UNIVERSITY OF CALIFORNIA

Los Angeles

State Estimation for Discrete Linear Systems with Additive Laplace Noise

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

by

Nhattrieu Chan Duong

2020

ABSTRACT OF THE DISSERTATION

State Estimation for Discrete Linear Systems
with Additive Laplace Noise

by

Nhattrieu Chan Duong
Doctor of Philosophy in Aerospace Engineering
University of California, Los Angeles, 2020
Professor Jason L. Speyer, Chair

State estimators are developed for discrete linear systems with scalar additive Laplace process and measurement noises and their properties are analyzed. For the scalar case, an analytic recursive estimator is presented, along with detailed analysis of its behavior with respect to noise parameters. In addition, a one-step model predictive controller is developed. Using an objective function with 1-norm control and terminal state costs, the expectation of the objective function with respect to the conditional probability density function is determined by using the computational structure developed for the estimator. Numerical simulations for both the estimator and one-step controller are presented to demonstrate their unique behavior, including robustness to noise spikes in the measurements. For the general vector system, update and propagation algorithms as well as a method for computing moments in closed form using characteristic functions are presented. An explicit state estimator is developed for the two-state case, and a numerical example is presented to demonstrate the algorithms and the unique properties of Laplace estimators.

The dissertation of Nhattrieu Chan Duong is approved.

Lieven Vandenberghe

Tetsuya Iwasaki

Jeffrey Eldredge

Jason L. Speyer, Committee Chair

University of California, Los Angeles

2020

To my family . . .

my dad Hai, my mom Ngoc Bich, and my sisters Jess and Uyen . . .

whose unwavering love and support

have made this possible

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Thank you to the Cauchy Team, for the long hours and hard work in realizing these results. Professor Speyer, I am so grateful for your patience, guidance and good humor. It's been a long journey, and I'm very appreciative of all you've done to help me through it. Moshe, thank you for your attention to detail, your unyielding demand for rigor, and your visits all these years. Jun, thank you for helping me derive those cpdfs. Nat, thank you for your work in finding algorithms for flattening the g's.

I'd like to thank two people whom I've never met yet provided key contributions. Dr. Rom Pinchasi, from the Technion in Haifa, Israel, for your work with Moshe in proving the key hyperplane conjecture. Dr. Miroslav Rada, from the University of Economics in Prague, Czech Republic, for providing the code and guidance for enumerating the faces of hyperplane arrangements. And thank you Gurobi, for providing fast, efficient LP solvers.

To my support group over the years, I am grateful for your encouragement and friendship. Besides my family, to whom I've dedicated this dissertation, I want to thank the following people: Sung, for your support, admonition and constant, bordering on aggravating, reminder to stay on track. Jay and Wei, for those hours spent preparing for prelims. Emmanuell and Andre, for the pointers and good times at SySense. Betancourt, Rahul, Adrian, Lauren, and Nhi, for checking in on me and making sure I was on track (with two you constantly reminding me that I had work to do). Curtis, for your careful proofreading. My roommates, Ryan, Rob, Amy and Molly, who can now stop calling me Dr. Tom. My sincere apologies to anyone else I've missed.

This work was supported by the National Science Foundation (NSF) Grant Nos. 1607502 and 1934467 and United States-Israel Binational Science Foundation (BSF) joint NSF-BSF ECCS program under Grant No. 2015702.

Lastly, I would be remiss not to acknowledge the COVID-19 pandemic, which helped encourage me to hole up in my room for three straight months finishing this dissertation!

VITA

2007	B.S. (Mechanical Engineering), California Institute of Technology, Pasadena, California.
2008	M.S. (Aerospace Engineering), UCLA, Los Angeles, California.
2012	Teaching Assistant, Mechanical and Aerospace Engineering, UCLA. During Spring 2012, led discussions for MAE 171A Introduction to Feedback and Control Systems.
2016	Teaching Assistant, Mechanical and Aerospace Engineering, UCLA. During Spring 2016, led discussions for MAE 171A Introduction to Feedback and Control Systems.
2009–present	Research Engineer, SySense Incorporated, El Segundo, California.

