_1 - \mu_3\right) + jd_i\,\mathrm{sgn}\left(\mathcal{B}_{2|1}^{M_i^2}\mathrm{H}^{\mathrm{T}}\right)\ \mathrm{sgn}\left(\mu_2 - \mu_3\right) \\
&\qquad = a_i + b_i\,\mathrm{sgn}\left(-\mathcal{B}_{2|1}^{M_i^1}\mathrm{A}\mathcal{B}_{2|1}^{M_i^3\mathrm{T}}\right)\ \mathrm{sgn}\left(-\mathcal{B}_{2|1}^{M_i^1}\mathrm{A}\mathcal{B}_{2|1}^{M_i^3\mathrm{T}}\right) \\
&\qquad\qquad + j\left(c_i\,\mathrm{sgn}\left(-\mathcal{B}_{2|1}^{M_i^1}\mathrm{A}\mathcal{B}_{2|1}^{M_i^3\mathrm{T}}\right) + d_i\,\mathrm{sgn}\left(-\mathcal{B}_{2|1}^{M_i^2}\mathrm{A}\mathcal{B}_{2|1}^{M_i^3\mathrm{T}}\right)\right)\ \mathrm{sgn}\left(\mathrm{HA}\nu\right).
\end{aligned} \tag{3.35}
$$

Substituting these into the bracket term in (3.27), and noting that the real parts of the denominators are similar to the real part of the exponential argument, yields a form that is compatible with the numerator forms given in Result 3 to produce coefficient functions for the intermediate terms denoted $\mathcal{G}_{2|2}^{i,m}(\nu)$ and given by

$$
\mathcal{G}_{2|2}^{i,m}(\nu) = a_{i,m} + b_{i,m}\,\mathrm{sgn}\left(\mathcal{B}_{2|2}^{M_{i,m}^1}\nu\right)\ \mathrm{sgn}\left(\mathcal{B}_{2|2}^{M_{i,m}^2}\nu\right) + jc_{i,m}\,\mathrm{sgn}\left(\mathcal{B}_{2|2}^{M_{i,m}^1}\nu\right) + jd_{i,m}\,\mathrm{sgn}\left(\mathcal{B}_{2|2}^{M_{i,m}^2}\nu\right). \tag{3.36}
$$

It is necessary to use Result 3 because the denominators in the bracket term in (3.27) form two sign functions, in the same way that the new exponential arguments (3.28) involve only two absolute value terms. Therefore, polynomial coefficients for all of the new intermediate terms can be expressed in the simple four-parameter form of (3.36). Note that the two fundamental directions involved here are the same as in the argument of the exponential. Hence, combining two terms with the same exponential arguments involves summing their polynomial coefficients $\mathcal{G}_{2|2}^{i,m}$, which is simply summing the corresponding parameters.

For the second update, three pairs of the nine new intermediate terms can be combined, leaving a total of six new terms in addition to the three old terms. The set of combined parameters for the second measurement update are appended to the bottom of $G(2|1) = G(1|1)$, producing the new array $G(2|2)$, which has four columns and the same number of rows as $M(2|2)$. Then, the coefficients for the new terms (i.e.,those corresponding to $i \in \{4, \ldots, 9\}$) are given by

$$\mathcal{G}_{2|2}^{i}(\nu) = a_i + b_i \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^1}\nu\right) \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^2}\nu\right) + jc_i \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^1}\nu\right) + jd_i \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^2}\nu\right), \tag{3.37}$$

where a_i, b_i, c_i, and d_i are the elements of the i^{th} row of $G(2|2)$.

3.3.4 Second Measurement Update - Old Terms

For $\ell = L_{k|k-1}^{i} + 1$ the constants $\{\tilde{\rho}_\ell, \bar{\rho}_\ell\} := \check{\rho}_\ell = 0$, which implies that both numerators in the bracket term in (3.27) are equal and hence can be pulled out of the brackets as a common factor. Therefore, the measurement updated old term is the same as the previous old term, except its coefficient is multiplied by the new bracket term,

$$\begin{aligned}\mathcal{G}_{2|2}^{i}(\nu) = &\left\{a_i + b_i \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^1}\nu\right) \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^2}\nu\right) + jc_i \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^1}\nu\right) + jd_i \operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^2}\nu\right)\right\} \\ &\times \frac{1}{2\pi}\left\{\left[j\theta_2^i + \gamma + \sum_{\ell=1}^{L_{2|2}^i=3} P_i^\ell\left(\mathcal{B}_{2|2}^{M_i^\ell}\mathrm{H}^{\mathrm{T}}\right)\operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^\ell}\nu\right)\right]^{-1}\right. \\ &\qquad\left. - \left[j\theta_2^i - \gamma + \sum_{\ell=1}^{L_{2|2}^i=3} P_i^\ell\left(\mathcal{B}_{2|2}^{M_i^\ell}\mathrm{H}^{\mathrm{T}}\right)\operatorname{sgn}\left(\mathcal{B}_{2|2}^{M_i^\ell}\nu\right)\right]^{-1}\right\}.\end{aligned} \tag{3.38}$$

Since the new bracket term in (3.38) has three sign functions, it cannot be manipulated using the results in Appendix A.

This implies that terms that have been created in previous measurement updates retain their exponential arguments during the current measurement update, and their coefficients are multiplied by new bracket terms as in (3.38), with one additional sign function than the coefficient had in the previous update. The difference between the current time step and the time step when the term was originally created is called the *age* of the term, and is given by $L_{k|k} - 2$. Although this complicated structure grows, acquiring a new bracket term every measurement update, the structure of each bracket term overlaps significantly with the others, and thus can be expressed efficiently.

3.3.5 Second Measurement Update - The Characteristic Function

Using (3.29) (3.30) (3.31) (3.32) (3.33), and denoting $N_{2|2} = 9$ for the number of terms, the CF of the ucpdf for the second measurement update is given by

$$\bar{\phi}_{2|2}(\nu) = \sum_{i=1}^{N_{2|2}} \mathcal{G}_{2|2}^i(\nu) \cdot \exp\left(\mathcal{E}_{2|2}^i(\nu)\right) \tag{3.39a}$$

where

$$\mathcal{E}_{2|2}^i(\nu) = -\sum_{\ell=1}^{L_{2|2}^i} P_i^\ell \left|\mathcal{B}_{2|2}^{M_i^\ell}\nu\right| + j\left(\sum_{\ell=1}^{2} Z_i^\ell \mathcal{B}_{2|2}^{M_i^\ell}\right)\nu \tag{3.39b}$$

and the coefficients are given by (3.37) and (3.38) for new and old terms, respectively. Note that the first three terms in $\bar{\phi}_{2|2}$ correspond to the old terms and have the same exponential parts as the three terms in $\bar{\phi}_{2|1}$. The subsequent six terms are new and involve only two fundamental directions. Although not used here, it will be useful to denote the number of new terms in $\bar{\phi}_{2|2}$ as $N_{2|2}^n = 6$. This structure repeats itself across subsequent measurement updates, and the next section will present the two-state estimator's recursion for the general k^{th} measurement update.

3.4 Estimator Recursion for the k^{th} Propagation and Update

This section presents the two-state Cauchy estimator algorithm for a state propagation and then the general measurement update for a general time step k. This algorithm was derived by induction based on a study of the first three measurement updates. It can be verified to be a special case of the approach developed in [14,15]. We first address the state propagation at time step k, and then follow with the measurement update recursion. Finally, we show how to determine the minimum variance estimate and its error variance.

3.4.1 Time Propagation

The general time propagation uses the same formula as in the first time propagation, given in (2.28),

$$\bar{\phi}_{k+1|k}(\nu) = \bar{\phi}_{k|k}(\Phi^{\mathrm{T}}\nu) \cdot e^{-\beta\left|\Gamma^{\mathrm{T}}\nu\right|}. \tag{3.40}$$

The characteristic function of the ucpdf for the once propagated conditional density can be written as

$$\bar{\phi}_{k+1|k}(\nu) = \sum_{i=1}^{N_{k+1|k}} \mathcal{G}^{i}_{k+1|k}(\nu) \cdot \exp\left(- \sum_{\ell=1}^{L^{i}_{k+1|k}} P^{\ell}_{i} \left| \mathcal{B}^{M^{\ell}_{i}}_{k+1|k} \nu \right| + j \left(\sum_{\ell=1}^{2} Z^{\ell}_{i} \mathcal{B}^{M^{\ell}_{i}}_{k+1|k} \right) \nu \right), \tag{3.41a}$$

where

$$P(k+1|k) = \begin{bmatrix} & \beta \\ P(k|k) & \beta \\ & \vdots \\ & \beta \end{bmatrix}, \qquad Z(k+1|k) = Z(k|k),$$

$$M(k+1|k) = \begin{bmatrix} & 2(k+1) \\ M(k|k) & 2(k+1) \\ & \vdots \\ & 2(k+1) \end{bmatrix}, \quad L_{k+1|k} = L_{k|k} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \mathcal{B}_{k+1|k} = \begin{bmatrix} \mathcal{B}_{k|k} \cdot \Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \end{bmatrix}, \tag{3.41b}$$

and both $N_{k+1|k} = N_{k|k}$ and $N^{n}_{k+1|k} = N^{n}_{k|k}$ because no new terms are created during the time propagation.

3.4.2 Measurement Update - The Convolution

We begin this section assuming we have a time propagated ucpdf's CF $\bar{\phi}_{k|k-1}$ and are given the k^{th} measurement $z(k)$. Consider the following generalization for the polynomial coefficients $\mathcal{G}^i_{k|k-1}$ produced for the old terms in the second measurement update. The form in (3.38) suggests that, in general, the time propagated coefficients will have the form

$$\begin{aligned}\mathcal{G}^i_{k|k-1}(\nu) = \Big\{ a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{k|k-1}\nu\right) \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{k|k-1}\nu\right) \\ +jc_i \operatorname{sgn}\left(\mathcal{B}^{M^1_i}_{k|k-1}\nu\right) + jd_i \operatorname{sgn}\left(\mathcal{B}^{M^2_i}_{k|k-1}\nu\right)\Big\} \\ \times \prod_{r=1}^{L^i_{k|k-1}-3} \frac{1}{2\pi}\left\{\frac{1}{j\theta^i_{k-r}+\gamma+\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k-1}\nu)} - \frac{1}{j\theta^i_{k-r}-\gamma+\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k-1}\nu)}\right\},\end{aligned} \tag{3.42a}$$

where

$$\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k-1}\nu) = \sum_{\ell=1}^{L^i_{k-r|k-r}} P^\ell_i \left(\mathcal{B}^{M^\ell_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}\left(\mathcal{B}^{M^\ell_i}_{k|k-1}\nu\right), \tag{3.42b}$$

$$\theta^i_{k-r} = z(k-r) - Z^1_i \mathcal{B}^{M^1_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} - Z^2_i \mathcal{B}^{M^2_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}, \tag{3.42c}$$

and $k-r$ is the time-step where the term involving $\mathcal{S}^i_{k-r}$ was created. The arrays M,L,P,$\mathcal{B}$, and Z for the general update will be constructed later but correspond to those of the first two updates already shown.

Using the measurement update formula (2.13) with the corresponding change of indices, $\bar{\phi}_{k|k}$ is given by

$$\begin{aligned}\bar{\phi}_{k|k}(\nu) = \sum_{i=1}^{N_{k|k-1}} \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{G}^i_{k|k-1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma) \\ \times \exp\left(-\sum_{\ell=1}^{L^i_{k|k-1}} P^\ell_i \left|\mathcal{B}^{M^\ell_i}_{k|k-1}(\nu - \mathrm{H}^{\mathrm{T}}\sigma)\right| - \gamma\left|-\sigma\right| + j\theta^i_k\sigma\right) d\sigma.\end{aligned} \tag{3.43}$$

It is necessary to divide the domain of integration into regions in which the polynomial coefficient is a constant and the exponential argument is continuous. For this, the summations that appear in (3.42) must involve the same fundamental directions as the exponential argument. Since the top bracket term in (3.42a) is the four parameter polynomial formed when the term was created, we can define $\bar{\rho}_\ell$ and $\bar{\rho}_m$ constants as in (3.24), where $\bar{\rho}_\ell = \tilde{\rho}_\ell = 0$ for $\ell = 3, \dots, L^i_{k|k-1} + 1$.

Similarly, the summations (3.42b) in the rest of the bracket terms can be written using constants ${}^{(r)}\breve{\rho}_\ell$ defined as

$$
{}^{(r)}\breve{\rho}_\ell = \begin{cases} P_i^\ell \left(\mathcal{B}^{M_i^\ell}_{k-r|k-r} \mathrm{H}^\mathrm{T} \right) \operatorname{sgn} \left(\mathcal{B}^{M_i^\ell}_{k|k-1} \mathrm{H}^\mathrm{T} \right) & \ell = 1, \dots, L^i_{k|k-1} - r \\ 0 & \ell = L^i_{k|k-1} - r + 1, \dots, L^i_{k|k-1} + 1. \end{cases} \tag{3.44}
$$

Using the shorthand $\breve{\rho}_\ell$ as in the second measurement update (3.25), construct μ_ℓ, θ^i_k, ρ_ℓ, $\bar{\rho}_\ell$, and $\tilde{\rho}_\ell$ constants as in (3.19b), (3.20b), (3.20c), and (3.24), respectively, in order to write the coefficients in the i^{th} integral in (3.43) as

$$
\begin{aligned}
&\mathcal{G}^i_{k|k-1}(\nu - \mathrm{H}^\mathrm{T}\sigma) = \\
&\qquad \mathcal{G}^i_{k|k-1} \Bigg(\sum_{\ell=1}^{L^i_{k|k-1}+1} \tilde{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \sum_{\ell=1}^{L^i_{k|k-1}+1} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \\
&\qquad\qquad \sum_{\ell=1}^{L^i_{k|k-1}+1} {}^{(L^i_{k|k-1}-2)}\breve{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \dots, \sum_{\ell=1}^{L^i_{k|k-1}+1} {}^{(1)}\breve{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \Bigg) \\
&\qquad\qquad := \mathcal{G}^i_{k|k-1} \left(\sum_{\ell=1}^{L^i_{k|k-1}+1} \breve{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right).
\end{aligned} \tag{3.45}
$$

The set of parameters needed in the sums in (3.45) are visualized in the array

$$
\begin{array}{cccccccc}
 & \tilde{\rho}_1 & \tilde{\rho}_2 & 0 & 0 & \cdots & 0 & 0 \\
 & \bar{\rho}_1 & \bar{\rho}_2 & 0 & 0 & \cdots & 0 & 0 \\
 & {}^{(L^i_k-2)}\breve{\rho}_1 & {}^{(L^i_k-2)}\breve{\rho}_2 & {}^{(L^i_k-2)}\breve{\rho}_3 & 0 & \cdots & 0 & 0 \\
\uparrow & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
r & {}^{(2)}\breve{\rho}_1 & {}^{(2)}\breve{\rho}_2 & \cdots & {}^{(2)}\breve{\rho}_{L^i_k-2} & 0 & 0 & 0 \\
 & {}^{(1)}\breve{\rho}_1 & {}^{(1)}\breve{\rho}_2 & \cdots & {}^{(1)}\breve{\rho}_{L^i_k-2} & {}^{(1)}\breve{\rho}_{L^i_k-1} & 0 & 0 \\
 & \rho_1 & \rho_2 & \cdots & \rho_{L^i_k-2} & \rho_{L^i_k-1} & \rho_{L^i_k} & \rho_{L^i_k+1} \\
 & & & & \ell \longrightarrow & & &
\end{array} \tag{3.46}
$$

where the bottom row is the set from the exponential argument, so that $\rho_{L_k^i} = \beta$ and $\rho_{L_k^i+1} = \gamma$. Then, the entire integral can be written compactly as

$$I_i = \int_{-\infty}^{\infty} \mathcal{G}_{k|k-1}^i \left(\sum_{\ell=1}^{L_{k|k-1}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \right) \cdot \exp\left(- \sum_{\ell=1}^{L_{k|k-1}^i+1} \rho_\ell \left|\mu_\ell - \sigma\right| + j\theta_k^i \sigma \right) d\sigma. \quad (3.47)$$

Using the definition of s_i^ℓ from (2.16), the arguments of the coefficient functions are constant, and hence $\mathcal{G}_{k|k-1}^i$ is a constant in the integral. Therefore, it can be rewritten as a sum of $L_{k|k-1}^i + 1$ integrals and evaluated as

$$\begin{aligned}
I_i =& \sum_{i=0}^{L_{k|k-1}^i+1} \left\{ \int_{\mu_i}^{\mu_{i+1}} \mathcal{G}_{k|k-1}^i \left(\sum_{\ell=1}^{L_{k|k-1}^i+1} \check{\rho}_\ell s_i^\ell \right) \exp\left(- \sum_{\ell=1}^{L_{k|k-1}^i+1} \rho_\ell \left(\mu_\ell - \sigma\right) s_i^\ell + jz(k)\sigma \right) d\sigma \right\} \\
=& \frac{1}{2\pi} \sum_{i=0}^{n} \left\{ \frac{\mathcal{G}_{k|k-1}^i \left(\sum_{\ell=1}^{L_{k|k-1}^i+1} \check{\rho}_\ell s_i^\ell \right) \exp\left(- \sum_{\ell=1}^{L_{k|k-1}^i+1} \rho_\ell \left(\mu_\ell - \mu_{i+1}\right) s_i^\ell + jz(k)\mu_{i+1} \right)}{jz(k) + \sum_{\ell=1}^{L_{k|k-1}^i+1} \rho_\ell s_i^\ell} \right\} \\
& - \frac{1}{2\pi} \sum_{i=0}^{n} \left\{ \frac{\mathcal{G}_{k|k-1}^i \left(\sum_{\ell=1}^{L_{k|k-1}^i+1} \check{\rho}_\ell s_i^\ell \right) \exp\left(- \sum_{\ell=1}^{L_{k|k-1}^i+1} \rho_\ell \left(\mu_\ell - \mu_i\right) s_i^\ell + jz(k)\mu_i \right)}{jz(k) + \sum_{\ell=1}^{L_{k|k-1}^i+1} \rho_\ell s_i^\ell} \right\}
\end{aligned} \quad (3.48)$$

Then, by following the same manipulations as in Section 2.1.1, (3.48) can be rewritten as

$$\begin{aligned}
I_i =& \sum_{m=1}^{L_{k|k-1}^i+1} \exp\left(- \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{k|k-1}^i+1} \rho_\ell \left|\mu_\ell - \mu_m\right| + j\theta_k^i \mu_m \right) \\
& \times \left\{ \frac{\mathcal{G}_{k|k-1}^i \left(+\check{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{k|k-1}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m) \right)}{j\theta_k^i + \rho_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{k|k-1}^i+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)} - \frac{\mathcal{G}_{k|k-1}^i \left(-\check{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{k|k-1}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m) \right)}{j\theta_k^i - \rho_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{k|k-1}^i+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)} \right\},
\end{aligned}$$

(3.49a)

where

$$\mathcal{G}^i_{k|k-1}\left(\pm\check{\rho}_m + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k|k-1}+1} \check{\rho}_\ell \ \mathrm{sgn}\left(\mu_\ell - \mu_m\right)\right) =$$

$$\left\{a_i + b_i\left(\frac{1}{2}\Big(\pm\tilde{\rho}_m + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k|k-1}+1}\tilde{\rho}_\ell\,\mathrm{sgn}\left(\mu_\ell-\mu_m\right)\Big)^2 - 1\right) + j\left(\pm\bar{\rho}_m + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k|k-1}+1}\bar{\rho}_\ell\,\mathrm{sgn}\left(\mu_\ell-\mu_m\right)\right)\right\}$$

$$\times\prod_{p=1}^{L^i_{k|k-1}-3}\frac{1}{2\pi}\left\{\frac{1}{j\theta^i_{k-r}+\gamma+\mathcal{S}^i_{k-r}(\pm^{(r)}\check{\rho}_m,\mu_\ell-\mu_m)} - \frac{1}{j\theta^i_{k-r}-\gamma+\mathcal{S}^i_{k-r}(\pm^{(r)}\check{\rho}_m,\mu_\ell-\mu_m)}\right\},$$

(3.49b)

and for $r \in \{1,\dots,L^i_{k|k-1}-3\}$,

$$\begin{aligned}
&\mathcal{S}^i_{k-r}(\pm^{(r)}\check{\rho}_m,\mu_\ell-\mu_m)\\
&\quad= \pm P^m_i\left(\mathcal{B}^{M^m_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right) + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k-r|k-r}} P^\ell_i\left(\mathcal{B}^{M^\ell_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\mathrm{sgn}\left(\mathcal{B}^{M^\ell_i}_{k|k-1}\mathrm{H}^{\mathrm{T}}\right)\mathrm{sgn}\left(\mu_\ell-\mu_m\right)\\
&= \pm P^m_i\left(\mathcal{B}^{M^m_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right) + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k-r|k-r}} P^\ell_i\left(\mathcal{B}^{M^\ell_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\mathrm{sgn}\left(-\mathcal{B}^{M^\ell_i}_{k|k-1}\mathrm{A}\mathcal{B}^{M^m_i\mathrm{T}}_{k|k-1}\cdot\mathcal{B}^{M^\ell_i}_{k|k-1}\mathrm{H}^{\mathrm{T}}\right)\mathrm{sgn}\left(\mathrm{HA}\nu\right).
\end{aligned}$$

(3.49c)

3.4.3 Measurement Update - Recovering the CF Structure

The old terms will be discussed later. For the new terms, i.e., $m = 1,\dots,L^i_{k|k-1}$, the numerators of the measurement updated coefficients (3.49b) can be reduced to a form compatible with Result 3. The four parameter bracket terms in (3.49b) can be rewritten as (3.36) from the second measurement update. Manipulate the bracket term outside the product in (3.49b) in the same manner as (3.34) and (3.35). Since Results 1 and 2 produce outputs that are of the same form as their respective inputs, they can be applied sequentially across the bracket terms. Then, the entire numerator form (3.49b) can be collapsed, using these results sequentially, into the form given in (3.36). Finally, the new four-parameter polynomial coefficient involving two fundamental directions (HA and a row of $\mathcal{B}_{k|k-1}$) can be computed as in the second measurement update using Result 3. Those parameters will be combined with other terms with the same exponential arguments, the set of new parameters will be appended to the bottom of $G(k|k-1)$ to form a new array denoted $G(k|k)$. Denote the number of the new terms in $\bar{\phi}_{k|k}$ as $N^n_{k|k}$ so that $N_{k|k} = N_{k|k-1} + N^n_{k|k}$.

Since one of the two fundamental directions for every new term will be HA, it is appended to $\mathcal{B}_{k|k-1}$ to obtain

$$\mathcal{B}_{k|k} = \begin{bmatrix} \mathcal{B}_{k|k-1} \\ \text{HA} \end{bmatrix}. \tag{3.50}$$

The array $M(k|k)$ is constructed by appending $N^n_{k|k}$ new rows to $M(k|k-1)$. The recursion is given by

$$M(k|k) = \begin{bmatrix} \multicolumn{2}{c}{M(k|k-1)} \\ M^1_{(r_k+1)} & 2k+1 \\ M^1_{(r_k+2)} & 2k+1 \\ \vdots & \vdots \\ M^1_{N_{k|k-1}} & 2k+1 \\ 2k-1 & 2k+1 \\ 2k & 2k+1 \\ 2k & 2k+1 \\ \vdots & \vdots \\ 2k & 2k+1 \end{bmatrix}, \tag{3.51}$$

where $r_k = N_{k-1|k-1} - N^n_{k-1|k-1}$, noting that $N_{k|k-1} = N_{k-1|k-1}$ and $N^n_{k|k-1} = N^n_{k-1|k-1}$. The left column of this new block has three parts. The first $N^n_{k-1|k-1}$ elements are the left elements of all the new rows of the previous measurement update, i.e. the new $N^n_{k-1|k-1}$ rows of $M(k-1|k-1)$. The next element corresponds to the new fundamental direction from the previous measurement update, i.e. the last row of $\mathcal{B}_{k-1|k-1}$. The remaining $N_{k-1|k-1}$ rows are all the same, and involve the new time propagation direction Γ^{T}.

The growth in the number of terms, which is given as a linear dynamic system of integers, is based on the pattern for the recursion of $M(k|k)$ given above. The number of terms in the sum in $\bar{\phi}_{k|k}$, given by $N_{k|k}$, is determined from the previous number of terms $N_{k-1|k-1}$

and the previous number of new terms $N^n_{k-1|k-1}$ by the following linear relationship:

$$\begin{bmatrix} N_{k|k} \\ N^n_{k|k} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} N_{k-1|k-1} \\ N^n_{k-1|k-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{3.52}$$

The recursions for the fundamental arrays $P(k|k)$ and $Z(k|k)$ were derived by induction from a study of the first three measurement updates. As in the first two measurement updates, the denominators of the elements of the i^{th} row of $Z(k|k)$ are equal. Therefore, it is useful to denote the diagonal matrix $D(k|k)$, where

$$D(k|k) = \operatorname{Diag}\left[\frac{1}{\mathcal{B}_{k|k}^{M^1_{r_k+1}}\mathrm{H}^{\mathrm{T}}}, \ldots, \frac{1}{\mathcal{B}_{k|k}^{M^1_{(r_k+N^n_{k|k-1})}}\mathrm{H}^{\mathrm{T}}}, \frac{1}{\mathrm{HA}\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}}, \frac{1}{\mathrm{H}\Gamma}, \ldots, \frac{1}{\mathrm{H}\Gamma}\right]. \tag{3.53}$$

Similarly, the denominators of the elements of the i^{th} row of $P(k|k)$ are also equal. Denote the diagonal matrix $\bar{D}(k|k)$, whose elements equal the absolute values of the corresponding elements in $D(k|k)$.

The new terms in $P(k|k)$ and $Z(k|k)$ are formed in the same manner as in the first and second measurement update, reducing sums of $\mu_\ell - \mu_m$ terms to constants times HA and reducing the polynomial coefficients to the four parameter structure in (3.3) using the Results in Appendix A, producing the new rows for $G(k|k)$. The recursion for $P(k|k)$ is

$$P(k|k) = \begin{bmatrix} P(k|k-1) \\ \bar{D}(k|k) \times \begin{bmatrix} \gamma & P^2_{(r_k+1)}\left|\mathcal{B}_{k-1|k-1}^{M^1_{(r_k+1)}}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi\right| + \beta\left|\mathcal{B}_{k|k}^{M^1_{(r_k+1)}}\mathrm{A}\Gamma\right| \\ \gamma & P^2_{(r_k+2)}\left|\mathcal{B}_{k|k-1}^{M^1_{(r_k+2)}}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi\right| + \beta\left|\mathcal{B}_{k|k}^{M^1_{(r_k+2)}}\mathrm{A}\Gamma\right| \\ \vdots & \vdots \\ \gamma & P^2_{r_k+N_{k|k-1}}\left|\mathcal{B}_{k-1|k-1}^{M^1_{(r_k+N^n_{k|k-1})}}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi\right| + \beta\left|\mathcal{B}_{k|k}^{M^1_{(r_k+N^n_{k|k-1})}}\mathrm{A}\Gamma\right| \\ \gamma & \gamma\left|\det\Phi\right| + \beta\left|\mathrm{HA}\Phi^{\mathrm{T}}\mathrm{A}\Gamma\right| \\ \gamma & \sum_{\ell=1}^{L^1_{k|k-1}-1} P^\ell_1\left|\mathcal{B}_k^{M^\ell_1}\mathrm{A}\Gamma\right| \\ \vdots & \vdots \\ \gamma & \sum_{\ell=1}^{L^{(N_{k|k-1})}_{k|k-1}-1} P^\ell_{N_{k-1|k-1}}\left|\mathcal{B}_k^{M^\ell_{(N_{k-1|k-1})}}\mathrm{A}\Gamma\right| \end{bmatrix} \end{bmatrix}. \tag{3.54}$$

Similarly, the recursion for $Z(k|k)$ is

$$Z(k|k) = \begin{bmatrix} Z(k|k-1) \\ D(k|k) \times \begin{bmatrix} z(k) & 0 \\ z(k) & 0 \\ z(k) & -Z^2_{(r_k+3)}\left(\mathcal{B}^{M^1_{(r_k+3)}}_{k-1|k-1}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi\right) \\ \vdots & \vdots \\ z(k) & -Z^2_{(r_k+N^n_{k-1})}\left(\mathcal{B}^{M^1_{(r_k+N^n_{k-1})}}_{k-1|k-1}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi\right) \\ z(k) & -z(k-1)\cdot\det\Phi \\ z(k) & -\sum_{\ell=1}^{2} Z^{\ell}_{(N_{k-1|k-1})}\left(\mathcal{B}^{M^{\ell}_1}_{k|k}\mathrm{A}\Gamma\right) \\ \vdots & \vdots \\ z(k) & -\sum_{\ell=1}^{2} Z^{\ell}_{(N_{k-1|k-1})}\left(\mathcal{B}^{M^{\ell}_{(N_{k-1|k-1})}}_{k|k}\mathrm{A}\Gamma\right) \end{bmatrix} \end{bmatrix}. \tag{3.55}$$

Since all the new rows in $M(k|k)$ have width two, $L_{k|k}$ is measurement updated by appending $N^n_{k|k}$ elements of the integer 2 to $L_{k|k-1}$ as

$$L^{\mathrm{T}}_{k|k} = \left[L^{\mathrm{T}}_{k|k-1}\,,\quad 2,\quad \ldots,\quad 2\right]^{\mathrm{T}}. \tag{3.56}$$

This produces all of the parameters necessary to express the ucpdf's CF for the k^{th} measurement update as

$$\bar{\phi}_{k|k}(\nu) = \sum_{i=1}^{N_{k|k}} \mathcal{G}^i_{k|k}(\nu)\cdot\exp\left(-\sum_{\ell=1}^{L^i_{k|k}} P^{\ell}_i\left|\mathcal{B}^{M^{\ell}_i}_{k|k}\nu\right| + j\left(\sum_{\ell=1}^{2} Z^{\ell}_i\mathcal{B}^{M^{\ell}_i}_{k|k}\right)\nu\right), \tag{3.57a}$$

where coefficients are given by

$$\begin{aligned}\mathcal{G}^i_{k|k}(\nu) = {} & \left\{a_i + b_i\,\mathrm{sgn}\left(\mathcal{B}^{M^1_i}_{k|k}\nu\right)\,\mathrm{sgn}\left(\mathcal{B}^{M^2_i}_{k|k}\nu\right) + jc_i\,\mathrm{sgn}\left(\mathcal{B}^{M^1_i}_{k|k}\nu\right) + jd_i\,\mathrm{sgn}\left(\mathcal{B}^{M^2_i}_{k|k}\nu\right)\right\} \\ & \times \prod_{r=1}^{L^i_{k|k}-2}\frac{1}{2\pi}\left\{\frac{1}{j\theta^i_{k-r}+\gamma+\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k}\nu)} - \frac{1}{j\theta^i_{k-r}-\gamma+\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k}\nu)}\right\},\end{aligned} \tag{3.57b}$$

where

$$\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k}\nu) = \sum_{\ell=1}^{L^i_{k-r|k-r}} P^{\ell}_i\left(\mathcal{B}^{M^{\ell}_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\,\mathrm{sgn}\left(\mathcal{B}^{M^{\ell}_i}_{k|k}\nu\right), \tag{3.57c}$$

$$\theta^i_{k-r} = z(k-r) - Z^1_i\mathcal{B}^{M^1_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} - Z^2_i\mathcal{B}^{M^2_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}, \tag{3.57d}$$

and $k-r$ is the time-step where the term involving $\mathcal{S}^i_{k-r}$ was created. Hence, the age of the i^{th} term after the k^{th} update is $L^i_{k|k}-2$, and new terms have age zero.

3.4.4 Method for Combining Terms

When searching for terms with the same exponenetial argument, in order to combine their coefficient functions, we use the indices in the rows of $M(k|k)$. However, certain sets of rows of $M(k|k)$ are equal but the corresponding rows of $P(k|k)$ and $Z(k|k)$ are not equal. One way to determine which terms to combine would be to compare rows of $M(k|k)$ and of $P(k|k)$ and $Z(k|k)$ as well, but the problem is that comparing floating point numbers is less efficient and less accurate than comparing integers.

There is a pattern in how the terms that have matching rows of $M(k|k)$ combine, and by using this pattern there is no need to compare floating point numbers at all. By keeping track of which rows are repeating and which rows are unique, as well as when those repeating rows were created, then it is possible to determine which intermediate terms to combine to generate each of the new terms corresponding to repeated rows of $M(k|k)$. This is because each new process noise term, those that go with $\left|\Gamma^{\mathrm{T}}\nu\right|$ and $|\mathrm{HA}\nu|$, acts like a new initial condition, and begins producing new terms of its own with each new update. Then, all terms that have matching exponential arguments have the same fundamental directions, indexed in $M(k|k)$, *and* that were produced by either that original term or any of its offspring are combined into one term. This term will have a unique expnential argument, distinct from all other terms that have the same fundamental directions. Therefore, by performing the convolutions in order of increasing i and then combining their terms with matching exponential parts, the correct combinations are achieved.

3.5 Second Propagation

Here we present the second propagation, using (2.28) to get

$$\phi_{3|2}(\nu) = \phi_{2|2}(\Phi^{\mathrm{T}}\nu) \cdot e^{-\beta\left|\Gamma^{\mathrm{T}}\nu\right|}. \tag{3.58}$$

The un-normalized characteristic function for the once propagated conditional density can be written as

$$\bar{\phi}_{3|2}(\nu) = \sum_{i=1}^{N_{3|2}} \mathcal{G}_{3|3}^{i}(\nu) \cdot \exp\left(-\sum_{\ell=1}^{L_{3|2}^{i}} P_i^{\ell} \left| \mathcal{B}_{3|2}^{M_i^{\ell}} \nu \right| + j\left(Z_i^1 \mathcal{B}_{3|2}^{M_i^1} + Z_i^2 \mathcal{B}_{3|2}^{M_i^2} \right) \nu \right) \tag{3.59a}$$

where $Z(3|2) = Z(2|2)$, $G(3|2) = G(2|2)$,

$$P(3|2) = \begin{bmatrix} & \beta \\ & \beta \\ P(2|2) & \vdots \\ & \beta \end{bmatrix}, \qquad \mathcal{B}_{3|2} = \begin{bmatrix} \mathcal{B}_{2|2} \cdot \Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \end{bmatrix},$$

$$M(3|2) = \begin{bmatrix} & 6 \\ & 6 \\ M(2|2) & \vdots \\ & 6 \end{bmatrix}, \quad L_{3|2} = L_{2|2} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \tag{3.59b}$$

and $N_{3|2} = N_{2|2} = 9$.

3.6 Third Measurement Update

This section will present only the exponential argument for the third measurement update, which was used to obtain the general recursion. The characteristic function for the third measurement update can be written as

$$\bar{\phi}_{3|3}(\nu) = \sum_{i=1}^{N_{3|3}} \mathcal{G}_{3|3}^{i}(\nu) \cdot \exp\left(-\sum_{\ell=1}^{L_{3|3}^{i}} P_i^{\ell} \left| \mathcal{B}_{3|3}^{M_i^{\ell}} \nu \right| + j\left(Z_i^1 \mathcal{B}_{3|3}^{M_i^1} + Z_i^2 \mathcal{B}_{3|3}^{M_i^2} \right) \nu \right) \tag{3.60}$$

The new four-parameter coefficient functions can be obtained as described previously. Apply the general recurision given in Section 3.4 to the state propagated CF given in (3.59). The set of fundamental vectors $\mathcal{B}_{3|3}$ is formed by appending HA to the bottom of $\mathcal{B}_{3|2}$, as

$$\mathcal{B}_{3|3} = \begin{bmatrix} \mathcal{B}_{3|2} \\ \mathrm{HA} \end{bmatrix} = \begin{bmatrix} \epsilon_1 \Phi^{\mathrm{T}2} \\ \epsilon_2 \Phi^{\mathrm{T}2} \\ \mathrm{HA}\Phi^{\mathrm{T}2} \\ \Gamma^{\mathrm{T}}\Phi^{\mathrm{T}} \\ \mathrm{HA}\Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \\ \mathrm{HA} \end{bmatrix}. \tag{3.61}$$

These are indexed by integers in the array $M(3|3)$, given by

$$M(3|3) = \begin{bmatrix} \multicolumn{2}{c}{M(3|2)} \\ 1 & 7 \\ 2 & 7 \\ 3 & 7 \\ 4 & 7 \\ 4 & 7 \\ 4 & 7 \\ 5 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \\ 6 & 7 \end{bmatrix}. \tag{3.62}$$

Then, the array $P(3|3)$ is given by

$$P(3|3) = \begin{bmatrix} & P(3|2) \\ \dfrac{\gamma}{|\epsilon_1 \Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\left(\alpha_2 \left|\det \Phi\right| + \beta \left|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|\right) \left|\det \Phi\right| + \beta \left|\epsilon_1 \Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right|}{|\epsilon_1 \Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\epsilon_2 \Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\left(\alpha_1 \left|\det \Phi\right| + \beta \left|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|\right) \left|\det \Phi\right| + \beta \left|\epsilon_2 \Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right|}{|\epsilon_2 \Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\mathrm{HA}\Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\left(\gamma \left|\det \Phi\right| + \beta \left|\mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|\right) \left|\det \Phi\right| + \beta \left|\mathrm{HA}\Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right|}{|\mathrm{HA}\Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\left(\alpha_1 \left|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right| + \alpha_2 \left|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|\right) \left|\det \Phi\right| + \beta \left|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|}{|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\left(\dfrac{\gamma}{|\epsilon_1 \mathrm{H}^{\mathrm{T}}|} \left|\epsilon_1 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right| + \dfrac{\alpha_2}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \left|\mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|\right) \left|\det \Phi\right| + \beta \left|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|}{|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\left(\dfrac{\gamma}{|\epsilon_2 \mathrm{H}^{\mathrm{T}}|} \left|\epsilon_2 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right| + \dfrac{\alpha_1}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \left|\mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|\right) \left|\det \Phi\right| + \beta \left|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|}{|\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \dfrac{\gamma}{|\mathrm{HA}\Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} & \dfrac{\gamma \left|\det \Phi\right| + \beta \left|\mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma\right|}{|\mathrm{HA}\Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}|} \\ \vdots & \vdots \\ & \text{(continued on next page)} \end{bmatrix} \tag{3.63a}$$

$$
P(3|3) = \begin{bmatrix}
\multicolumn{2}{c}{\text{(continued from previous page)}} \\
\vdots & \vdots \\
\frac{\gamma}{|H\Gamma|} & \frac{\alpha_1 \left|\epsilon_1 \Phi^{T2} A\Gamma\right| + \alpha_2 \left|\epsilon_2 \Phi^{T2} A\Gamma\right| + \beta \left|\Gamma^T \Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|\epsilon_1 H^T|} \left|\epsilon_1 \Phi^{T2} A\Gamma\right| + \frac{\alpha_2}{\epsilon_1 H^T} \left|HA\Phi^{T2} A\Gamma\right| + \beta \left|\Gamma^T \Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|\epsilon_2 H^T|} \left|\epsilon_2 \Phi^{T2} A\Gamma\right| + \frac{\alpha_1}{\epsilon_2 H^T} \left|HA\Phi^{T2} A\Gamma\right| + \beta \left|\Gamma^T \Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|\epsilon_1 \Phi^T H^T|} \left|\epsilon_1 \Phi^{T2} A\Gamma\right| + \left(\frac{\alpha_2 \left|\det \Phi\right| + \beta \left|\epsilon_1 \Phi^T A\Gamma\right|}{|\epsilon_1 \Phi^T H^T|}\right) \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|\epsilon_2 \Phi^T H^T|} \left|\epsilon_2 \Phi^{T2} A\Gamma\right| + \left(\frac{\alpha_1 \left|\det \Phi\right| + \beta \left|\epsilon_2 \Phi^T A\Gamma\right|}{|\epsilon_2 \Phi^T H^T|}\right) \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|HA\Phi^T H^T|} \left|HA\Phi^{T2} A\Gamma\right| + \left(\frac{\gamma \left|\det \Phi\right| + \beta \left|HA\Phi^T A\Gamma\right|}{|HA\Phi^T H^T|}\right) \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|H\Gamma|} \left|\Gamma^T \Phi^T A\Gamma\right| + \left(\frac{\alpha_1 \left|\epsilon_1 \Phi^T A\Gamma\right| + \alpha_2 \left|\epsilon_2 \Phi^T A\Gamma\right|}{|H\Gamma|}\right) \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|H\Gamma|} \left|\Gamma^T \Phi^T A\Gamma\right| + \left(\frac{\frac{\gamma}{|\epsilon_1 H^T|} \left|\epsilon_1 \Phi^T A\Gamma\right| + \frac{\alpha_2}{|\epsilon_1 H^T|} \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|}\right) \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|} \\
\frac{\gamma}{|H\Gamma|} & \frac{\frac{\gamma}{|H\Gamma|} \left|\Gamma^T \Phi^T A\Gamma\right| + \left(\frac{\frac{\gamma}{|\epsilon_2 H^T|} \left|\epsilon_2 \Phi^T A\Gamma\right| + \frac{\alpha_1}{|\epsilon_2 H^T|} \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|}\right) \left|HA\Phi^T A\Gamma\right|}{|H\Gamma|}
\end{bmatrix}. \tag{3.63b}
$$

The array $Z(3|3)$ is given by

$$Z(3|3) = \begin{bmatrix} \multicolumn{2}{c}{Z(3|2)} \\ \dfrac{z(3)}{\epsilon_1 \Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}} & 0 \\ \dfrac{z(3)}{\epsilon_2 \Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}} & 0 \\ \dfrac{z(3)}{\mathrm{H}\mathrm{A}\Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}} & \dfrac{z(1) \cdot (\det \Phi)^2}{\mathrm{H}\mathrm{A}\Phi^{\mathrm{T}2} \mathrm{H}^{\mathrm{T}}} \\ \dfrac{z(3)}{\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & 0 \\ \dfrac{z(3)}{\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & \dfrac{\left(\dfrac{z(1)}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_1 \Phi^{\mathrm{T}} \mathrm{A} \Gamma \right) \det \Phi}{\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} \\ \dfrac{z(3)}{\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & \dfrac{\left(\dfrac{z(1)}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_2 \Phi^{\mathrm{T}} \mathrm{A} \Gamma \right) \det \Phi}{\Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} \\ \dfrac{z(3)}{\mathrm{H}\mathrm{A}\Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} & \dfrac{z(2) \cdot \det \Phi}{\mathrm{H}\mathrm{A}\Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} \\ \vdots & \vdots \\ \multicolumn{2}{c}{\text{(continued on next page)}} \end{bmatrix} \tag{3.64a}$$

$$
Z(3|3) = \begin{bmatrix}
\multicolumn{2}{c}{\text{(continued from previous page)}} \\
\vdots & \vdots \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & 0 \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(1)}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_1 \Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right)}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(1)}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_2 \Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right)}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(2)}{\epsilon_1 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_1 \Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right)}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(2)}{\epsilon_2 \Phi^{\mathrm{T}} \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_2 \Phi^{\mathrm{T}2} \mathrm{A}\Gamma\right)}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(2)}{\mathrm{HA}\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}} \cdot \mathrm{HA}\Phi^{\mathrm{T}2} \mathrm{A}\Gamma + \dfrac{z(1) \cdot \det \Phi}{\mathrm{HA}\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}} \cdot \mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma\right)}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(2)}{\mathrm{H}\Gamma}\right) \cdot \Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{A}\Gamma}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(2)}{\mathrm{H}\Gamma}\right) \cdot \Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{A}\Gamma - \left(\dfrac{-z(1)}{\epsilon_1 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_1 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right) \cdot \mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma}{\mathrm{H}\Gamma} \\
\dfrac{z(3)}{\mathrm{H}\Gamma} & \dfrac{-\left(\dfrac{z(2)}{\mathrm{H}\Gamma}\right) \cdot \Gamma^{\mathrm{T}} \Phi^{\mathrm{T}} \mathrm{A}\Gamma - \left(\dfrac{-z(1)}{\epsilon_2 \mathrm{H}^{\mathrm{T}}} \cdot \epsilon_2 \Phi^{\mathrm{T}} \mathrm{A}\Gamma\right) \cdot \mathrm{HA}\Phi^{\mathrm{T}} \mathrm{A}\Gamma}{\mathrm{H}\Gamma}
\end{bmatrix} \tag{3.64b}
$$