PUBLICATIONS

Duong, Nhattrieu, Speyer, Jason L., Yoneyama, Jun, Idan. Moshe. Laplace Estimator for Linear Scalar Systems. 2018 IEEE Conference on Decision and Control (CDC). Miami Beach, FL, USA. 17-19 December 2018

Duong, Nhattrieu, Speyer, Jason L., Idan. Moshe. Laplace Controller for Linear Scalar Systems. 2019 27th Mediterranean Conference on Control and Automation (MED). Akko, Israel, Israel. 1-4 July 2019

CHAPTER 1

Introduction

In many engineering applications, random processes or noises have volatility that are not well-modeled by Gaussian distributions. The Gaussian distribution is considered a light-tailed distribution, whose tails decay at a faster rate than an exponential [Bry74]. While its structure lends itself to compact, closed-form analytical results, this is in fact a constraint on the robustness of its modeling. The light tails poorly model systems with noise spikes, such as radar, sonar [KFR98], and stock market volatility [Lin05], and algorithms built on Gaussian distributions are susceptible to such outliers. Ad-hoc methods have been developed to compensate for this limitation. Ideas such as pre-filters have been industry standard for decades.

Past efforts in deriving analytic recursive estimators have used Cauchy distributions, whose heavy tails better capture volatile phenomena [IS14],[FSI15]. However, developing an estimator using the Cauchy probability density functions (pdf) directly becomes intractable beyond the scalar case, and the multivariate estimator was developed using characteristic functions [IS14]. In this work, we explore the use of Laplace distributions, whose densities have tails which do decay exponentially but at a slower rate than those of Gaussian densities. Furthermore, the structure of the Laplace pdfs allows for a more direct treatment of the conditional densities and have interesting properties with respect to objective functions with 1-norm costs.

1.1 Laplace densities

The Laplace probability density function (pdf) with zero mean and variance $2\alpha^2$ has the form

$$f_L(x) = \frac{1}{2\alpha} e^{-\frac{1}{\alpha}|x|} \tag{1.1}$$

Contrast this with the pdfs for the Cauchy distribution with zero median and Gaussian distribution with zero mean and variance σ^2,

$$\begin{aligned} f_C(x) &= \frac{\beta\pi}{\beta^2 + x^2} \\ f_G(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \end{aligned} \tag{1.2}$$

Figure 1.1 shows plots of the Cauchy, Laplace, and Gaussian probability density functions where the spread parameters α, β, and σ have been chosen to minimize the square of the difference between the pdfs. For a given α, $\sigma \approx 1.06\alpha$ and $\beta \approx 1.4\sigma$ [IS14]. We observe the heaviness of the Cauchy tails compared to that of the other two. While the Laplace tails still over-bound those of the Gaussian, we can see that it does decay exponentially and is much lighter than those of the Cauchy pdf.

1.2 Particle filters and Laplace distributions

With the advent of fast, inexpensive computational capabilities, simulations methods have been used to fill in the gap where analytical filters have been absent. Particle filters have had widespread use in non-linear systems [MBL11] in robotics [KSO16], navigation, and image processing, using both Gaussian and non-Gaussian noise. Laplace densities have been used in areas such as image [RVG06] and speech [LBG10] processing. However, these techniques are approximate by nature and do not produce explicit closed-form expressions for the minimum variance estimate.