Finally, the array $L_{3|3}$ is constructed from $L_{3|2}$. The i^{th} element in the $L_{3|3}$ array is an integer that equals the number of terms in the i^{th} row of $P(3|3)$, so $L_{3|3}$ is formed by simply appending an array of 2s of length $N^n_{3|3} = 16$.

$$
L_{3|3} = \begin{bmatrix} L^{\mathrm{T}}_{3|2} & 2 & 2 & 2 & \cdots & 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}} \tag{3.65}
$$

3.7 Evaluating the Conditional Mean and Estimation Error Variance

It is shown in Section 2.3 that $\bar{\phi}_{k|k}(\nu)$ is twice continuously differentiable. The mean can be found from the unnormalized characteristic function [1] by taking its partial derivative and then taking the limit as ν goes to the origin.

$$\hat{x}(k) = E\left[x(k)|\mathcal{Z}_k\right] = \frac{1}{jf_{Z_k}(\mathcal{Z}_k)}\left(\frac{\partial\bar{\phi}_{k|k}(\nu)}{\partial\nu}\right)^{\mathrm{T}}\Bigg|_{\nu\to 0}, \tag{3.66a}$$

where

$$f_{Z_k}(\mathcal{Z}_k) = \bar{\phi}_{k|k}(\nu)\big|_{\nu\to 0}. \tag{3.66b}$$

The second moment can be found by taking the same limit of the second partial derivative of the characteristic function, as

$$E\left[x(k)x(k)^{\mathrm{T}}|\mathcal{Z}_k\right] = \frac{-1}{f_{Z_k}(\mathcal{Z}_k)}\cdot\frac{\partial^2\bar{\phi}_{X_k|Z_k}(\nu)}{\partial\nu\nu^{\mathrm{T}}}\Bigg|_{\nu\to 0}, \tag{3.66c}$$

and the error variance is given by

$$\Xi(k) = E\left[x(k)x(k)^{\mathrm{T}}|\mathcal{Z}_k\right] - \hat{x}(k)\hat{x}(k)^{\mathrm{T}}. \tag{3.66d}$$

The limits above must be taken along valid directions, due to the structure of $\bar{\phi}_{k|k}(\nu)$ in (3.41a). Since we are taking a limit, we need only consider a small neighborhood around the origin. Within this neighborhood, ν cannot cross a boundary where $\mu_\ell - \nu = 0$ for all ℓ because then the sign functions would be undefined. Hence, in order to take this limit properly, fix ν to a particular value ν_0 that is not orthogonal to any row of $\mathcal{B}_{k|k}$, i.e. $\mathcal{B}^{\ell}_{k|k}\nu_0 \neq 0$ for all ℓ. Doing so makes the sign functions constant, so that the constant f_{Z_k} is given by

$$f_{Z_k}(\mathcal{Z}_k) = \sum_{i=1}^{N_{k|k}} \mathcal{G}^i_{k|k}(\nu_0), \tag{3.67}$$

the first derivative of $\bar{\phi}_{k|k}(\nu)$ evaluated at the origin is given by

$$\frac{\partial}{\partial\nu}\bar{\phi}_{k|k}(\nu)\Big|_{\nu=0} = \sum_{i=1}^{N_{k|k}}\mathcal{G}^i_{k|k}(\nu_0)\cdot\left[-\sum_{\ell=1}^{L^i_{k|k}} P^\ell_i \operatorname{sgn}\left(\mathcal{B}^{M^\ell_i}_{k|k}\nu_0\right)\mathcal{B}^{M^\ell_i}_{k|k} + j\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{k|k}\right]^{\mathrm{T}}, \tag{3.68}$$

and its second derivative evaluated at the origin is given by

$$\frac{\partial}{\partial \nu}\bar{\phi}_{k|k}(\nu)\Big|_{\nu=0} = \sum_{i=1}^{N_{k|k}} \mathcal{G}^i_{k|k}\left(\nu_0\right) \cdot \left[-\sum_{\ell=1}^{L^i_{k|k}} P^\ell_i \operatorname{sgn}\left(\mathcal{B}^{M^\ell_i}_{k|k}\nu_0\right)\mathcal{B}^{M^\ell_i}_{k|k} + j\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{k|k}\right]^{\mathrm{T}}$$
$$\times \left[-\sum_{\ell=1}^{L^i_{k|k}} P^\ell_i \operatorname{sgn}\left(\mathcal{B}^{M^\ell_i}_{k|k}\nu_0\right)\mathcal{B}^{M^\ell_i}_{k|k} + j\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{k|k}\right]. \tag{3.69}$$

3.8 Finite Horizon Approximation

In order to arrest the growth in computational complexity, we approximate the full information CF with one using a fixed sliding window of the most recent measurements, where number of measurements in this horizon is denoted N_Z. Hence, the first N_Z measurement updates in the estimation are performed normally. Then, for every measurement update $k > N_Z$, we initialize a new finite horizon (FH) estimator and perform N_Z measurement updates over the fixed window $\{z(k-N_Z+1), \ldots, z(k)\}$. This new initial condition for the FH estimator is of the form

$$\phi^k_{W1|0}(\nu) = \exp\left[-\alpha^W_1 \left|\mathcal{B}^1_{W1|0}\nu\right| - \alpha^W_2 \left|\mathcal{B}^2_{W1|0}\nu\right| + j\bar{x}^W_1\nu_1 + j\bar{x}^W_2\nu_2\right],$$
$$\mathcal{B}_{W1|0} = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}, \tag{3.70}$$

where $\mathcal{B}_{W1|0}$ is a rotation matrix. Denote the windowed first measurement update mean and variance by $\hat{x}^W_k(1)$ and $\Xi^W_k(1)$. The exact values of α^W_1, α^W_2, $\bar{x}_1$, $\bar{x}_2$, and φ are determined, given the measurement $z(k-N_Z+1)$, by equating the first two moments of $\bar{\phi}_{k-N_Z+1|k-N_Z+1}$ and of the updated FH first update $\phi^k_{W1|1}$. This process involves solving five nonlinear equations with the five unknowns stated above, which is carried out using standard numerical tools. It is useful here to use the closed form expressions for the first measurement update's mean and variance given in (2.21). It is also necessary to make use of the decomposition in (1.23) in order to apply the proposed algorithm to the initial condition in (3.70).

This local first measurement updated CF, then, has the same mean and variance as the original CF we are approximating. The remaining $N_Z - 1$ measurement updates are performed over the measurements in the window, ultimately producing the N_Z-measurement updated CF $\bar{\phi}^k_{WN_Z|N_Z}$. This CF is taken as the approximation of $\bar{\phi}_{k|k}$, i.e., $\bar{\phi}^k_{WN_Z|N_Z} \approx \bar{\phi}_{k|k}$. Hence, for $k - N_Z + 1 \leq N_Z$ the FH initial condition in (3.70) approximates $\bar{\phi}_{k|k}$, which is conditioned on the entire measurement history. Then, for $k - N_Z + 1 > N_Z$, the FH initial condition approximates the mean and variance of $\bar{\phi}^{k-N_Z+1}_{WN_Z|N_Z}$, produced by a previous iteration of this process.

Numerical comparisons have shown that the local initial condition found in this way performs well in reproducing the full information mean and variance. Moreover, simulations have shown that the finite horizon mean and variances agree very closely with the full information case even with horizon lengths as small as 8, as shown in the next section.

3.9 Numerical Examples

We present a set of four examples demonstrating the performance of our proposed two-state estimator. The main challenge in implementing this estimator is the growth, with each measurement, of the number of terms needed to express the cpdf's CF. The proposed two-state estimator is more efficient and produces far fewer terms than the general-state estimator presented in [14, 15]. The improvement in performance is quantified in the table below, comparing the number of terms in the sum for a two-state implementation of [14, 15] to the number produced by the proposed algorithm, given by (3.52).

Measurement Update k	1	2	3	5	8	10	12
$N_{k\|k}$ of Previous [15]	3	12	51	942	75036	1389207	25719609
$N_{k\|k}$ of Proposed	3	9	25	177	3193	21891	150049
Percent Retained	100%	75%	49%	19%	4.3%	1.6%	0.58%

However, the proposed estimator algorithm still suffers from the same fundamental issue of growing complexity. This motivated the use of a fixed window of the most recent measurements discussed in Section 3.8, the performance of which is discussed next.

3.9.1 Finite Horizon Accuracy

Figure 3.1 shows, on a logarithmic scale, differences between the elements of the estimated state and error variance between the finite-horizon and full-information estimators, normalized to the full-information values; denote this normalized difference of a given element as $e(\cdot)$. The system parameters used are $\beta = 0.5$ or 0.1, $\gamma = 0.1$ or 0.5, $\alpha_1 = \alpha_2 = 0.8$, $\text{eig}(\Phi) = 0.8 \pm 0.55\jmath$, $\text{H} = [1 \;\; 1]$, and $\Gamma = [0.5 \;\; 1]^{\text{T}}$. We compare the performance of horizon lengths of $N_Z = 8$ (dashed lines) and $N_Z = 10$ (solid lines). The subscripts indicate which element of the state estimate vector and error variance matrix are being compared. These results show that this finite horizon approximation is very accurate, with errors approximately between 0.01% and 0.0001% for our example and these two horizon lengths.

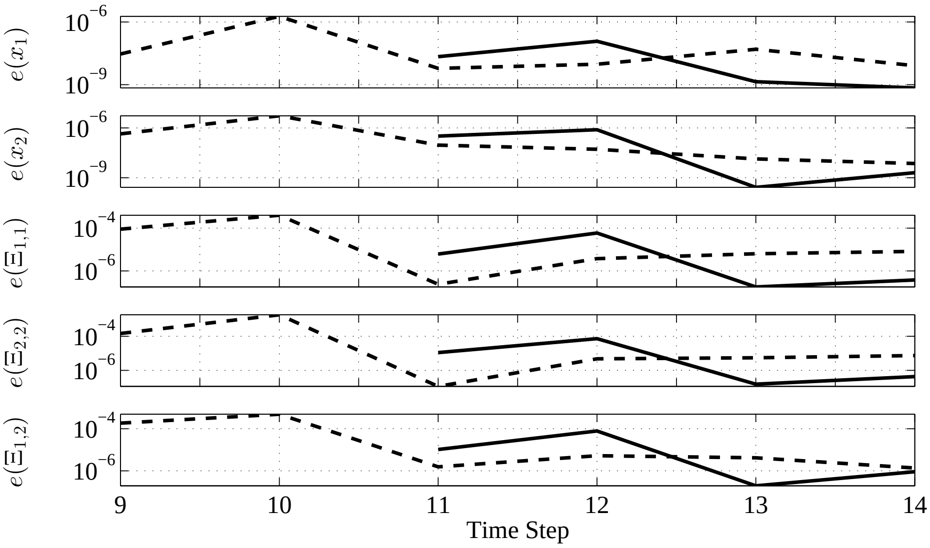


Figure 3.1: Comparison of two finite horizon estimators (dashed is 8, solid is 10) to the full information estimator's means and error variances.

3.9.2 Simulations

The simulations in Figs 3.2, 3.3, and 3.4 all use the same dynamics, where the eigenvalues of the transition matrix are $\text{eig}(\Phi) = 0.8 \pm 0.55\jmath$, $\text{H} = [1 \;\; 1]$, and $\Gamma = [0.5 \;\; 1]^{\text{T}}$. The initial condition has a zero median and $\alpha_1 = \alpha_2 = 0.8$. All simulations use a measurement horizon length of $N_Z = 10$. In Fig. 3.2, $\gamma = 0.5$ and $\beta = 0.1$ so that the measurement noise dominates the process noise; in Fig. 3.3 the parameters are interchanged so that the process noise dominates the measurement noise. Gaussian parameters used for the LEG and for Gaussian noises are closest in the $S\alpha S$ sense to their corresponding Cauchy distributions. For clarity of presentation in the figures, the first update occurs at $k = 0$ instead of $k = 1$.

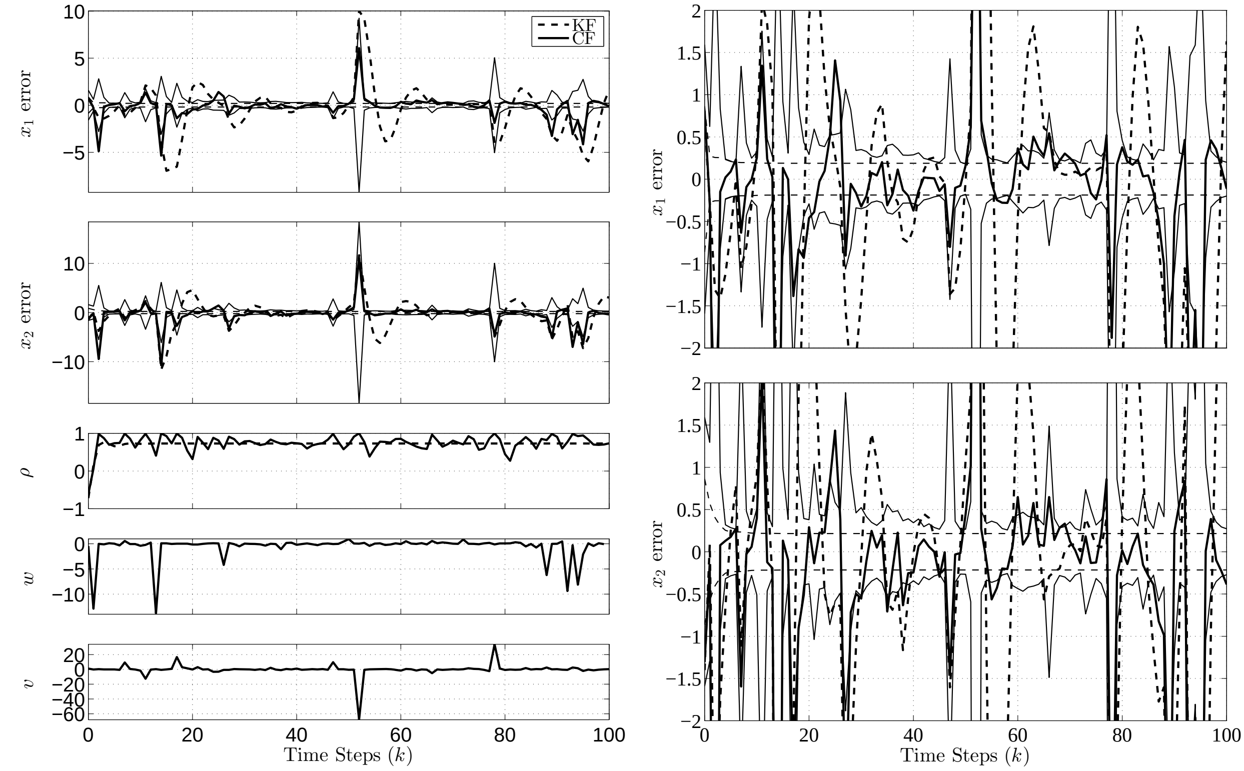
Figures 3.2 and 3.3 compare the Cauchy and Kalman filters' responses to Cauchy distributed noises, and Fig. 3.4 compares their response to Gaussian distributed noises. Figures 3.2(b) and 3.3(b) show the same data as Figs. 3.2(a) and 3.3(a) when zoomed in around zero to demonstrate more clearly how the controllers respond when noise impulses are encountered.

When facing Cauchy distributed noises the proposed estimator outperforms the Kalman filter, especially when the measurement noise dominates the process noise as in Fig. 3.2. The Kalman filter's error is sometimes far from that of the conditional mean generated by the Cauchy estimator, such as when there is an impulsive input. Moreover, these impulsive inputs can cause the conditional variance computed by the Kalman filter to be orders of magnitude smaller than the exact conditional variance computed by the Cauchy estimator.

In Fig. 3.3, where the process noise dominates the measurement noise, the Cauchy and Gaussian filters appear to have similar performance in the state error. However, the Cauchy estimate of the conditional variance is exact and quite different from that of the Kalman filter when the noise environment is impulsive. Moreover, despite their similar state estimate errors, the Cauchy filter much more accurately estimates the conditional variance under Gaussian noise conditions. For both cases presented here, the Kalman filter's closed loop eigenvalues are both complex: for the $\gamma > \beta$ case the eigenvalues are $0.66 \pm 0.56j$, and for the $\beta > \gamma$ case they are $0.31 \pm 0.69j$. The magnitudes of the eigenvalues for the $\beta > \gamma$ case in Fig. 3.3(a) are smaller due to the larger Kalman filter gain, and hence the estimator weighs the measurements more than the state and performs better.

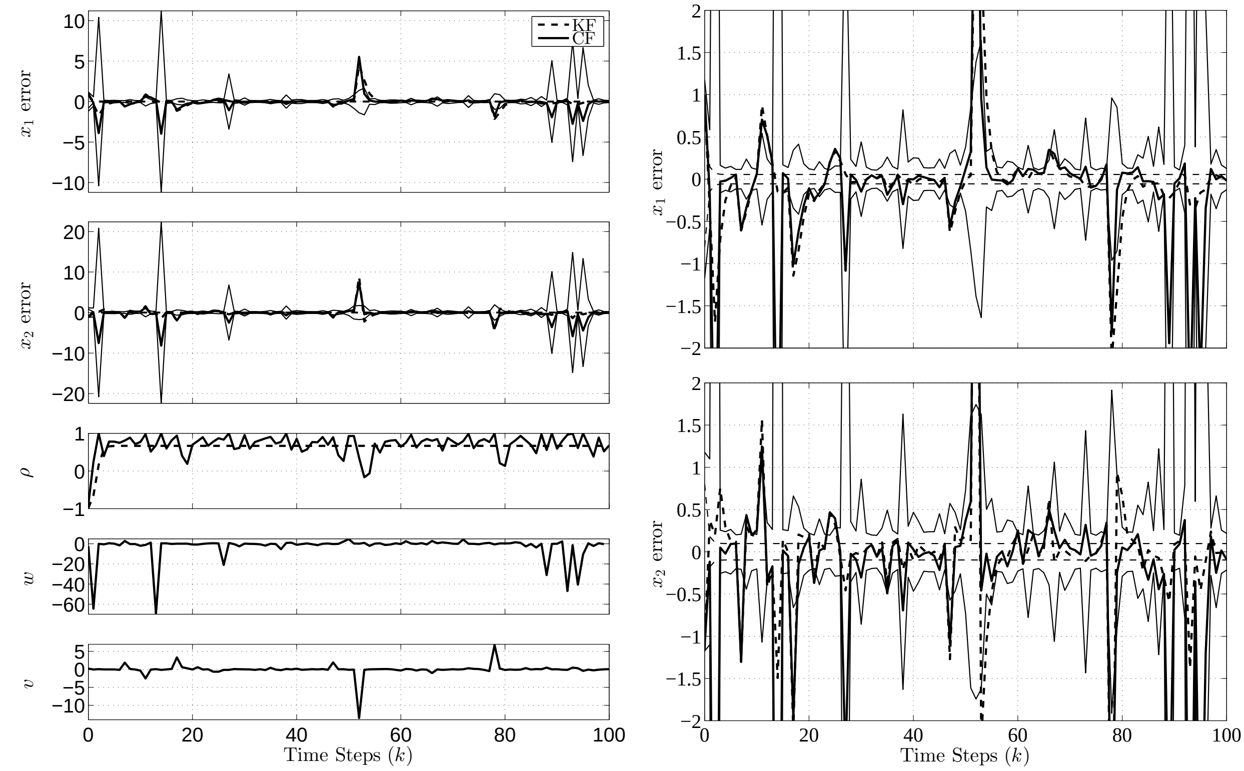
During periods in the simulation without large impulses, the Cauchy and Kalman filters have similar performance, as shown in Figs. 3.2(b) and 3.3(b). This suggests that in a non-impulsive noise setting, the two estimators would have similar performance. In a Gaussian noise setting, shown in Fig. 3.4, the Cauchy filter performs very well. It approximates the variance of the optimal Kalman filter and tracks its mean very closely. For the Cauchy estimator, the optimal conditional variance is shown to be well estimated even in a Gaussian noise environment. This behavior is possible due to the dependence of the Cauchy estimator's error variance on the measurement history.

To better understand the robustness of the Cauchy estimator against non-impulsive, Gaussian noises we performed Monte Carlo experiments, shown in Fig. 3.5. These simulations average 500 independent runs of 101 measurement updates, and used a horizon length of $N_Z = 8$. These simulations fix the Gaussian measurement and process noise parameters, given by V and W, and the Cauchy parameters are closest in the $S\alpha S$ to these Gaussian parameters. These simulations corroborate Fig. 3.4 by showing that, although the Cauchy estimator's estimate error variance is larger than the optimal Kalman filter's, the estimate error and the correlation ρ are very similar. It is interesting to note that when the process noise dominates the measurement noise, the Cauchy estimator's average estimate error variance is larger than when the measurement noise is dominant.



(a) Full view. (b) Zoomed view.

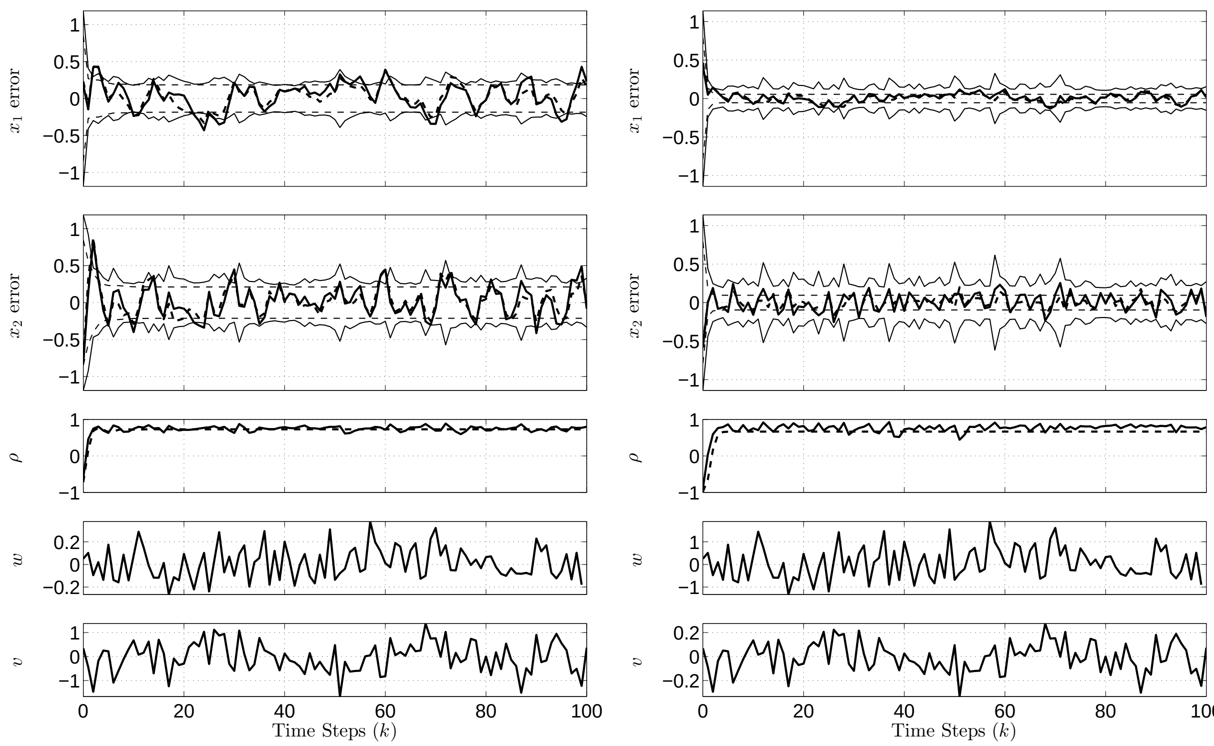
Figure 3.2: Cauchy and Kalman estimators for $\gamma > \beta$; thick lines are the estimate errors, and thin lines are the standard deviations.



(a) Full view.

(b) Zoomed view.

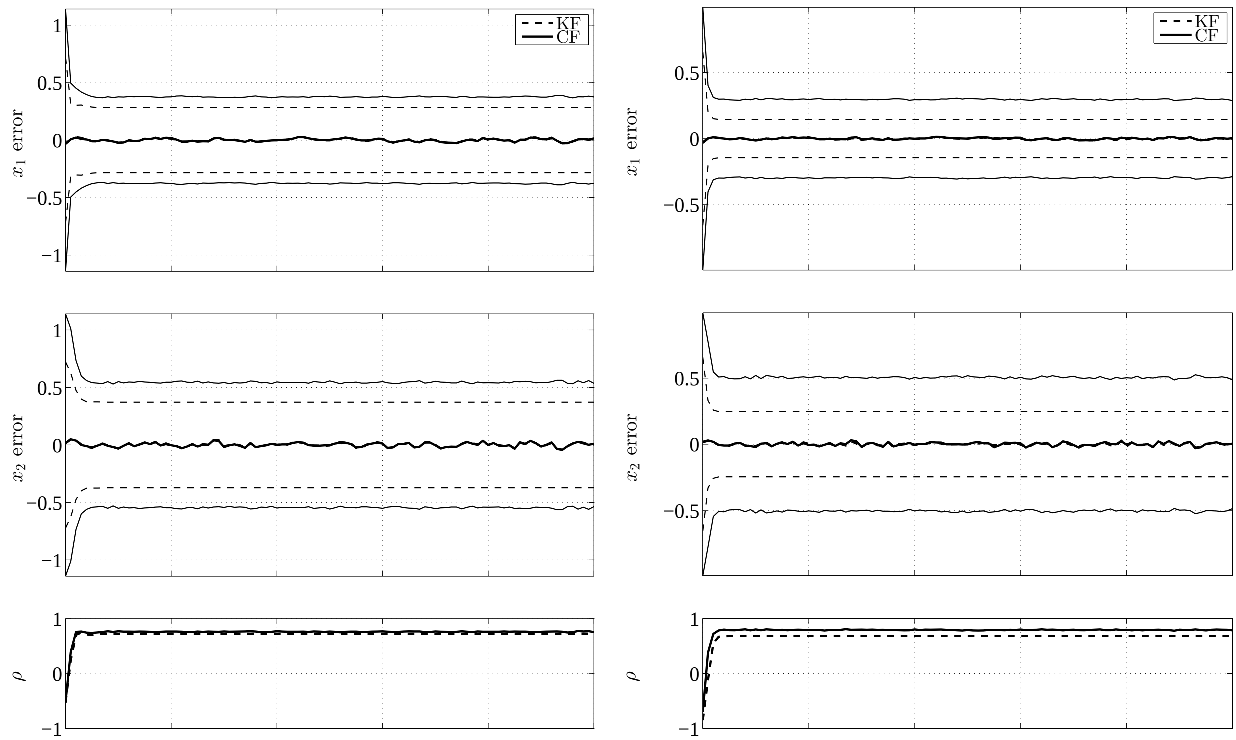
Figure 3.3: Cauchy and Kalman estimators for $\beta > \gamma$; thick lines are the estimate errors, and thin lines are the standard deviations.



(a) Gaussian noises, $\beta = 0.1$ and $\gamma = 0.5$.

(b) Gaussian noises, $\beta = 0.5$ and $\gamma = 0.1$.

Figure 3.4: Cauchy and Kalman estimators against Gaussian noises closest in the $S\alpha S$ sense to the given Cauchy parameters; thick lines are the estimate errors, and thin lines are the standard deviations.



(a) $V = 0.71$ and $W = 0.14$.

(b) $V = 0.14$ and $W = 0.71$.

Figure 3.5: Monte Carlo experiments for the Cauchy and Kalman estimators against Gaussian noises; thick lines are the estimate errors, and thin lines are the standard deviations.

CHAPTER 4

Control with Characteristic Functions

Before his eyes a kaleidoscopic range of phantasmal images played, all of them dissolving at intervals into the picture of a vast, unplumbed abyss of night wherein whirled suns and worlds of an even profounder blackness. He thought of the ancient legends of Ultimate Chaos, at whose centre sprawls the blind idiot god Azathoth, Lord of All Things, encircled by his flopping horde of mindless and amorphous dancers, and lulled by the thin monotonous piping of a demoniac flute held in nameless paws.

H.P. Lovecraft, “The Haunter of the Dark”

The goal of our model predictive control problem is to maximize the conditional expectation of an objective function. This objective function is cast as a product of functions resembling Cauchy pdfs, which are easily transformed into a function of the spectral variables. Consequently, the conditional performance index is found in a closed form. Due to its complexity, the optimal control signal is determined by numerically optimizing this conditional performance index in a model predictive control setting [19]. The main innovation here is evaluating this conditional expectation using the CF directly by applying Parseval’s Identity.

First, the conditional performance index for a general vector state system using the CF structure in [14, 15] is discussed, demonstrating the generality of the method. Then, we assume a two-state system and, using the CF structure from Chapter 3, evaluate the conditional performance index in closed form. It is optimized numerically at each time step, and examples are given to demonstrate its performance.

4.1 Derivation of the Cost Using Characteristic Functions

Our proposed controller is an m-step model predictive controller [29] that uses current and past measurements, and averages over future process noise. At each time step, the conditional performance index is computed. Since the performance index will be shown to be a nonconvex function of the control sequence, it is maximized numerically. Once the optimal control sequence of length m is computed, only the first control in that sequence is applied. At the next step, a new measurement is taken and the process is repeated, producing a new optimal control sequence and applying only the first one. In this paper, we study the optimal stochastic state regulation problem, noting that the tracking problem can be handled in a similar fashion. Our regulation problem will have a finite horizon of length m such that the terminal state occurs at time-step $p = k + m$.

Let the state, measurement, and control histories used in the control problem formulation be defined as

$$\mathcal{X}_{k+1}^{p} := \{x(k+1), \ldots, x(p)\}, \tag{4.1a}$$

$$\mathcal{Z}_k := \{z(1), \ldots, z(k)\}, \tag{4.1b}$$

$$\mathcal{U}_k^{p-1} := \{u(k), \ldots, u(p-1)\},\ \mathcal{U}_k^{p-1} \in \mathcal{F} \tag{4.1c}$$

where $\mathcal{F}$ is the class of piecewise continuous functions adapted to the σ-algebra σ_k generated by the measurement history, i.e. the control is a random variable that is measurable with respect to events in σ_k [30]. The measurement history can be decomposed using (1.23) and (1.24) as $\mathcal{Z}_k = \tilde{\mathcal{Z}}_k + \bar{\mathcal{Z}}_k$, where

$$\tilde{\mathcal{Z}}_k = \{\tilde{z}(1), \cdots, \tilde{z}(k)\}, \quad \bar{\mathcal{Z}}_k = \{\bar{z}(1), \cdots, \bar{z}(k)\}. \tag{4.2}$$

In the following, from [13], it is shown that the control is measurable on events generated by $\tilde{\mathcal{Z}}_k$ only.

Theorem 3. *Consider the filtration $\sigma - algebra$ $\tilde{\sigma}_k$ generated by $\tilde{\mathcal{Z}}_k$, with the decomposition $\mathcal{Z}_k = \tilde{\mathcal{Z}}_k + \bar{\mathcal{Z}}_k$. For $\tilde{\mathcal{Z}}_k \in \tilde{\sigma}_k$ and $\tilde{\sigma}_{k-1} \subset \tilde{\sigma}_k$, $\bar{\mathcal{Z}}_k$ is adapted to $\tilde{\sigma}_{k-1}$ and $u(k)$ is adapted to $\tilde{\sigma}_k$.*

Proof. Start with $k = 0$. The initial state is decomposed as $x(1) = \tilde{x}(1) + \bar{x}(1)$, where $\bar{z}(1)$ is a given non-random parameter. The measurement decomposes as $z(1) = \tilde{z}(1) + \bar{z}(1)$, where $\bar{z}(1) = \mathrm{H}\bar{x}(1)$ is a given non-random parameter and $\tilde{z}(1) = \tilde{\mathcal{Z}}_1 \in \tilde{\sigma}_1$. Then, $u(1)$, which is determined by $z(1)$, is adapted to $\tilde{\sigma}_1$. At $k = 2$, both $\bar{x}(2) = \Phi\bar{x}(1) + u(1)$ and $\bar{z}(2) = \mathrm{H}\bar{x}(2)$ are adapted to $\tilde{\sigma}_1$, and thus $\bar{\mathcal{Z}}_2$ is adapted to $\tilde{\sigma}_1$. For the measurement at $k = 2$, $\tilde{z}(2) \in \tilde{\sigma}_2$, $\tilde{\mathcal{Z}}_2 \in \tilde{\sigma}_2$, and $\tilde{\sigma}_2 \subset \tilde{\sigma}_2$. Hence, since $u(2)$ is determined by $\mathcal{Z}_2 = \tilde{\mathcal{Z}}_2 + \bar{\mathcal{Z}}_2$, it is adapted to $\tilde{\sigma}_2$. Recursively to any k, $\bar{\mathcal{Z}}_k$ is adapted to $\tilde{\sigma}_{k-1}$. With $\tilde{\mathcal{Z}}_k \in \tilde{\sigma}_k$, and $\tilde{\sigma}_{k-1} \subset \tilde{\sigma}_k$, $u(k)$ that is determined by $\mathcal{Z}_k = \tilde{\mathcal{Z}}_k + \bar{\mathcal{Z}}_k$ is adapted to $\tilde{\sigma}_k$. □

Due to the result of Theorem 3, i.e., that the control is adapted to $\tilde{\sigma}_k$, the conditioning on $\mathcal{Z}_k$ can be replaced by $\tilde{\mathcal{Z}}_k$. With these substitutions, the optimization or maximization step of the model predictive control problem is restated as

$$J^*_{\mathcal{Z}_k} = \max_{\mathcal{U}_k^{p-1}} E\left[\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right)\middle| \mathcal{Z}_k\right] = \max_{\mathcal{U}_k^{p-1}} E\left[\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right)\middle| \tilde{\mathcal{Z}}_k\right]$$
$$\triangleq \max_{\mathcal{U}_k^{p-1}} J_{\tilde{\mathcal{Z}}_k} \triangleq J^*_{\tilde{\mathcal{Z}}_k}. \quad (4.3)$$

In the model predictive control operation mode, although the optimal control sequence is determined over the prediction interval from k to p, only the current control input $u(k)$ at time step k is applied to the system. Then, at subsequent time steps, the performance index is maximized again to compute a new optimal control sequence, the first element of which is applied to the system.

Similar to the scalar control problem presented in [12, 13], the control objective function is chosen as a product of functions resembling Cauchy pdfs (1.22), given by

$$\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right) = \prod_{i=k}^{p-1}\left(\frac{\zeta_i/\pi}{u^2(i) + \zeta_i^2} \cdot \prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{x_r^2(i+1) + \eta_{i+1,r}^2}\right). \quad (4.4)$$

Then, the the performance index conditioned on the current measurement history and averaged over future process noises is given by

$$J^*_{k,p} = \max_{\mathcal{U}_k^{p-1} \in \mathcal{F}} E\left[\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right)\right] = \max_{\mathcal{U}_k^{p-1} \in \mathcal{F}} E\left[E\left[\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right)\middle|\tilde{\mathcal{Z}}_k\right]\right]$$
$$= E\left[\max_{\mathcal{U}_k^{p-1} \in \mathcal{F}} E\left[\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right)\middle|\tilde{\mathcal{Z}}_k\right]\right] \triangleq E\left[J^*_{\tilde{\mathcal{Z}}_k}\right], \quad (4.5)$$

where the interchange of the maximum and expectation operations is due to the fundamental theorem in [1].

We are now concerned with determining the analytic form for the conditional performance index $J_{\tilde{\mathcal{Z}}_k}$, where using (4.4) it becomes

$$\begin{aligned} J_{\tilde{\mathcal{Z}}_k} = E\left[\psi\left(\mathcal{X}_{k+1}^p, \mathcal{U}_k^{p-1}\right) \Big| \tilde{\mathcal{Z}}_k\right] &= E\left[\prod_{i=k}^{p-1}\left(\prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2} \cdot \frac{\zeta_i/\pi}{u^2(i)+\zeta_i^2}\right) \Big| \tilde{\mathcal{Z}}_k\right] \\ &= \int_{-\infty}^{\infty} \prod_{i=k}^{p-1}\left(\prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2} \cdot \frac{\zeta_i/\pi}{u^2(i)+\zeta_i^2}\right) \\ &\quad \times f_W(\tilde{x}(p)|\tilde{x}(p-1)) \cdots f_W(\tilde{x}(k+2)|\tilde{x}(k+1)) \\ &\quad \times f_{\tilde{X}_{k+1}|\tilde{\mathcal{Z}}_k}(\tilde{x}(k+1)|\tilde{\mathcal{Z}}_k) d\tilde{x}_1(k+1) \dots d\tilde{x}_n(k+1) \\ &\qquad \times d\tilde{x}_1(k+2) \dots d\tilde{x}_n(k+2) \dots d\tilde{x}_1(p) \dots d\tilde{x}_n(p) \end{aligned} \tag{4.6}$$

For presentation simplicity, in this derivation we will only consider weighting on the terminal state, $x(p)$, and on the m scalar control inputs. The control weighting functions $\mathcal{M}^{\mathcal{U}} \triangleq \prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{u^2(i)+\zeta_i^2}$ can come out of the integral in (4.6). Then, the product over i inside the integrand in (4.6) has only the term for $i = p-1$, and the integral is only over $\{\tilde{x}_1(p), \dots, \tilde{x}_n(p)\}$. Thus, for notational convenience we can drop the time-step index in this integral and write it over $\{\tilde{x}_1, \dots, \tilde{x}_n\}$ as

$$J_{\tilde{\mathcal{Z}}_k} = \mathcal{M}^{\mathcal{U}} \int_{-\infty}^{\infty} \left(\prod_{r=1}^{n} \frac{\eta_{p,r}/\pi}{\left(\tilde{x}_r + \bar{x}_r\right)^2 + \eta_{p,r}^2}\right) \cdot f_{\tilde{X}_p|\tilde{\mathcal{Z}}_k}(\tilde{x}|\tilde{\mathcal{Z}}_k) d\tilde{x}_1 \dots d\tilde{x}_n. \tag{4.7}$$

The cpdf $f_{\tilde{X}_p|\tilde{\mathcal{Z}}_k}(\tilde{x}|\tilde{\mathcal{Z}}_k)$ can be evaluated in closed form for scalar systems [10]. However, for vector state systems only the characteristic function of the cpdf, $\phi_{p|k}(\nu)$, can be evaluated in closed form. Therefore, when computing the conditional performance index, we need to be able to integrate over the spectral variable ν instead of the pdf variable $\tilde{x}$.

Define the product over r in the integral in (4.7) as ℓ_x and its Fourier transform as L_x:

$$\ell_x(\tilde{x} + \bar{x}) = \prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{\left(\tilde{x}_r + \bar{x}_r\right)^2 + \eta_{i+1,r}^2}, \tag{4.8a}$$

$$L_x(\nu) = \prod_{r=1}^{n} e^{-\eta_{p,r}\,|\nu_r| - j\bar{x}_r}. \tag{4.8b}$$

Using these definitions, we can apply Parseval's equation over each variable in (4.7) to express the conditional performance index as an integral over the spectral variable ν,

$$
\begin{aligned}
J_{\tilde{\mathcal{Z}}_k} &= \frac{\mathcal{M}^{\mathcal{U}}}{(2\pi)^n} \int_{-\infty}^{\infty} L_x^*(\nu) \cdot \phi_{p|k}(\nu)\, d\nu_1 \dots d\nu_n \\
&= \frac{\mathcal{M}^{\mathcal{U}}}{(2\pi)^n} \int_{-\infty}^{\infty} \left(\prod_{r=1}^{n} e^{-\eta_{p,r}\,|\nu_r| \,+\, j\bar{x}_r(p)\nu_r} \right) \phi_{p|k}(\nu)\, d\nu_1 \dots d\nu_n, \qquad (4.9)
\end{aligned}
$$

where L_x^* is the complex conjugate of L_x. The next section shows how to evaluate these n nested integrals sequentially in closed form.