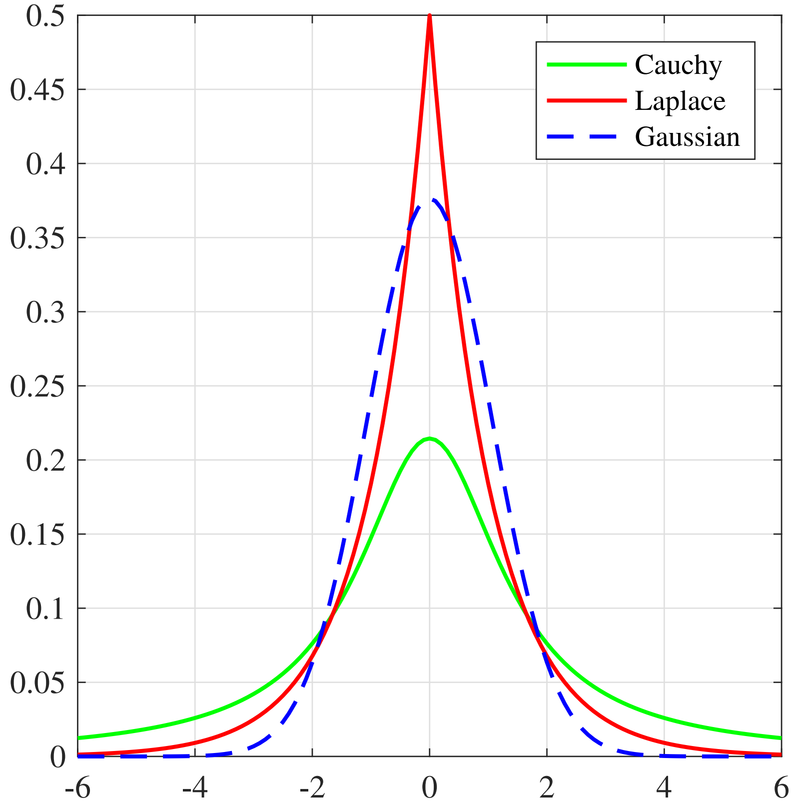


Figure 1.1: Comparison of pdfs with same effective variance

1.3 Analytic linear estimators

Prior to this work, there were two known analytic recursive linear estimators: the Kalman filter and the Cauchy estimator. The Kalman filter has been in use for decades, and the use of Gaussian noise has been successfully applied to solve a large class of problems. However, its limitations have been well-documented. The Cauchy estimator was developed in [IS10],[IS12] and [IS14]. It is considered a heavy-tailed estimator, and its behavior is quite different compared to the Kalman filter. In particular, the variance was shown to be a function of the measurements as well as the noise parameters. In contrast, the Kalman filter variance can be computed *a priori*. One can see the consequences of this when the Kalman filter processes Cauchy noise [IS12].

There has some work done in the way of estimating Laplace random vector corrupted by

Gaussian noise [Sel08] and state estimation for linear systems driven by Laplace noise using a bank of Kalman filters [FMS16]. However, in this work, we derive a recursive analytical estimator in closed-form. Like the Kalman filter and Cauchy estimator, the solution is the exact optimal estimator for a linear system with additive Laplace noise, not an approximation.

These three analytic linear estimators can be distinguished by their *a posteriori* conditional pdfs (cdpf). The Kalman filter's *a posteriori* cpdf is Gaussian and is therefore both symmetric and unimodal. In contrast, the Cauchy *a posteriori* cpdf is neither symmetric nor unimodal. True to its nature of sitting between the other two, the Laplace *a posteriori* cpdf is not symmetric but is unimodal.

1.4 Laplace controllers

Another interesting aspect of the functional form of the Laplace density is the 1-norm in the exponential. The commonly-used least-squares optimal solutions to control problems minimize the square of a cost. This is a matter of applying the tools that are available rather than the ones that are desired. For many problems where one wishes to optimize the 1-norm, such as fuel consumption, the least-squares solution undervalues the control cost near zero, leading to unnecessary control inputs when the error is small. Since the Laplace pdf has the same form as a an objective function with 1-norm costs, the equations developed to obtain the expectations for the estimator can also be used to obtain the expected function of the cost.

1.5 Outline

In Chapter 2, we state the estimation problem for a discrete scalar linear system driven by additive Laplace noise process with additive Laplace noise measurements. We then develop the conditional probability density functions (cpdf) through the first few update and propagation steps to motivate a general form for the k-th step, and we state the recursive algorithm to update and propagate from step k to step $k+1$. We then use the Fourier transform of the cpdf to derive the closed-form solutions for the normalization and first two moments in order to obtain the state estimate and estimation error variance. Finally, we present a numerical example tying everything together and discuss some of the features of a Laplace estimator. In Chapter 3, following a similar process to the scalar case, we generalize the derivation to that of a discrete vector linear system. We present a method for obtaining the moments in n-dimensions. We then explicitly derive the closed-form expressions for the normalization, first and second moments and present a numerical example for the 2-dimensional problem. In Chapter 5, we go back to the scalar case and derive a one-step model-predictive controller by minimizing a cost function with 1-norm cost. We present a numerical example and discuss some of the unique properties of such a Laplace controller. Finally, in Chapter 6, we offer concluding remarks and provide suggestions for future extension to this concept.

1.6 Notation

The Laplace densities involve summing many terms, including a handful of parameters and variables, and to keep track of every component explicitly will cause the notation to become quite cluttered. Therefore, we will make certain sacrifices to the notation to help keep things legible. When the notation deviates from the standard notation, it will be made clear in context.