4.2 The Conditional Performance Index

Consider the integral over ν_n in (4.9),

$$
\begin{aligned}
I_n &= \int_{-\infty}^{\infty} \left(\prod_{r=1}^{n} e^{-\eta_{p,r}|\langle \epsilon_r, \nu \rangle|} \right) e^{j\langle \bar{x}(p), \nu \rangle} \phi_{p|k}(\nu) d\nu_n \\
&= \int_{-\infty}^{\infty} e^{-\sum_{r=1}^{n} \eta_{p,r}|\langle \epsilon_r, \nu \rangle| + j\langle \bar{x}(p), \nu \rangle} \phi_{p|k}(\nu) d\nu_n, \qquad (4.10)
\end{aligned}
$$

where ϵ_r is the r^{th} column of the n-dimensional identity matrix. The cpdf for the state $\tilde{x}(k)$ is denoted as $f_{\tilde{X}_k|\tilde{Z}_k}$. The unnormalized cpdf (ucpdf) is denoted as $\bar{f}_{\tilde{X}_k|\tilde{\mathcal{Z}}_k} = f_{\tilde{X}_k|\tilde{\mathcal{Z}}_k} \cdot f_{\tilde{\mathcal{Z}}_k}$, where $f_{\tilde{\mathcal{Z}}_k}$ is the pdf of the measurement history and has a known value. In [14, 15, 17, 18], the characteristic function of the ucpdf $\bar{\phi}_{p|k}(\nu)$ is recursively propagated; the characteristic function of the normalized cpdf is $\phi_{k|k}(\nu) = \bar{\phi}_{k|k}(\nu)/f_{\tilde{\mathcal{Z}}_k}$, where $f_{\tilde{\mathcal{Z}}_k} = \bar{\phi}_{\tilde{X}_k|\tilde{\mathcal{Z}}_k}(\nu)|_{\nu=0}$. From [14, 15, 17, 18] the form of the characteristic function of the ucpdf at time k is shown to be

$$
\bar{\phi}_{k|k} = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(y_{gi}(\nu)) e^{y_{ei}^{k|k}(\nu)}, \qquad (4.11a)
$$

where

$$
y_{gi}^{k|k}(\nu) = \sum_{\ell=1}^{n_{ei}^{k|k}} q_{i,\ell}^{k|k} \operatorname{sgn}\left(\left\langle a_{i\ell}^{k|k}, \nu \right\rangle \right) \in \mathbb{R}^k, \qquad (4.11b)
$$

$$
y_{ei}^{k|k}(\nu) = -\sum_{\ell=1}^{n_{ei}^{k|k}} p_{i,\ell}^{k|k} \left| \left\langle a_{i\ell}^{k|k}, \nu \right\rangle \right| + j \left\langle b_i^{k|k}, \nu \right\rangle, \qquad (4.11c)
$$

and the parameters $n_t^{k|k}$, $n_{e,i}^{k|k}$, $q_{i\ell}^{k|k}$, $p_{i\ell}^{k|k}$, $a_{i\ell}^{k|k}$, $b_i^{k|k}$ are generated sequentially from $k = 1$.

For the MPC algorithm, the characteristic function of the ucpdf is to be propagated through the stochastic dynamics to time $k + m = p$ using the propagation formula given in [14,15,17,18]. The characteristic function of the m-step propagated cpdf is denoted $\bar{\phi}_{p|k}(\nu)$ and given by

$$\begin{aligned}
\bar{\phi}_{p|k}(\nu) &= \bar{\phi}_{\tilde{X}_k|\tilde{\mathcal{Z}}_k}(\Phi^{m\mathrm{T}}\nu)\phi_W((\Phi^{m-1}\Gamma)^{\mathrm{T}}\nu)\cdots\phi_W((\Phi\Gamma)^{\mathrm{T}}\nu)\phi_W(\Gamma\nu) \\
&= \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(y_{gi}(\Phi^{m\mathrm{T}}\nu))e^{y_{ei}^{k|k}(\Phi^{m\mathrm{T}}\nu)}\exp\Big(-\beta\left|\left\langle\Phi^{m-1}\Gamma,\nu\right\rangle\right|-\cdots-\beta\left|\langle\Phi\Gamma,\nu\rangle\right|-\beta\left|\langle\Gamma,\nu\rangle\right|\Big).
\end{aligned} \tag{4.12}$$

In (4.12) we effectively add m terms to the sum in $y_{ei}^{k|k}(\nu)$ of (4.11c). By combining the exponent in (4.10) with that in (4.12), the combined exponent in the integrand of (4.10) has a total of $n_{ei}^{k|k}+m+n$ real terms, where the imaginary part is composed of two components. Define the following terms

$$\begin{aligned}
&\bar{p}_{i\ell} = p_{i\ell}, \quad \bar{a}_{i\ell}^{k|k} = \Phi^m a_{i\ell}^{k|k} && \text{for } \ell = 1,\ldots,n_{ei}^{k|k} \\
&\bar{p}_{i\ell} = \beta, \qquad \bar{a}_{i\ell}^{k|k} = \Phi^t\Gamma && \text{for } \ell = 1,\ldots,n_{ei}^{k|k},\ \ t = \ell-(n_{ei}^{k|k}+1) \\
&\bar{p}_{i\ell} = \eta_{p\ell}, \qquad \bar{a}_{i\ell}^{k|k} = \epsilon_r && \text{for } \ell = 1,\ldots,n_{ei}^{k|k},\ \ r = \ell-(n_{ei}^{k|k}+m)
\end{aligned} \tag{4.13a}$$

and

$$\begin{aligned}
\bar{b}_i^{k|k} = b_i^{k|k}+\bar{x}_p, \qquad \bar{n}_{ei}^{k|k} = n_{ei}^{k|k}+m+n \\
q_{t\ell}^{k|k} = 0 \qquad \text{for } \ell = n_{ei}^{k|k}+1,\ldots,\bar{n}_{ei}^{k|k}
\end{aligned} \tag{4.13b}$$

Using these definitions, the integrand in (4.10) becomes

$$\psi(\nu) = \sum_{i=1}^{n_t^{k|k}}\psi_i(\nu) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(\bar{y}_{gi}^{k|k}(\nu))\cdot e^{\bar{y}_{ei}^{k|k}(\nu)}, \tag{4.14a}$$

where

$$\bar{y}_{gi}^{k|k}(\nu) = \sum_{\ell=1}^{\bar{n}_{ei}^{k|k}} q_{i\ell}^{k|k}\,\mathrm{sgn}\left(\left\langle\bar{a}_{i\ell}^{k|k},\nu\right\rangle\right), \tag{4.14b}$$

$$\bar{y}_{ei}^{k|k}(\nu) = \sum_{\ell=1}^{\bar{n}_{ei}^{k|k}} \bar{p}_{i\ell}^{k|k}\left|\left\langle\bar{a}_{i\ell}^{k|k},\nu\right\rangle\right| + j\left\langle\bar{b}_{i\ell}^{k|k},\nu\right\rangle. \tag{4.14c}$$

The integration in (4.10) is performed for each element of ν in turn. Beginning with ν_n, decompose $\nu = \begin{bmatrix} \hat{\nu} & \nu_n \end{bmatrix}^{\mathrm{T}}$ where $\hat{\nu} \in \mathbb{R}^{n-1}$. Then,

$$I_n = \int_{-\infty}^{\infty} \psi(\nu) d\nu = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{i=1}^{n_t^{k|k}} \int_{-\infty}^{\infty} \psi_i(\nu) d\nu = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{i=1}^{n_t^{k|k}} \left[\int_{-\infty}^{\infty} \psi_i(\nu_n, \hat{\nu}) d\nu_n \right] d\hat{\nu}. \tag{4.15}$$

The objective is to reduce the inner integral in (4.15) to a form that is obtained in closed form using the integral formula developed in [14, 15, 17, 18]. First, since $\bar{a}_{i\ell}^{k|k}$ multiplies ν in the sign function in (4.14b) and the absolute value function in (4.14c), they are decomposed as $\bar{a}_{i\ell}^{k|k} = \begin{bmatrix} \hat{a}_{i\ell}^{k|k} & \tilde{a}_{i\ell}^{k|k} \end{bmatrix}^{\mathrm{T}}$, where $\tilde{a}_{i\ell}^{k|k}$ is a scalar and $\hat{a}_{i\ell}^{k|k} \in \mathbb{R}^{n-1}$. Therefore, the inner products in (4.14b) and (4.14c) become

$$\left\langle \bar{a}_{i\ell}^{k|k}, \nu \right\rangle = \left\langle \hat{a}_{i\ell}^{k|k}, \hat{\nu} \right\rangle - \left(-\tilde{a}_{i\ell}^{k|k} \nu_n \right). \tag{4.16}$$

In order to rewrite (4.16) in a form consistent with the integral formula in [14, 15, 17, 18], $-\tilde{a}_{i\ell}^{k|k}$ is factored out of the second term. If $\tilde{a}_{i\ell}^{k|k} = 0$, then the term $e^{\left\langle \hat{a}_{i\ell}^{k|k}, \hat{\nu} \right\rangle}$ loses dependence on ν_n and it is removed from the inner integral in (4.15). Therefore, only $\tilde{a}_{i\ell}^{k|k} \neq 0$ needs to be considered. Let (4.16) be rewritten as

$$\left\langle \bar{a}_{i\ell}^{k|k}, \nu \right\rangle = \left| \tilde{a}_{i\ell}^{k|k} \right| \operatorname{sgn}\left(-\tilde{a}_{i\ell}^{k|k} \right) \left(\mu_{i\ell}^{k|k} - \nu_n \right), \tag{4.17a}$$

where $\mu_{i\ell}^{k|k} = \left\langle \frac{\hat{a}_{i\ell}^{k|k}}{-\tilde{a}_{i\ell}^{k|k}}, \hat{\nu} \right\rangle$. Therefore, the elements in (4.14b) and (4.14c) are

$$q_{i\ell}^{k|k} \operatorname{sgn}\left(\left\langle \bar{a}_{i\ell}^{k|k}, \nu \right\rangle \right) = \bar{q}_{i\ell}^{k|k} \operatorname{sgn}\left(\mu_{i\ell}^{k|k} - \nu_n \right), \tag{4.17b}$$

$$\bar{p}_{i\ell}^{k|k} \left| \left\langle \bar{a}_{i\ell}^{k|k}, \hat{\nu} \right\rangle \right| = \rho_{i\ell}^{k|k} \left| \mu_{i\ell}^{k|k} - \nu_n \right|, \tag{4.17c}$$

where $\bar{q}_{i\ell}^{k|k} = q_{i\ell}^{k|k} \operatorname{sgn}\left(-\tilde{a}_{i\ell}^{k|k} \right)$ and $\rho_{i\ell}^{k|k} = p_{i\ell}^{k|k} \left| \tilde{a}_{i\ell}^{k|k} \right|$. Using these definitions, the inner integral of (4.15) is restated as

$$\int_{-\infty}^{\infty} \psi_i(\nu) d\nu_n =$$

$$e^{j\left\langle \hat{b}_i^{k|k}, \hat{\nu} \right\rangle} \int_{-\infty}^{\infty} g_i^{k|k} \left(\sum_{\ell=1}^{n_{ei}^{k|k}} \bar{q}_{i\ell}^{k|k} \operatorname{sgn}\left(\mu_{i\ell}^{k|k} - \nu_n \right) \right) \exp\left(\sum_{\ell=1}^{\bar{n}_{i\ell}^{k|k}} \rho_{i\ell}^{k|k} \left| \mu_{i\ell}^{k|k} - \nu_n \right| + j \tilde{b}_i^{k|k} \nu_n \right) d\nu_n. \tag{4.18}$$

The convolution integral in (4.18) is shown in [14, 15, 17, 18] to have a closed form solution composed as a sum with $\bar{n}_{i\ell}^{k|k}$ terms, each of which is structurally similar to the terms in $\psi_i(\nu)$. That is, there will be a new g function which is a function of signs of inner products of $\hat{\nu}$.

Therefore, this integration process can be repeated until all of the integrals are taken, and a closed form solution of the conditional performance index is determined. For a vector state system of general order n, a closed form analytical solution is not attainable due to the complexity of the coefficient functions $g_i^{k|k}$. However, a closed form expression for the conditional performance index can be obtained for a two-state system, presented next, by using the structure from [17, 18].

4.3 The Conditional Performance Index for a Second Order System

Now, let us limit our discussion to a second order system in order to use the structure for the cpdf's characteristic function presented in [16,19]. This alternate, two state structure for the characteristic function of the cpdf takes advantage of relationships not yet generalized to the general vector-state case that drastically reduces the number of terms needed to express the characteristic function. In particular, there is a simpler structure for the exponential argument in (4.11c), as well as a simpler, closed form representation for the $g_i^{k|k}$ coefficient functions in (4.11a).

The structure for the characteristic function for the ucpdf is givenin (3.57) as

$$\bar{\phi}_{k|k}(\nu) = \sum_{i=1}^{N_{k|k}} \mathcal{G}_{k|k}^{i}(\nu) \exp\left(-\sum_{\ell=1}^{L_{k|k}^{i}} P_i^{\ell} \left| \mathcal{B}_{k|k}^{M_i^{\ell}} \nu \right| + j \sum_{\ell=1}^{2} \left(Z_i^{\ell} \mathcal{B}_{k|k}^{M_i^{\ell}} \right) \nu \right), \tag{4.19}$$

which consists of a sum of $N_{k|k}$ similar terms. The coefficient functions $\mathcal{G}_{k|k}^{i}(\nu)$ equal the $g_i^{k|k}$

from (4.11a), and are rational functions of polynomials of sums of sign functions given by

$$\mathcal{G}^i_{k|k}(\nu) = \frac{1}{(2\pi)^n}\left\{a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k|k}\nu\right)\operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k|k}\nu\right) + jc_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k|k}\nu\right) + jd_i \operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k|k}\nu\right)\right\}$$
$$\times \prod_{r=1}^{L^i_{k|k}-2}\left\{\frac{1}{j\theta^i_{k-r}+\gamma+\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k}\nu)} - \frac{1}{j\theta^i_{k-r}-\gamma+\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k}\nu)}\right\}, \tag{4.20a}$$

where

$$\mathcal{S}^i_{k-r}(\mathcal{B}_{k|k}\nu) = \sum_{\ell=1}^{L^i_{k-r|k-r}} P^\ell_i\left(\mathcal{B}^{M_i^\ell}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\operatorname{sgn}\left(\mathcal{B}^{M_i^\ell}_{k|k}\nu\right), \tag{4.20b}$$

$$\theta^i_{k-r} = z(k-r) - Z^1_i\mathcal{B}^{M_i^1}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} - Z^2_i\mathcal{B}^{M_i^2}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}. \tag{4.20c}$$

The arguments $y^{k|k}_{gi}(\nu)$ of $g^{k|k}_i$, given in (4.11b), correspond to the $\mathcal{S}^i_{k-r|k-r}(\mathcal{B}_{k|k}\nu)$ defined above.

The state propagation uses the same formula as in (4.12) and [14,15,17,18]. This process affects the exponential argument of $\bar{\phi}_{k|k}$ by adding new terms to the real part and a transformation on the fundamental directions $\mathcal{B}^{M_i^\ell}_{k|k}$. The coefficient functions remain unchanged as polynomials of sign functions, but the arguments of the sign functions have the same transformation as in the exponential argument, so that $\mathcal{G}^i_{p|k}(\nu) = \mathcal{G}^i_{k|k}(\Phi^{m\mathrm{T}}\nu)$. The state propagation does not add any new terms to the sum, so that $N_{p|k} = N_{k|k}$. Hence, the m-step state propagated cpdf's characteristic function is given by

$$\bar{\phi}_{p|k}(\nu) = \sum_{i=1}^{N_{p|k}} \mathcal{G}^i_{p|k}(\nu)\exp\left(-\sum_{\ell=1}^{L^i_{p|k}} P^\ell_i\left|\mathcal{B}^{M_i^\ell}_{p|k}\nu\right| + j\sum_{\ell=1}^{2}\left(Z^\ell_i\mathcal{B}^{M_i^\ell}_{p|k}\right)\nu\right), \tag{4.21}$$

Using this expression for the unnormalized CF, we can derive an expression for the performance index in closed form. The transformed objective function in the performance index (4.8b) is now given by

$$L^*_{x(p)}(\nu) = e^{-\eta_{p,1}|\nu_1| + j\bar{x}_1\nu_1 - \eta_{p2}|\nu_2| + j\bar{x}_2\nu_2} \tag{4.22}$$

where, since k will be a constant through this process, the time subscript of $\bar{x}(p)$ is replaced with element subscripts as $\bar{x}(p) = [\bar{x}_1 \quad \bar{x}_2]$. The 2-state specific version of (4.9) then becomes

$$\begin{aligned} J_{\tilde{\mathcal{Z}}_k} &= \frac{\mathcal{M}^{\mathcal{U}}}{(2\pi)^2} \int_{-\infty}^{\infty} L^*_{x(p)}(\nu_1, \nu_2) \cdot \phi_{p|k}(\nu_1, \nu_2)\, d\nu_1 \nu_2 \\ &= \frac{\mathcal{M}^{\mathcal{U}}}{(2\pi)^2 f_{\tilde{\mathcal{Z}}_k}} \int_{-\infty}^{\infty} \exp\left(-\eta_1 |\nu_1| - \eta_2 |\nu_2| + j\bar{x}_1\nu_1 + j\bar{x}_2\nu_2\right) \\ &\qquad \times \sum_{i=1}^{N_{p|k}} \mathcal{G}^i_{p|k}(\nu) \exp\left[-\sum_{\ell=1}^{L^i_{p|k}} P^\ell_i \left| \mathcal{B}^{M^\ell_i}_{p|k} \nu \right| + j \sum_{\ell=1}^{2} \left(Z^\ell_i \mathcal{B}^{M^\ell_i}_{p|k} \right) \nu \right] d\nu. \end{aligned} \tag{4.23}$$

The integral with respect to ν_2 is now taken. We use a second subscript to denote the individual elements of the M^ℓ_i row of $\mathcal{B}_{p|k}$ as $\mathcal{B}^{M^\ell_i}_{p|k} = \left[\mathcal{B}^{M^\ell_i}_{k,1} \quad \mathcal{B}^{M^\ell_i}_{k,2} \right]$. This allows us to decompose the complex part of the exponential argument as

$$j \sum_{\ell=1}^{2} \left(Z^\ell_i \mathcal{B}^{M^\ell_i}_{p|k} \right) \nu = j \left(\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{k,1} \right) \nu_1 + j \left(\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{k,2} \right) \nu_2, \tag{4.24}$$

and rewrite $J_{\tilde{\mathcal{Z}}_k}$ as

$$\begin{aligned} J_{\tilde{\mathcal{Z}}_k} &= \frac{1}{(2\pi)^2 f_{\tilde{\mathcal{Z}}_k}} \left(\prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{u^2(i) + \zeta_i^2} \right) \\ &\qquad \times \sum_{i=1}^{N_{p|k}} \int_{-\infty}^{\infty} \exp\left(-\eta_1 |\nu_1| + j\bar{x}_1\nu_1 + j \sum_{\ell=1}^{2} \left(Z^\ell_i \mathcal{B}^{M^\ell_i}_{k,1} \right) \nu_1 \right) \\ &\times \int_{-\infty}^{\infty} \mathcal{G}^i_{p|k}(\nu) \cdot \exp\left(-\sum_{\ell=1}^{L^i_{p|k}} P^\ell_i \left| \mathcal{B}^{M^\ell_i}_{p|k} \nu \right| - \eta_2 |\nu_2| + j \sum_{\ell=1}^{2} \left(Z^\ell_i \mathcal{B}^{M^\ell_i}_{k,2} \right) \nu_2 + j\bar{x}_2\nu_2 \right) d\nu_2 d\nu_1. \end{aligned} \tag{4.25}$$

Denote the i^{th} inner integral with respect to ν_2 in (4.25) as $I_{i,2}$:

$$\begin{aligned} I_{i,2} &= \int_{-\infty}^{\infty} \mathcal{G}^i_{p|k}(\nu_1, \nu_2) \\ &\times \exp\left[-\sum_{\ell=1}^{L^i_{p|k}} P^\ell_i \left| \mathcal{B}^{M^\ell_i}_{k,1} \nu_1 + \mathcal{B}^{M^\ell_i}_{k,2} \nu_2 \right| - \eta_2 |0 + \nu_2| + j \left(\sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M^\ell_i}_{k,2} \right) \nu_2 + j\bar{x}_2\nu_2 \right] d\nu_2. \end{aligned} \tag{4.26}$$

The integral $I_{i,2}$ is over the ν_2 variable, but it also contains ν_1 in the absolute value terms in the argument of the exponential and the coefficient function $\mathcal{G}^i_{p|k}$. In order to solve it, we need to use the integral of absolute values method presented in [16]. This method involves defining a set of scalar constants ρ_ℓ and ξ_k^i, as well as scalar variables μ_ℓ that depend on ν_1, and thus are constants in this integration.

If $\mathcal{B}_{k,2}^{M_i^\ell} = 0$, then when this constant multiplies ν_2 the variable disappears, and that term comes out of the integral. This will affect the specific form of second integral over ν_1, but the method presented here can still be used. However, for simplicity of presentation, in this derivation we assume that all the $\mathcal{B}_{k,2}^{M_i^\ell} \neq 0$. Then, we can construct a $\{\mu_\ell\}_1^{L_k^i+1}$ set that transforms $I_{i,2}$ into the integral of absolute values structure. This assumption does not affect generality, because if one of the $\mathcal{B}_{k,2}^{M_i^\ell}$ equals zero for some i, then that absolute value term would not be a function of ν_2 and would come out of $I_{i,2}$ and integrated later.

The set of variables, denoted $\{\mu_\ell\}$, is constructed as

$$\begin{aligned} \mathcal{B}_{k,1}^{M_i^\ell}\nu_1 + \mathcal{B}_{k,2}^{M_i^\ell}\nu_2 &= -\mathcal{B}_{k,2}^{M_i^\ell}\left(\frac{-\mathcal{B}_{k,1}^{M_i^\ell}\nu_1}{\mathcal{B}_{k,2}^{M_i^\ell}} - \nu_2\right) \\ &= -\mathcal{B}_{k,2}^{M_i^\ell}\left(\mu_\ell - \nu_2\right) \end{aligned} \tag{4.27a}$$

where

$$\mu_\ell = \begin{cases} \frac{-\mathcal{B}_{k,1}^{M_i^\ell}\nu_1}{\mathcal{B}_{k,2}^{M_i^\ell}} & \ell \in \{1,\ldots,L_{p|k}^i\} \\ 0 & \ell = L_{p|k}^i + 1. \end{cases} \tag{4.27b}$$

Similarly, for the argument of the exponential, we can construct a set of $\{\rho_\ell\}_1^{L_{p|k}^i+1}$ as

$$\rho_\ell = \begin{cases} P_i^\ell \left|-\mathcal{B}_{k,2}^{M_i^\ell}\right| & \ell \in \{1,\ldots,L_{p|k}^i\} \\ \eta_2 & \ell = L_{p|k}^i + 1 \end{cases} \tag{4.27c}$$

and a scalar number ξ_k^i as

$$\xi_k^i = \bar{x}_2 + \sum_{\ell=1}^{2} Z_i^\ell \mathcal{B}_{k,2}^{M_i^\ell}. \tag{4.27d}$$

The solution to an integral of an exponent of absolute values requires dividing the domain of integration into regions in which the integrand is continuous. Since $\mathcal{G}_{p|k}^i(\nu)$ is piecewise-constant, its discontinuities lie on the boundaries of these regions, and hence $\mathcal{G}_{p|k}^i(\nu)$ is treated as a constant in each integral. In order for the $\mathcal{G}_{p|k}^i(\nu)$ coefficients to be consistent with the form in (4.31), a procedure mirroring that of the measurement update process from (3.24)

[16,17] must be used. This involves rewriting the first bracket term in (4.20a) as a sum of sign functions, using the identity $\operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k|k}\nu\right)\,\operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k|k}\nu\right) = \frac{1}{2}\left(\operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k|k}\nu\right) + \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k|k}\nu\right)\right)^2 - 1$ and the substitutions

$$\begin{aligned} &\bar{\rho}_1 = c_i \operatorname{sgn}\left(-\mathcal{B}^{M_i^1}_{k,2}\right), \quad &&\bar{\rho}_2 = d_i \operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k,2}\right), \quad &&\bar{\rho}_\ell = 0 \text{ for } \ell > 2, \\ &\tilde{\rho}_1 = 1, &&\tilde{\rho}_2 = 1, &&\tilde{\rho}_\ell = 0 \text{ for } \ell > 2. \end{aligned} \tag{4.28}$$

Furthermore, for the rest of the bracket terms in (4.20a) we can rewrite the sums given by $\mathcal{S}_{k-r}$ in (4.20b) by defining constants ${}^{(r)}\check{\rho}_\ell$ as

$${}^{(r)}\check{\rho}_\ell = \begin{cases} P_i^\ell\left(\mathcal{B}^{M_i^\ell}_{k-r|k-r}\mathrm{H}^\mathrm{T}\right)\,\operatorname{sgn}\left(-\mathcal{B}^{M_i^\ell}_{k,2}\right) & \ell = 1,\ldots,L^i_{p|k} - r \\ 0 & \ell = L^i_{p|k} - r + 1,\ldots,L^i_{p|k} + 1. \end{cases} \tag{4.29}$$

in order to write $\mathcal{G}^i_{p|k}$ in (4.26) as

$$\begin{aligned} \mathcal{G}^i_{p|k}(\nu - \mathrm{H}^\mathrm{T}\sigma) = \mathcal{G}^i_{p|k}&\left(\sum_{\ell=1}^{L^i_{p|k}+1} \tilde{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \sum_{\ell=1}^{L^i_{p|k}+1} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \ldots\right. \\ &\left.\ldots, \sum_{\ell=1}^{L^i_{p|k}+1} {}^{(L^i_{p|k}-2)}\check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \ldots, \sum_{\ell=1}^{L^i_{p|k}+1} {}^{(1)}\check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\right) \\ &:= \mathcal{G}^i_{p|k}\left(\sum_{\ell=1}^{L^i_{p|k}+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\right). \end{aligned} \tag{4.30}$$

Here, the variable $\check{\rho}_\ell$ is a shorthand variable representing all of the ${}^{(r)}\check{\rho}_\ell$, $\tilde{\rho}_\ell$, and $\bar{\rho}_\ell$ constants in all of the sums in $\mathcal{G}^i_{p|k}$.

Then, let $\sigma := \nu_2$ in order to write the one-dimensional integral as

$$I_{i,2} = \int_{-\infty}^{\infty} \mathcal{G}^i_{p|k}\left(\sum_{\ell=1}^{L^i_{p|k}+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\right) \exp\left(-\sum_{\ell=1}^{L^i_{p|k}+1} \rho_\ell\,|\mu_\ell - \sigma| + j\xi^i_k\sigma\right) d\sigma. \tag{4.31}$$

The solution is given as a sum of $L^i_{p|k}+1$ terms as

$$I_{i,2}(\nu_1) = \sum_{m=1}^{L^i_{p|k}+1} \exp\left(-\sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \rho_\ell \left|\mu_\ell-\mu_m\right| + j\xi^i_k\mu_m \right)$$
$$\times \left\{ \frac{\mathcal{G}^i_{p|k}\left(+\check{\rho}_m + \sum\limits_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \check{\rho}_\ell \ \mathrm{sgn}\left(\mu_\ell-\mu_m\right)\right)}{j\xi^i_k + \rho_m + \sum\limits_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \rho_\ell \ \mathrm{sgn}\left(\mu_\ell-\mu_m\right)} - \frac{\mathcal{G}^i_{p|k}\left(-\check{\rho}_m + \sum\limits_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \check{\rho}_\ell \ \mathrm{sgn}\left(\mu_\ell-\mu_m\right)\right)}{j\xi^i_k - \rho_m + \sum\limits_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \rho_\ell \ \mathrm{sgn}\left(\mu_\ell-\mu_m\right)} \right\}. \tag{4.32a}$$

where

$$\mathcal{G}^i_{p|k}\left(\pm\check{\rho}_m + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \check{\rho}_\ell \ \mathrm{sgn}\left(\mu_\ell-\mu_m\right)\right) =$$
$$\left\{ a_i + b_i\left(\frac{1}{2}\Big(\pm\tilde{\rho}_m + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \tilde{\rho}_\ell \,\mathrm{sgn}\left(\mu_\ell-\mu_m\right)\Big)^2 - 1\right) + j\left(\pm\bar{\rho}_m + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{p|k}+1} \bar{\rho}_\ell \,\mathrm{sgn}\left(\mu_\ell-\mu_m\right)\right)\right\}$$
$$\times \prod_{p=1}^{L^i_{k|k}-2} \frac{1}{2\pi}\left\{ \frac{1}{j\theta^i_{k-r}+\gamma+\mathcal{S}^i_{k-r}(\pm^{(r)}\check{\rho}_m, \mu_\ell-\mu_m)} - \frac{1}{j\theta^i_{k-r}-\gamma+\mathcal{S}^i_{k-r}(\pm^{(r)}\check{\rho}_m, \mu_\ell-\mu_m)}\right\}, \tag{4.32b}$$

and for $r \in \{1,\ldots,L^i_{k|k}-2\}$,

$$\begin{aligned}
&\mathcal{S}^i_{k-r}(\pm^{(r)}\check{\rho}_m, \mu_\ell-\mu_m)\\
&\quad = \pm P^m_i\left(\mathcal{B}^{M^m_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right) + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k-r|k-r}} P^\ell_i\left(\mathcal{B}^{M^\ell_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\mathrm{sgn}\left(-\mathcal{B}^{M^\ell_i}_{k,2}\right)\mathrm{sgn}\left(\mu_\ell-\mu_m\right)\\
&\quad = \pm P^m_i\left(\mathcal{B}^{M^m_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\\
&\qquad + \sum_{\substack{\ell=1\\ \ell\neq m}}^{L^i_{k-r|k-r}} P^\ell_i\left(\mathcal{B}^{M^\ell_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\mathrm{sgn}\left(-\mathcal{B}^{M^\ell_i}_{k,2}\cdot\left(-\frac{\mathcal{B}^{M^\ell_i}_{k,1}}{\mathcal{B}^{M^\ell_i}_{k,2}} + \frac{\mathcal{B}^{M^m_i}_{k,1}}{\mathcal{B}^{M^m_i}_{k,2}}\right)\right)\cdot\mathrm{sgn}\left(\nu_1\right).
\end{aligned} \tag{4.32c}$$

In the two-state estimator [17], algebraic relationships are used to reduce complicated bracket terms produced by the update process into simpler, polynomial forms. Similarly, those same algebraic relationships can be used here to simplify (4.32a). Denote the coefficient of the m^{th} term in the sum as $\mathcal{G}_{p|k}^{i,m}$, consider the numerators of the coefficient function in the brackets in (4.32a) given by (4.32b). This complicated term can be reduced to a simple function of $\operatorname{sgn}(\nu_1)$. This process begins by rewriting the numerator functions (4.32b) in a four parameter form involving only $\operatorname{sgn}(\nu_1)$ as

$$a_i^m \pm b_i^m \cdot \operatorname{sgn}(\nu_1) \pm j c_i^m + j d_i^m \cdot \operatorname{sgn}(\nu_1). \tag{4.33}$$

Then, the real part of the exponential argument in (4.32a) and the sums in the denominators of the coefficient function in (4.32b) can be reduced to a single constant times $|\nu_1|$, where that constant is defined as

$$D_{i,m}\,|\nu_1| = \sum_{\substack{\ell=1\\ \ell\neq m}}^{L_{p|k}^i+1} \rho_\ell\,|\mu_\ell - \mu_m| = \left(\sum_{\substack{\ell=1\\ \ell\neq m}}^{L_{p|k}^i+1} \rho_\ell \left|\frac{-\mathcal{B}_{k,1}^{M_i^\ell}}{\mathcal{B}_{k,2}^{M_i^\ell}} - \frac{-\mathcal{B}_{k,1}^{M_i^m}}{\mathcal{B}_{k,2}^{M_i^m}}\right|\right)|\nu_1| \tag{4.34}$$

$$\check{D}_{i,m}\operatorname{sgn}(\nu_1) = \sum_{\substack{\ell=1\\ \ell\neq m}}^{L_{p|k}^i+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m) = \left(\sum_{\substack{\ell=1\\ \ell\neq m}}^{L_{p|k}^i+1} \rho_\ell \operatorname{sgn}\left(\frac{-\mathcal{B}_{k,1}^{M_i^\ell}}{\mathcal{B}_{k,2}^{M_i^\ell}} - \frac{-\mathcal{B}_{k,1}^{M_i^m}}{\mathcal{B}_{k,2}^{M_i^m}}\right)\right)\operatorname{sgn}(\nu_1) \tag{4.35}$$

Expressing the numerators of (4.32b) as (4.33), and using $\check{D}_{i,m}$, then Result 4 from Appendix A can be used to express the coefficient function as a two parameter form, where those parameters are denoted $a_{i,m}$ and $d_{i,m}$. For the exponential argument, use $D_{i,m}$ and rewrite the complex part as

$$\hat{\xi}_k^i \nu_1 = \left(\xi_k^i \cdot \frac{-\mathcal{B}_{k,1}^{M_i^\ell}}{\mathcal{B}_{k,2}^{M_i^\ell}}\right) \cdot \nu_1 \tag{4.36}$$

in order to express (4.32a) in a much simpler form given by

$$I_{i,2}(\nu_1) = \sum_{m=1}^{L_{p|k}^i+1} \{a_{i,m} + j\, d_{i,m} \operatorname{sgn}(\nu_1)\} \cdot e^{-D_{i,m}\,|\nu_1| + j\,\hat{\xi}_k^i \nu_1}. \tag{4.37}$$

In order to obtain the form in (4.33) for (4.32b), a procedure mirroring that of the estimator update process is applied, which uses Results 1 and 2 in Appendix A. For $m = 1$ and $m = 2$, the numerators of (4.32a) are not equal and hence cannot come out of the bracket term. Then, the numerators for $m = 1$ are given by

$$
\begin{aligned}
a_i + b_i \left(\frac{1}{2} \Bigg(\pm \tilde{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{p|k}+1} \tilde{\rho}_\ell \operatorname{sgn}(\mu_2 - \mu_1) \Bigg)^2 - 1 \right) + j \left(\pm \bar{\rho}_1 + \sum_{\ell=2}^{L^i_{p|k}+1} \bar{\rho}_\ell \operatorname{sgn}(\mu_2 - \mu_1) \right) \\
= a_i \pm b_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k,2} \cdot \mathcal{B}^{M_i^2}_{k,2} \right) \operatorname{sgn}(\mu_2 - \mu_1) \\
\pm j c_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k,2} \right) + j d_i \operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k,2} \right) \operatorname{sgn}(\mu_2 - \mu_1) \\
= a_i \pm b_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k,2} \cdot \mathcal{B}^{M_i^2}_{k,2} \right) \operatorname{sgn}\left(-\frac{\mathcal{B}^{M_i^2}_{k,1}}{\mathcal{B}^{M_i^2}_{k,2}} + \frac{\mathcal{B}^{M_i^1}_{k,1}}{\mathcal{B}^{M_i^1}_{k,2}} \right) \operatorname{sgn}(\nu_1) \\
\pm j c_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k,2} \right) + j d_i \operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k,2} \right) \operatorname{sgn}\left(-\frac{\mathcal{B}^{M_i^2}_{k,1}}{\mathcal{B}^{M_i^2}_{k,2}} + \frac{\mathcal{B}^{M_i^1}_{k,1}}{\mathcal{B}^{M_i^1}_{k,2}} \right) \operatorname{sgn}(\nu_1) .
\end{aligned}
\tag{4.38}
$$

The same manipulations can be done for the numerators for $m = 2$ by interchanging μ_1 and c_i with μ_2 and d_i, respectively. The $+$ and $-$ in (4.38) refer to the left and right numerators, respectively, in (4.32a) that are given by (4.32b).

Hence, since the first bracket term in (4.32b) can be expressed as (4.38), it can be combined with a bracket term from the product in (4.32b) by substituting in (4.32c). This product is given in Result 2 and produces a term of the same form as (4.38). Hence, this result can be combined with the next bracket term in the product in (4.32c), eventually expressing both numerators as (4.33).

For $m > 2$, this first bracket term produces a simpler form due to two sign functions canceling each other out. Recalling that $\bar{\rho}_\ell = \tilde{\rho}_\ell = 0$ for $\ell > 2$, the new term obtained for

$m = 3$ corresponds to the coefficient

$$
\begin{aligned}
a_i + b_i &\left(\frac{1}{2} \Big(\pm \tilde{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L^i_{p|k}+1} \tilde{\rho}_\ell \operatorname{sgn}(\mu_1 - \mu_3) \Big)^2 - 1 \right) + j \left(\pm \bar{\rho}_1 + \sum_{\ell=2}^{L^i_{p|k}+1} \bar{\rho}_\ell \operatorname{sgn}(\mu_2 - \mu_3) \right) \\
&\qquad = a_i + b_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k,2} \cdot \mathcal{B}^{M_i^2}_{k,2} \right) \operatorname{sgn}(\mu_1 - \mu_3) \operatorname{sgn}(\mu_2 - \mu_3) \\
&\qquad\quad + j c_i \operatorname{sgn}\left(\mathcal{B}^{M_i^1}_{k,2} \right) \operatorname{sgn}(\mu_1 - \mu_3) + j d_i \operatorname{sgn}\left(\mathcal{B}^{M_i^2}_{k,2} \right) \operatorname{sgn}(\mu_2 - \mu_3) \\
&= \left[a_i + b_i \operatorname{sgn}\left(-\frac{\mathcal{B}^{M_i^1}_{k,1}}{\mathcal{B}^{M_i^1}_{k,2}} + \frac{\mathcal{B}^{M_i^3}_{k,1}}{\mathcal{B}^{M_i^3}_{k,2}} \right) \operatorname{sgn}\left(-\frac{\mathcal{B}^{M_i^2}_{k,1}}{\mathcal{B}^{M_i^2}_{k,2}} + \frac{\mathcal{B}^{M_i^3}_{k,1}}{\mathcal{B}^{M_i^3}_{k,2}} \right) \right] \\
&\qquad + j \left[c_i \operatorname{sgn}\left(-\frac{\mathcal{B}^{M_i^1}_{k,1}}{\mathcal{B}^{M_i^1}_{k,2}} + \frac{\mathcal{B}^{M_i^3}_{k,1}}{\mathcal{B}^{M_i^3}_{k,2}} \right) + d_i \operatorname{sgn}\left(-\frac{\mathcal{B}^{M_i^2}_{k,1}}{\mathcal{B}^{M_i^2}_{k,2}} + \frac{\mathcal{B}^{M_i^3}_{k,1}}{\mathcal{B}^{M_i^3}_{k,2}} \right) \right] \operatorname{sgn}(\nu_1). \qquad (4.39)
\end{aligned}
$$

Denote the outer integral with respect to ν_1 in (4.25) by

$$
\begin{aligned}
I_{i,1} = &\sum_{m=1}^{L^i_{p|k}+1} \int_{-\infty}^{\infty} \{ a_{i,m} + j\, d_{i,m} \operatorname{sgn}(\nu_1) \} \\
&\times \exp\left(-D_{i,m} |\nu_1| + j \hat{\xi}^i_k \nu_1 \right) \exp\left(-\eta_1 |\nu_1| + j \bar{x}_1 \nu_1 \right) \exp\left(j \sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M_i^\ell}_{k,1} \right) d\nu_1 \\
= &\sum_{m=1}^{L^i_{p|k}+1} \int_{-\infty}^{\infty} \{ a_{i,m} + j\, d_{i,m} \operatorname{sgn}(\nu_1) \} \\
&\times \exp\left(-\Big[D_{i,m} + \eta_1 \Big] |\nu_1| + j \left[\hat{\xi}^i_k + \bar{x}_1 + \sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M_i^\ell}_{k,1} \right] \nu_1 \right) d\nu_1. \qquad (4.40)
\end{aligned}
$$

This integral has a form identical to the measurement update process for a scalar system [10]. Its solution is given by

$$
\begin{aligned}
I_{i,1} &= \sum_{m=1}^{L^i_{p|k}+1} \frac{a_{i,m} - j\, d_{i,m}}{j \left[\hat{\xi}^i_k + \bar{x}_1 + \sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M_i^\ell}_{k,1} \right] + \Big[D_{i,m} + \eta_1 \Big]} - \frac{a_{i,m} + j\, d_{i,m}}{j \left[\hat{\xi}^i_k + \bar{x}_1 + \sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M_i^\ell}_{k,1} \right] - \Big[D + \eta_1 \Big]} \\
&= \sum_{m=1}^{L^i_{p|k}+1} 2 \cdot \frac{a_{i,m} \Big[D_{i,m} + \eta_1 \Big] - d_{i,m} \left[\bar{x}_1 + \sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M_i^\ell}_{k,1} \right]}{\left[\hat{\xi}^i_k + \bar{x}_1 + \sum_{\ell=1}^{2} Z^\ell_i \mathcal{B}^{M_i^\ell}_{k,1} \right]^2 + \Big[D_{i,m} + \eta_1 \Big]^2}. \qquad (4.41)
\end{aligned}
$$

Finally, using (4.41) in (4.25), the conditional performance index is given by

$$J_{\tilde{\mathcal{Z}}_k} = \frac{1}{2\pi^2 f_{\tilde{\mathcal{Z}}_k}} \left(\prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{u^2(i)+\zeta_i^2} \right) \sum_{i=1}^{N_{p|k}} \sum_{m=1}^{L^i_{p|k}+1} \left(\frac{a_{i,m}\left[D_{i,m}+\eta_1\right] - d_{i,m}\left[\bar{x}_1 + \sum_{\ell=1}^{2} Z_i^{\ell} \mathcal{B}_{k,1}^{M_i^{\ell}}\right]}{\left[\hat{\xi}_k^i + \bar{x}_1 + \sum_{\ell=1}^{2} Z_i^{\ell} \mathcal{B}_{k,1}^{M_i^{\ell}}\right]^2 + \left[D_{i,m}+\eta_1\right]^2} \right). \tag{4.42}$$

This closed form conditional performance index was tested numerically and found to be non-convex and depends on the control input sequence $\{u(k),\ldots,u(p-1)\}$ in a complex way. Specifically, $\begin{bmatrix}\bar{x}_1 & \bar{x}_2\end{bmatrix}^{\mathrm{T}} = \bar{x}(p)$ depends on the control sequence, along with the parameters $a_{i,m}$ and $d_{i,m}$ that also depend on the measurement sequence $\tilde{\mathcal{Z}}_k$. Therefore, we maximize (4.42) numerically. The optimization is done in two steps: first, the global optimum of the conditional performance index, expressed as a double sum in (4.42), is optimized with respect to the control input with a coarse gradient search in order to find an approximate optimal control sequence. Then, we use the accelerated gradient search method [31] starting from that approximate sequence in order to refine the optimal control solution used.

4.4 Linear Exponential Gaussian Model Predictive Control

The numerical examples presented in the next section compare our proposed Cauchy model predictive controller with an analogous linear exponential Gaussian model predictive controller. Here, we present the derivation and equations for the LEG MPC. Denote the conditional performance index for the LEG MPC as

$$J^G_{\tilde{\mathcal{Z}}_k} = E\left[\prod_{i=k}^{p-1} \left(\exp\left(-\frac{1}{2} r_i \cdot u^2(i)\right) \cdot \prod_{r=1}^{n} \exp\left(-\frac{1}{2} q_{i+1,r} \cdot x_r^2(i+1)\right)\right) \middle| \tilde{\mathcal{Z}}_k \right], \tag{4.43}$$

where $q_{i+1,r}$ are the weights for the states and r_i are the weights for the control inputs. As in the Cauchy MPC problem, we will only penalize the terminal state $x(p)$. Then, in order to be consistent with the Cauchy MPC, these weights are constructed as

$$q_{p,r} = \frac{1}{\kappa^2 \eta_{p,r}^2}, \tag{4.44a}$$

$$r_i = \frac{1}{\kappa^2 \zeta_i^2}, \quad i \in \{k,\ldots,p-1\}. \tag{4.44b}$$

Note that if there is no weighting on the control inputs, then $r_i = 0 \, \forall \, i$.

Then, the optimal control sequence $u_G^* \triangleq \{u(k), \dots, u(p-1)\}$ is given by

$$u_G^* = -\left[S^{\mathrm{T}}\left(M_g^{-1} - M_g^{-1}\left(Q + M_g^{-1}\right)M_g^{-1}\right)S + R\right]^{-1} \times S\left(M_g^{-1} - M_g^{-1}\left(Q + M_g^{-1}\right)M_g^{-1}\right)\Phi^m \hat{x}(k) \quad (4.45)$$

where $\hat{x}(k)$ is the Kalman filter's state estimate conditioned on $\tilde{\mathcal{Z}}_k$, M_g is the m-step propagated error variance of the state estimate $\hat{x}(k)$, S is a matrix given by

$$S = \begin{bmatrix} \Phi^{m-1}\Lambda, & \dots, & \Phi\Lambda, & \Lambda \end{bmatrix}, \quad (4.46)$$

and Q and R are diagonal matrices constructed as

$$Q = \operatorname{diag}\begin{bmatrix} q_{p,1}, & \dots, & q_{p,n} \end{bmatrix}, \quad (4.47a)$$

$$R = \operatorname{diag}\begin{bmatrix} r_k, & \dots, & r_{p-1} \end{bmatrix}. \quad (4.47b)$$

4.5 Numerical Examples

Here, we present two sets of examples, the first of which shows the optimal control versus the measurement for the first time step only, and the second set shows two multi-step examples. All of these examples use a two-step horizon, i.e. $m = 2$, so that there exists a control sequence that can drive our two-state system to the origin over this horizon length. However, as we are using model predictive control, only the first control input of this sequence is applied at that time step. Gaussian parameters used for the LEG and for Gaussian noises are closest in the $S\alpha S$ sense to their corresponding Cauchy distributions. For clarity of presentation in the figures, the first update occurs at $k = 0$ instead of $k = 1$. Furthermore, since the control weightings remain constant across time steps, we will drop the time notation from the parameters η and ζ.

All the simulations use the same system dynamics with $\mathrm{H} = [1 \quad 1]$, $\Gamma^{\mathrm{T}} = [0.5 \quad 1]$, $\Lambda^{\mathrm{T}} = [0.5 \;\; 1]$, $m = 2$, and the eigenvalues of Φ are $0.8 \pm 0.55j$. The terminal state weightings are $\eta = [1 \quad 1]$, and when control weightings are used they are $\zeta = [10 \quad 10]$. The initial condition's scaling parameters are given by $\alpha = [0.8 \;\; 0.8]$, and the process and measurement noise parameters β and γ are either 0.5 or 0.1.

All of our examples compare our Cauchy optimal model predictive controller with a similar LEG model predictive optimal controller [1]. The LEG estimator assumes that the stochastic inputs are described by the Gaussian pdfs that are closest, in a least squared sense, to the given Cauchy pdfs. The LEG controller assumes that its objective functions of the state and control resemble scaled Gaussian pdfs that are closest, in the least-squared sense, to the scaled Cauchy pdfs in (4.4). The LEG design details can be found in [13,15–18], where similar comparisons and least square fits were used. The LEG controllers' responses are shown in dashed lines in the figures.

4.5.1 One-Step Control Examples

The first set of examples are shown in Fig. 4.1. These figures show the applied optimal control input at the first time step given the first measurement. In the two cases presented, all the systems parameters are the same, except in Fig. 4.1(a) $\gamma > \beta$ (i.e. more measurement than state uncertainty), and in Fig. 4.1(b) $\beta > \gamma$ (i.e. more state than measurement uncertainty).

The example in Fig. 4.1(a) shows that the Cauchy controller is nearly linear for small measurements and reduces its control effort to zero as the measurement deviations become large. This is in contrast to the LEG controller, which is linear and thus responds strongly to large measurement deviations. This behavior in the Cauchy controller occurs when the measurement uncertainty is larger than the state uncertainty. In the opposite case shown in Fig. 4.1(b), the measurement has less uncertainty than the state. Here, the Cauchy controller's response closely matches that of the LEG in a neighborhood of the origin, and in fact responds even more strongly than the LEG for large measurement deviations.

The three different curves in both of these figures depict the control signals for three different control weights: no control weight, $\zeta = 10$, and $\zeta = 5$. As expected, heavier control weights (i.e. smaller ζ) reduce the control effort. Even without any control weighting, the response in Fig. 4.1(a) goes to zero for large measurement deviations. The fact that this behavior is seen when there is no control weighting implies that the attenuation of the control signal for large measurement deviations is due to the cpdf and not the objective function. Moreover, this behavior is not shared by the LEG controller that uses a similar objective function but assumes light-tailed, Gaussian distributions.

4.5.2 Multi-Step Control Examples

The complexity of evaluating the conditional performance index grows as the number of terms increases across time steps, as indicated in (4.42). For implementable control, this growth needs to be arrested. The full information characteristic function of the ucpdf (4.19) is approximated by a characteristic function of a ucpdf conditioned on a fixed sliding window of the most recent measurements, as described in Section 3.8, where the number of measurements in this horizon is denoted N_Z. Hence, the first N_Z measurement updates are performed normally, and the control optimization is performed using this CF. Then, for time steps $k > N_Z$, we initialize a new finite horizon (FH) estimator and perform N_Z measurement updates over the fixed window $\{z(k - N_Z + 1), \ldots, z(k)\}$. Then, evaluating the conditional performance index (4.9) uses this FH CF.

The remaining examples in Figs 4.2, 4.3, 4.5, 4.4, and 4.6 are all multistep examples over 100 measurements, and all use a horizon length of $N_Z = 8$. The variations between the simulations are in the stochastic parameters, alternating which of process or measurement noise dominates the other and whether the noises are Gaussian or Cauchy.

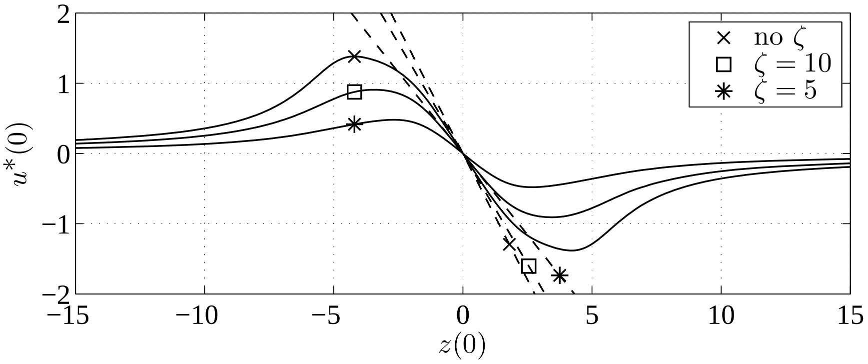
In Fig. 4.2, there is more uncertainty in the state process noise than in the measurement noise. When large measurement deviations occur (such as at $k = 52$), the Cauchy controller's effort is very small. In contrast, the LEG controller responds with a large control effort that drives the states from their regulated state of zero. However, when large process noise inputs occur, the state deviates and the Cauchy controller applies a larger control effort than the LEG, thus regulating the state more effectively. It is interesting to compare Fig. 4.2 with Fig. 4.3 in light of Fig. 4.1. They suggest that when the measurement noise density parameter dominates the process noise density parameter in constructing the Cauchy controller, the effect of measurement outliers is mitigated, while still responding to state deviations due to process noise.

On the other hand, in Fig. 4.3 there is more uncertainty in the state than in the measurement, and the Cauchy controller behaves very much like the linear LEG controller. The state trajectories and control inputs of the Cauchy and LEG controllers appear equal, but actually their differences are much smaller than the scale of the axis and cannot be seen. Similar behavior was observed in Fig. 4.1(b), suggesting that when the stochastic parameters allow it, the Cauchy controller follows the measurement more. It responds in a more linear fashion to the measurements, similar to the performance of the LEG controller in that setting.

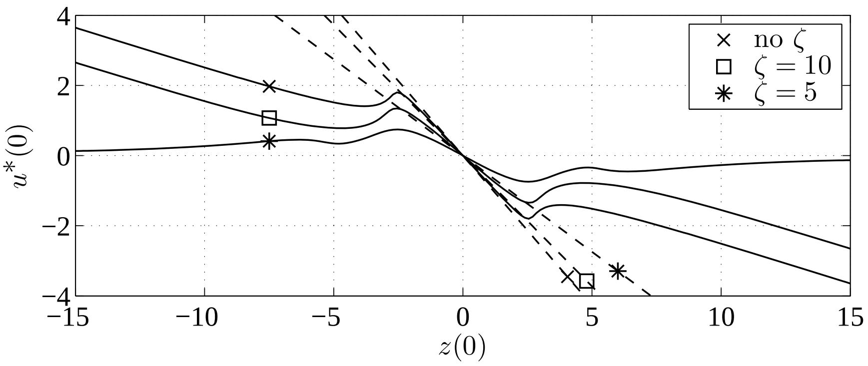
This behavior is seen again when both controllers face Gaussian noises, as in Figure 4.4. Here, the Cauchy controller closely follows both the control and state trajectories of the LEG, which is the true optimal solution. This demonstrates that the Cauchy controller is robust under non-impulsive noise environments, as it closely approximates the true optimal solution given by the LEG.

Fig. 4.5 shows two examples using control weighting $\zeta = [10 \quad 10]$, where otherwise Fig. 4.5(a) has the same parameters as Fig. 4.2, and similarly with Figs. 4.5(b) and 4.3. Applying control weighting appears to slow down the reaction from the Cauchy controller, as seen in the response to the initial process noise pulses in Fig. 4.2. In Fig. 4.5(a) with control weighting, the Cauchy controller seems to wait for a time step, coasting over the dynamics, before applying the control input. This behavior appears again in Fig. 4.5(b), where the Cauchy controller is very delayed in applying control, leading to much larger deviations of the state than the LEG allowed. This suggests that control weighting can adversely affect the regulation performance of the Cauchy controller in response to process noise pulses. However, in a situation where the measurement noise dominates, the control weightings have a much smaller effect, and improve performance in ignoring measurement outliers.

The final set of results in Fig. 4.6 show situations where either measurement or process noise is Cauchy and the other is Gaussian. In Fig. 4.6(a) the measurement noise is Cauchy and the process noise is Gaussian. Hence, there are no large state deviations, and it is clear that the LEG controller responds to all of the pulses, inducing errors in the state regulation. In Fig. 4.6(b), the measurement noise is Gaussian and the process noise is Cauchy. Here, the large deviations occur in the state, but there are no large measurement deviations. The LEG controller therefore, as a linear controller, will respond to all the measurements, and hence will perform well in regulating the state. The Cauchy controller's performance closely matches that of the LEG in this case, except for the early time steps.

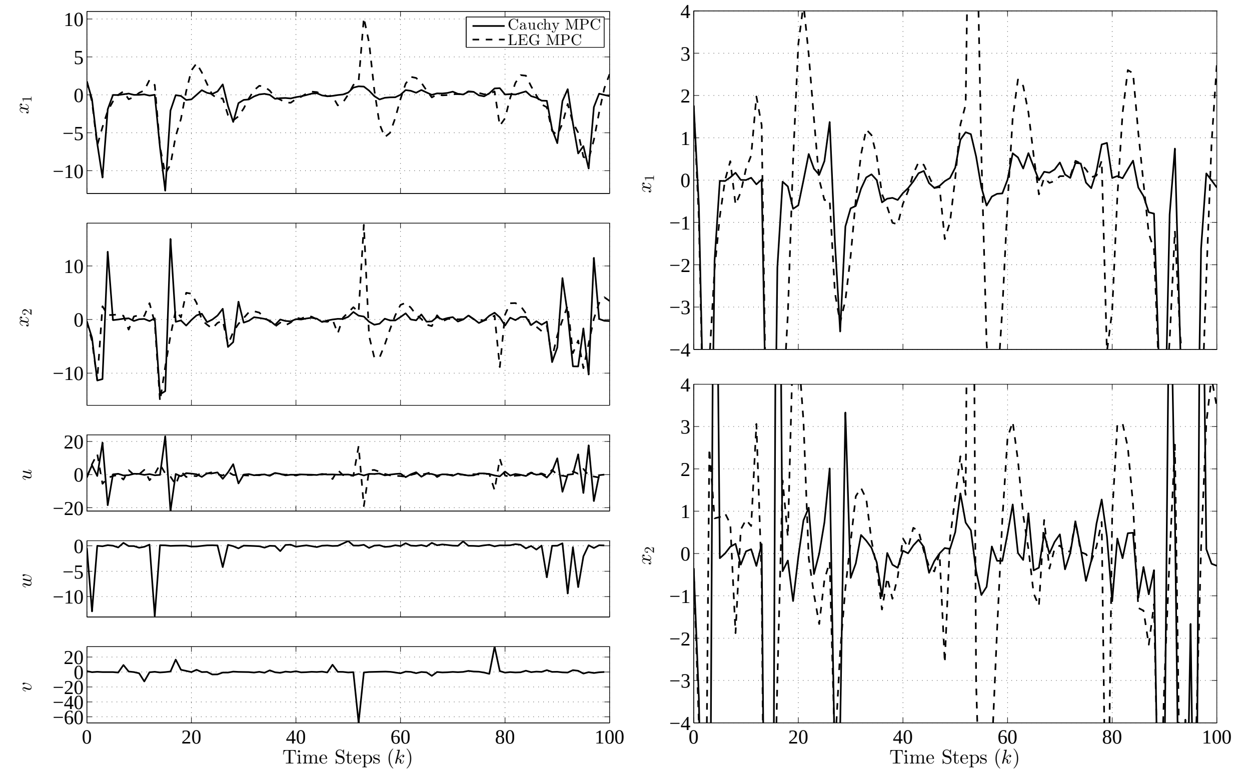


(a) $\alpha_i = 0.1$, $\gamma = 0.5$.



(b) $\alpha_i = 0.5$, $\gamma = 0.1$.

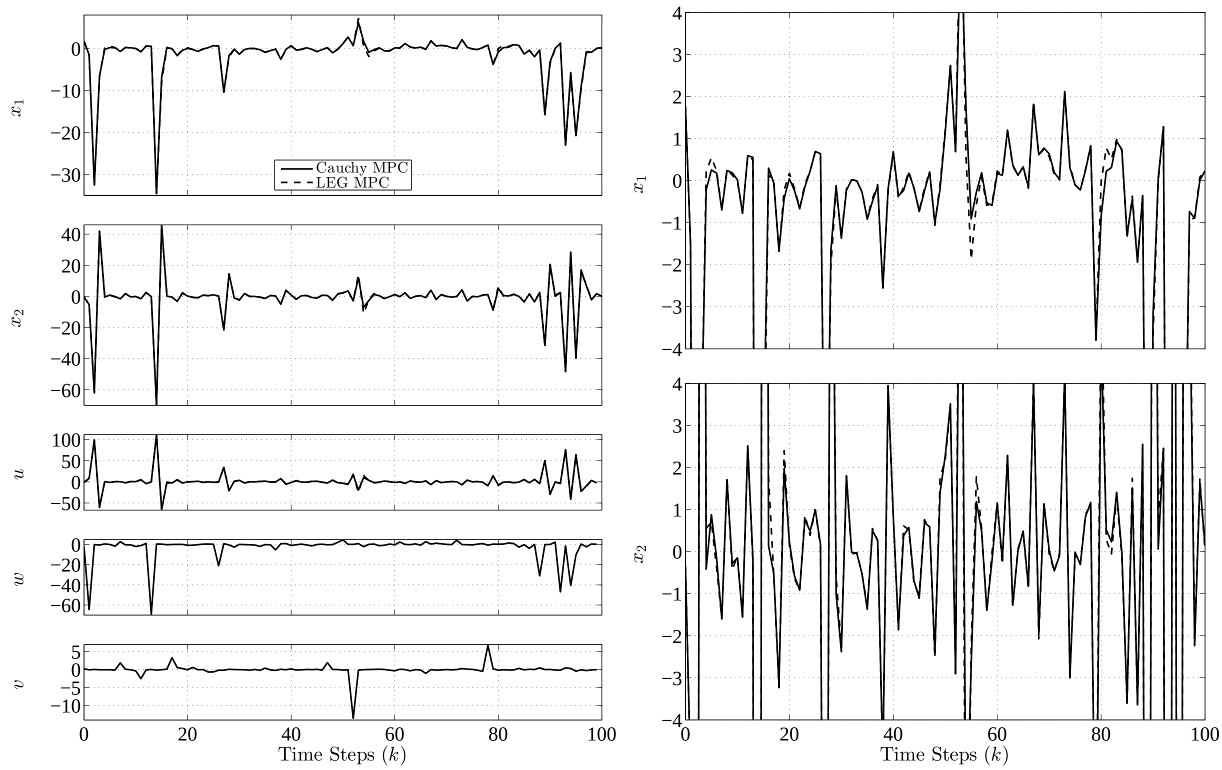
Figure 4.1: Optimal control vs the measurement for the first time step.



(a) Full view.

(b) Zoomed view.

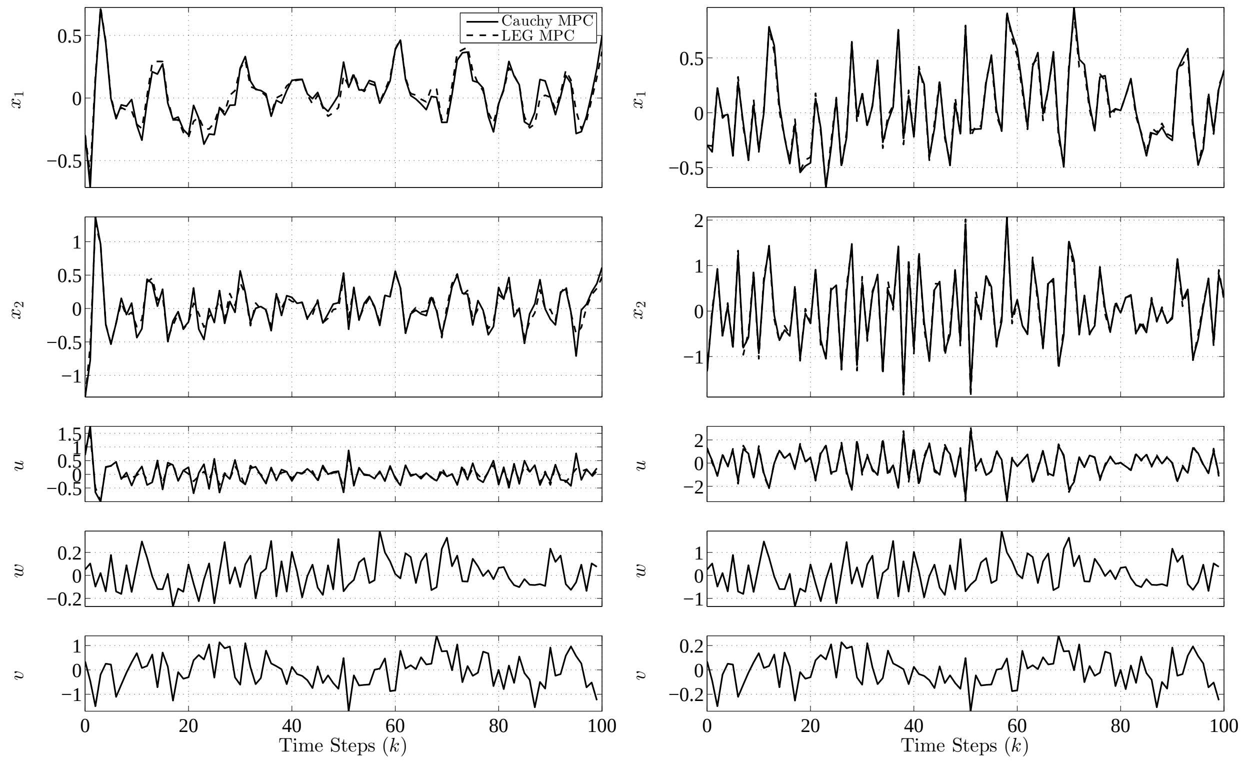
Figure 4.2: Cauchy and LEG controllers' performance when the measurement noise dominates the process noise, $\gamma = 0.5$ and $\beta = 0.1$, and without control weighting.



(a) Full view.

(b) Zoomed view.

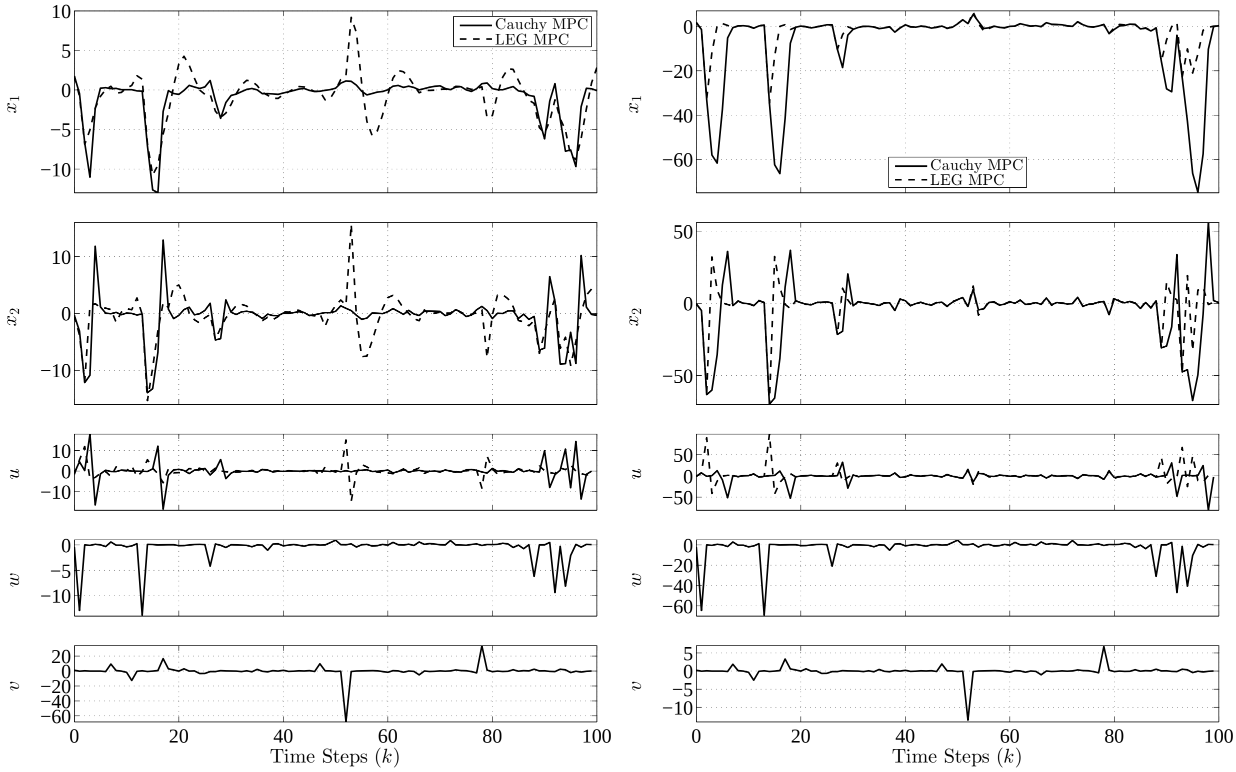
Figure 4.3: Cauchy and LEG controllers' performance when the measurement noise dominates the process noise, $\gamma = 0.1$ and $\beta = 0.5$, and without control weighting.



(a) $\gamma = 0.5$, $\beta = 0.1$.

(b) $\gamma = 0.1$, $\beta = 0.5$.

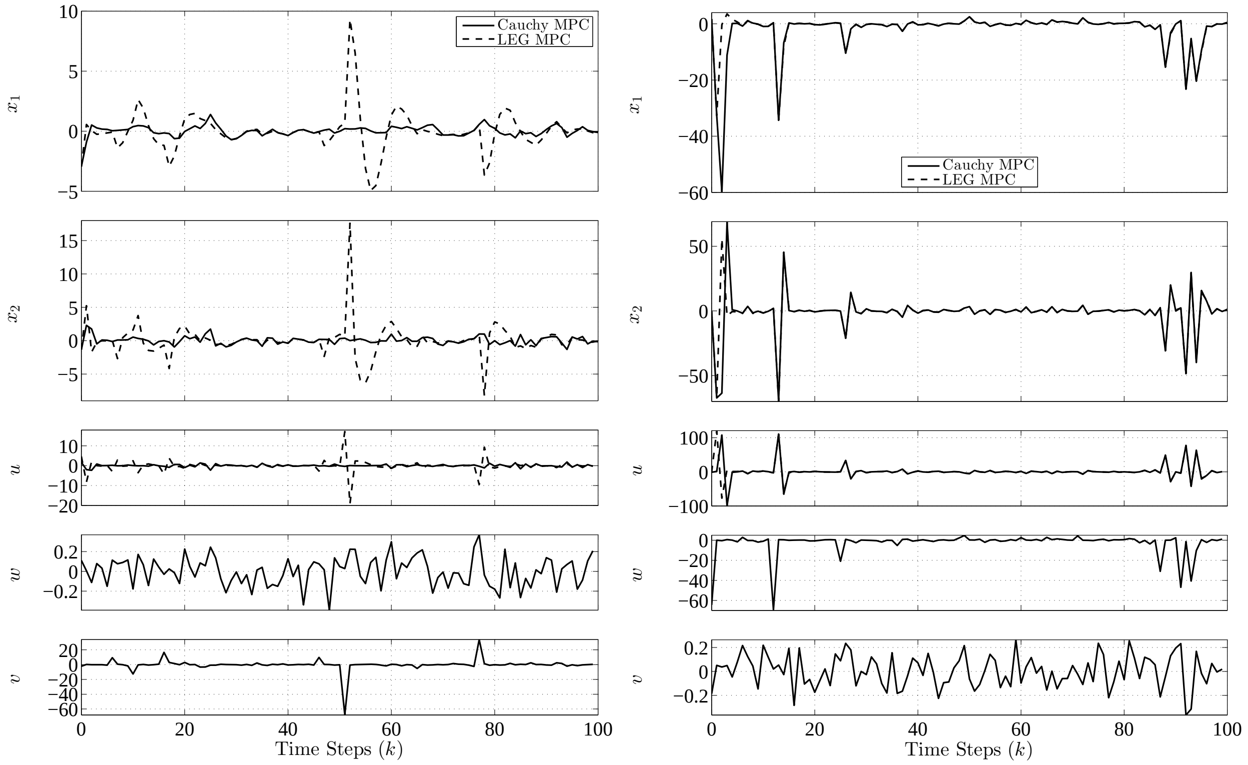
Figure 4.4: Cauchy and LEG controllers' performance against Gaussian noises closest in the $S\alpha S$ sense to the given Cauchy parameters.



(a) $\gamma = 0.5$, $\beta = 0.1$.

(b) $\gamma = 0.1$, $\beta = 0.5$.

Figure 4.5: Cauchy and LEG controllers' performance with control weighting, $\zeta = [10 \;\; 10]$.



(a) Gaussian process noise, Cauchy measurement noise, $\gamma = 0.5$ and $\beta = 0.1$.

(b) Cauchy process noise, Gaussian measurement noise, $\gamma = 0.1$ and $\beta = 0.5$.

Figure 4.6: Cauchy and LEG controllers' performance against mixed Cauchy and Gaussian noises.

CHAPTER 5

Conclusions

What was it? What had it ever been? Of what was it made? Why was there no evidence of separate blocks in the glassy, bafflingly homogenous walls? Why were there no traces of doors, either interior or exterior? I knew only that I was in a round, roofless, doorless edifice of some hard, smooth, perfectly transparent, non-refractive and non-reflective material, a hundred yards in diameter, with many corridors, and with a small circular room at the centre. More than this I could never learn from a direct investigation.

H.P. Lovecraft, "Within the Walls of Eryx"

The main contribution of this work has been a methodology for performing optimal control using characteristic functions directly, and an efficient and recursive estimator for linear systems driven by Cauchy noises. The proposed estimator is based on the general vector-state estimator, but it takes advantage of all the known relationships and analysis to make the estimator as efficient as we were able. This limits the efficient estimator to two-state systems. Finally, by determining the finite horizon approximation, this efficient estimator can be run over an indefinite number of measurements.

The simplified estimator structure allowed the derivation of a closed form expression of the conditional performance index for the model predictive control problem, using a new, computable objective function. This closed-form conditional performance index is optimized numerically using an accelerated gradient search. This improved estimator performance allowed us to study the performance of a vector state Cauchy controller over a large number of measurement updates. Hence, we present here all the methods necessary for efficient control of two-state state systems driven by Cauchy noises, and thereby provide direction for the development of an efficient general vector-state controller and estimator.

Future work should focus on improving the efficiency of the general vector-state Cauchy estimator. The main lesson form the two-state case is the importance of the fundamental directions in $\mathcal{B}$ and how they proliferate and combine as new terms are created. The pattern of this proliferation is expressed in the array M, where the *ordering of the elements in each row* and the *ordering of the rows themselves* are important. Whether for systems of order greater than two these relationships can be expressed in a more complex M array, or if multiple such arrays are needed, is currently unclear.

The development of the two-state estimator presented in this work was based on the two-state estimator for a system without process noise, i.e. a system where only the dynamics Φ affect the state, and the only Cauchy inputs are the initial condition and the measurement noise. It was by studying this problem that the M array of integers was first conceived and used, in conjunction with the set of fundamental directions $\mathcal{B}$ to construct the full two-state estimator. I believe that any efficient algorithm for a vector state system should begin with a study of the three-state case without process noise, in order to understand how the update structure evolves without the obfuscating complexity of the process noise.

APPENDIX A

The Algebraic Results

Result 1. *The product of two terms given by*

$$\{a_i^m + jd_i^m \cdot \operatorname{sgn}(\mathrm{HA}\nu)\} \times \left\{\frac{1}{j\theta + \gamma + D \cdot \operatorname{sgn}(\mathrm{HA}\nu)} - \frac{1}{j\theta - \gamma + D \cdot \operatorname{sgn}(\mathrm{HA}\nu)}\right\}$$
$$= a_o^m + jd_o^m \cdot \operatorname{sgn}(\mathrm{HA}\nu) \quad \text{(A.1a)}$$

where the i subscript denotes the input term and the o subscript denotes the output term,

$$\begin{bmatrix} a_o^m \\ d_o^m \end{bmatrix} = \frac{1}{2\pi \cdot \bar{e}} \cdot \begin{bmatrix} \bar{a} & -\bar{d} \\ \bar{d} & \bar{a} \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ d_i^m \end{bmatrix}, \quad \text{(A.1b)}$$

and

$$\bar{a} = -2\gamma \cdot (D^2 - \gamma^2 - \theta^2), \qquad \bar{d} = 4\theta\gamma D, \qquad \bar{e} = (D^2 - \gamma^2 - \theta^2)^2 + 4(\theta D)^2 \quad \text{(A.1c)}$$

Result 2. *The product of two terms given by*

$$\{a_i^m + b_i^m \cdot \operatorname{sgn}(\mathrm{HA}\nu) + jc_i^m + jd_i^m \cdot \operatorname{sgn}(\mathrm{HA}\nu)\}$$
$$\times \left\{\frac{1}{j\theta + \gamma + \rho_m + D \cdot \operatorname{sgn}(\mathrm{HA}\nu)} - \frac{1}{j\theta - \gamma + \rho_m + D \cdot \operatorname{sgn}(\mathrm{HA}\nu)}\right\}$$
$$= a_o^m + b_o^m \cdot \operatorname{sgn}(\mathrm{HA}\nu) + jc_o^m + jd_o^m \cdot \operatorname{sgn}(\mathrm{HA}\nu), \quad \text{(A.2a)}$$

where the i subscript denotes the input term and the o subscript denotes the output term,

$$\begin{bmatrix} a_o^m \\ b_o^m \\ c_o^m \\ d_o^m \end{bmatrix} = \frac{1}{\Delta} \cdot \begin{bmatrix} \bar{e} & -\bar{f} & 0 & 0 \\ -\bar{f} & \bar{e} & 0 & 0 \\ 0 & 0 & \bar{e} & -\bar{f} \\ 0 & 0 & -\bar{f} & \bar{e} \end{bmatrix} \cdot \begin{bmatrix} \bar{a} & \bar{b} & -\bar{c} & -\bar{d} \\ \bar{b} & \bar{a} & -\bar{d} & -\bar{c} \\ \bar{c} & \bar{d} & \bar{a} & \bar{b} \\ \bar{d} & \bar{c} & \bar{b} & \bar{a} \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ b_i^m \\ c_i^m \\ d_i^m \end{bmatrix}, \quad \text{(A.2b)}$$

and

$$\bar{a} = -2\gamma\left(\rho_m^2 + D^2 - \gamma^2 - \theta^2\right), \qquad \bar{b} = -4\gamma\rho_m D, \qquad \bar{c} = 4\theta\gamma\rho_m, \qquad \bar{d} = 4\theta\gamma D,$$
$$\bar{e} = \left(\rho_m^2 + D^2 - \gamma^2 - \theta^2\right)^2 + 4\left(\rho_m D\right)^2 + 4\left(\rho_m\theta\right)^2 + 4\left(\theta D\right)^2, \tag{A.2c}$$
$$\bar{f} = 4\left(\rho_m D\right)\cdot\left(\rho_m^2 + D^2 - \gamma^2\right) + 8\theta^2\rho_m D, \qquad \Delta = \pi\cdot\left(\bar{e}^2 - \bar{f}^2\right).$$

Result 3. *The term given by*

$$\left\{\frac{a_i^m + b_i^m\cdot\operatorname{sgn}(\mathrm{HA}\nu) + jc_i^m + jd_i^m\cdot\operatorname{sgn}(\mathrm{HA}\nu)}{j\theta + \rho_m - \gamma\cdot\operatorname{sgn}\left(\mathcal{B}_{k|k}^{M_i^m}\nu\right) + D\cdot\operatorname{sgn}(\mathrm{HA}\nu)}\right.$$
$$\left. - \frac{a_i^m - b_i^m\cdot\operatorname{sgn}(\mathrm{HA}\nu) - jc_i^m + jd_i^m\cdot\operatorname{sgn}(\mathrm{HA}\nu)}{j\theta - \rho_m - \gamma\cdot\operatorname{sgn}\left(\mathcal{B}_{k|k}^{M_i^m}\nu\right) + D\cdot\operatorname{sgn}(\mathrm{HA}\nu)}\right\}$$

$$= a_o^m + b_o^m\cdot\operatorname{sgn}\left(\mathcal{B}_{k|k}^{M_i^m}\nu\right)\operatorname{sgn}(\mathrm{HA}\nu) + jc_o^m\cdot\operatorname{sgn}\left(\mathcal{B}_{k|k}^{M_i^m}\nu\right) + jd_o^m\cdot\operatorname{sgn}(\mathrm{HA}\nu) \tag{A.3a}$$

where the i subscript denotes the input term and the o subscript denotes the output term,

$$\begin{bmatrix} a_o^m \\ b_o^m \\ c_o^m \\ d_o^m \end{bmatrix} = \frac{1}{\Delta}\cdot\begin{bmatrix} \bar{e} & -\bar{f} & 0 & 0 \\ -\bar{f} & \bar{e} & 0 & 0 \\ 0 & 0 & \bar{e} & -\bar{f} \\ 0 & 0 & -\bar{f} & \bar{e} \end{bmatrix}\cdot\begin{bmatrix} \bar{a} & \bar{b} & \bar{c} & \bar{d} \\ \bar{b} & \bar{a} & \bar{d} & \bar{c} \\ -\bar{c} & -\bar{d} & \bar{a} & \bar{b} \\ -\bar{d} & -\bar{c} & \bar{b} & \bar{a} \end{bmatrix}\cdot\begin{bmatrix} -\rho_m & D & -\theta & 0 \\ 0 & -\gamma & 0 & 0 \\ 0 & 0 & -\gamma & 0 \\ 0 & \theta & D & -\rho_m \end{bmatrix}\cdot\begin{bmatrix} a_i^m \\ b_i^m \\ c_i^m \\ d_i^m \end{bmatrix}, \tag{A.3b}$$

and

$$\bar{a} = \gamma^2 + D^2 - \rho_m^2 - \theta^2, \qquad \bar{b} = -2\gamma D, \qquad \bar{c} = -2\theta\gamma, \qquad \bar{d} = 2\theta D,$$
$$\bar{e} = \bar{a}^2 + \bar{b}^2 + \bar{c}^2 + \bar{d}^2, \qquad \bar{f} = 2\cdot\left(\bar{a}\bar{b} + \bar{c}\bar{d}\right), \qquad \Delta = \pi\cdot\left(\bar{e}^2 - \bar{f}^2\right). \tag{A.3c}$$

Result 4. *The term given by*

$$\left\{\frac{a_i^m + b_i^m\cdot\operatorname{sgn}(\nu_1) + jc_i^m + jd_i^m\cdot\operatorname{sgn}(\nu_1)}{j\theta + \rho_m + \check{D}\cdot\operatorname{sgn}(\nu_1)} - \frac{a_i^m - b_i^m\cdot\operatorname{sgn}(\nu_1) - jc_i^m + jd_i^m\cdot\operatorname{sgn}(\nu_1)}{j\theta - \rho_m + \check{D}\cdot\operatorname{sgn}(\nu_1)}\right\}$$

$$= a_{i,m} + jd_{i,m}\cdot\operatorname{sgn}(\nu_1) \tag{A.4a}$$

where the parameters $a_{i,m}$ and $d_{i,m}$ are generated from the other parameters by the linear *relationships*

$$\begin{bmatrix} a_{i,m} \\ d_{i,m} \end{bmatrix} = \frac{1}{\Delta} \cdot \begin{bmatrix} \bar{a} & \bar{b} & 0 & 0 \\ 0 & 0 & \bar{a} & -\bar{b} \end{bmatrix} \cdot \begin{bmatrix} \rho_m & -\check{D} & -\theta & 0 \\ \theta & 0 & \rho_m & -\check{D} \\ 0 & \theta & -\check{D} & \rho_m \\ -\check{D} & \rho_m & 0 & -\theta \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ b_i^m \\ c_i^m \\ d_i^m \end{bmatrix}, \tag{A.4b}$$

where

$$\begin{gathered} \bar{a} = \check{D}^2 - \rho_m^2 - \theta^2, \qquad \bar{b} = -2\theta\gamma, \\ \Delta = \pi \cdot \left(\bar{a}^2 + \bar{b}^2\right). \end{gathered} \tag{A.4c}$$

References

[1] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation, and Control.* SIAM, 2008.

[2] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable.* Random House, 2007.

[3] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, "Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution," *IEEE Communications Letters*, vol. 2, no. 10, pp. 282–284, Oct. 1998.

[4] P. Reeves, "A non-gaussian turbulence simulation," Air Force Flight Dynamics Laboratory, Tech. Rep. AFFDL-TR-69-67, 1969.

[5] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance.* New York: Chapman & Hall, 1994.

[6] P. Tsakalides and C. L. Nikias, *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.* Birkhauser, 1998.

[7] G. A. Tsihrintzis, *Statistical Modeling and Receiver Design for Multi-User Communication Networks;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.* Birkhauser, 1998.

[8] G. A. Hewer, R. D. Martin, and J. Zeh, "Robust preprocessing for kalman filtering of glint noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 1, pp. 120–128, January 1987.

[9] J. P. Nolan, J. G. Gonzalez, and R. C. Núñez, "Stable filters: A robust signal processing framework for heavy-tailed noise." in *In Proceedings of the 2010 IEEE Radar Conference*, 2010.

[10] M. Idan and J. L. Speyer, "Cauchy estimation for linear scalar systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1329–1342, 2010.

[11] ——, "State estimation for linear scalar dynamic systems with additive cauchy noises: Characteristic function approach," *SIAM J. Control Optim.*, vol. 50, no. 4, pp. 1971–1994, 2012.

[12] J. L. Speyer, M. Idan, and J. Fernández, "Multi-Step Prediction Optimal Control for a Scalar Linear System with Additive Cauchy Noise," in *IEEE Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.

[13] ——, "A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise," *Automatica*, 2013.

[14] M. Idan and J. L. Speyer, “Multivariate Cauchy Estimator with Scalar Measurement and Process Noises,” in *Proceedings of the* 52^{nd} *IEEE Conference on Decision and Control*, Florence, Italy, December 2013.

[15] ——, “Multivariate cauchy estimator with scalar measurement and process noises,” *Submitted to SIAM*, 2013.

[16] J. L. Speyer, M. Idan, and J. Fernández, “The Two-State Estimator for Linear System with Additive Measurement and Process Cauchy Noise,” in *IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2012.

[17] J. H. Fernández, J. L. Speyer, and M. Idan, “Linear Dynamic Systems with Additive Cauchy Noises - Part 1: Stochastic Estimation for Two-State Systems,” *Submitted to IEEE Transactions*, 2013.

[18] J. Fernández, J. L. Speyer, and M. Idan, “A Stochastic Controller for Vector Linear Systems with Additive Cauchy Noises,” in *Proceedings of the* 52^{nd} *IEEE Conference on Decision and Control*, Florence, Italy, December 2013.

[19] J. H. Fernández, J. L. Speyer, and M. Idan, “Linear Dynamic Systems with Additive Cauchy Noises - Part 2: Stochastic Model Predictive Control,” *Submitted to IEEE Transactions*, 2013.

[20] T. S. Ferguson, “A Representation of the Symmetric Bivariate Cauchy Distribution,” *Institute of Mathematical Statistics*, vol. 33, no. 4, pp. 1256–1266, Dec 1962.

[21] K.-C. Chu, “Estimation and Decision for Linear Systems with Elliptical Random Processes,” October 1973.

[22] D. H. Jacobson, “Optimal Stochastic Linear Systems with Exponential Criteria and Their Relalation to Deterministic Differential Games,” *IEEE Transactions on Automatic Control*, vol. 18, no. 2, April 1973.

[23] M. Idan and J. L. Speyer, “Characteristic Function Approach for Estimation of Scalar Systems with Cauchy Noises,” in *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario Canada, Aug. 2010.

[24] D. Champeney, *A Handbook of Fourier Theorems.* New York: Cambridge University Press, 1987.

[25] C.-T. Chen, *Linear System Theory and Design*, 3rd ed. New York: Oxford University Press, 1999.

[26] W. Rudin, *Principles of mathematical analysis*, 3rd ed. New York: McGraw-Hill Book Co., 1976, international Series in Pure and Applied Mathematics.

[27] E. W. Hansen, *Fourier Transforms and Complex Variables.* Hanover, NH: Thayer School of Engineering at Dartmouth College, 2005, unpublished Course Notes.

[28] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. John Wiley & Sons, Inc., 1999.

[29] M. Morari and E. Zafiriou, *Robust Process Control*. New Jersey: Prentice-Hall, 1989.

[30] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*. New York: Springer-Verlag, 1975.

[31] R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," *The Computer Journal*, vol. 6, no. 2, pp. 163–168, 1963.