Random variables at step k will be represented by capital, alphabetical letters, such as W_k and X_k, and their realization will have a corresponding lower-case version, w_k and x_k. Vector variables have the same convention, except that they will be boldfaced, as $\boldsymbol{X}_k$ and $\boldsymbol{x}_k$. However, elements of $\boldsymbol{x}_k \in \mathbb{R}^n$, $x_1, \cdots, x_n$, lose the step k in favor of the element index. When explicitly using the elements, the step k should be obvious in context. Matrix and vector parameters will use capital letters without indices, such as Φ, Γ and H.

pdf	Probability density function
cpdf	Conditional probability density function ($f_{X\|Y}$)
ucpdf	Unnormalized conditional probability density function ($\bar{f}_{X\|Y}$)

Table 1.1: List of initialisms

We distinguish between update and propagation steps by noting the indices for the state followed the measurement sequence, separated by a vertical line. For example, $k+1|k$ refers to $\boldsymbol{x}_{k+1}$ conditioned on $\boldsymbol{y}_k$.

Throughout this work, we will refer to the $\operatorname{sgn}(x)$ function, which is defined using the convention

$$\operatorname{sgn}(x) = \begin{cases} -1, & x \leq 0 \\ +1, & x > 0 \end{cases}. \tag{1.3}$$

Note that this may differ from other conventions, where $\operatorname{sgn}(0) = 0$. We will also make exceptions to this convention, where sometimes $\operatorname{sgn}(0) = 1$, but that will be noted explicitly.

CHAPTER 2

Scalar Laplace Estimator

In this chapter, we present the derivation, analysis, and demonstration of the minimum variance estimator for the scalar case. In addition to providing a framework for an analytic, recursive algorithm, we introduce a new concept for simplifying the relatively complicated coefficient function used to define the Laplace conditional probability density function (cpdf).

In Section 2.1, we define the discrete scalar linear system. In Section 2.2, we derive the unnormalized conditional pdfs (ucpdf) for the first few steps, and then in Section 2.3, we hypothesize the general form of the ucpdf at $k|k-1$ and prove it by induction. In Section 2.4, we discuss some properties of the structure of the ucpdf. In Section 2.5, we derive the closed-form equations for the mean and variance and discuss some properties of the estimator for the first update. Finally, in Section 2.6, we present and discuss a 50-step numerical simulation which shows the robustness of the scalar Laplace estimator.

2.1 Scalar dynamical system

The discrete scalar linear system with scalar state $\tilde{x}_k$, measurement $\tilde{z}_k$, independent measurement noise v_k, process noise w_k, and deterministic control input $\bar{u}_k$ is given by

$$\begin{aligned} \tilde{x}_{k+1} &= \Phi \tilde{x}_k + w_k + \bar{u}_k \\ \tilde{z}_k &= H \tilde{x}_k + v_k \end{aligned} \tag{2.1}$$

where $\tilde{x}_1, w_k$ and v_k are all Laplace distributed as

$$f_{\tilde{X}_1}(\tilde{x}_1) = \frac{1}{2\alpha} e^{-\frac{1}{\alpha}|\tilde{x}_1 - \bar{x}_1|} \tag{2.2}$$

$$f_W(w_k) = \frac{1}{2\beta} e^{-\frac{1}{\beta}|w_k|} \tag{2.3}$$

$$f_V(v_k) = \frac{1}{2\gamma} e^{-\frac{1}{\gamma}|v_k|} \tag{2.4}$$

and $\bar{x}_1$ is the mean of x_1. For convenience, we decompose the system into deterministic and stochastic parts, so that

$$\tilde{x}_k = \bar{x}_k + x_k, \tag{2.5}$$

where the deterministic part is

$$\bar{x}_{k+1} = \Phi \bar{x}_k + \bar{u}_k, \qquad \bar{z}_k = H\bar{x}_k, \tag{2.6}$$

and the stochastic part is

$$x_{k+1} = \Phi x_k + w_k, \qquad z_k = Hx_k + v_k. \tag{2.7}$$

We define the measurement history up to step k as a random sequence

$$\boldsymbol{Y}_k = \{Z_1, \cdots, Z_k\} \tag{2.8}$$

with the associated realization

$$\boldsymbol{y}_k = \{z_1, \cdots, z_k\}. \tag{2.9}$$

The purpose for using z as elements of $\boldsymbol{y}$ is because we refer to both the random variable and realization pretty often, and it's clearer to no have things like $\boldsymbol{y}_k$ and y_k in the same expression. It's the only major exception to the standard notation. For the remainder of the derivation, we consider only the stochastic part of the system.

2.2 First two updates and propagations

We work through the first few updates and propagations to deduce the general form of the ucpdf. These steps will also form the base case for the subsequent proof by induction.

2.2.1 Measurement update at $k = 1$

For the update at step $k = 1$, we make use of Bayes' Theorem, which states that

$$f_{A|B} = \frac{f_{B|A} f_A}{f_B}. \tag{2.10}$$

Therefore, the pdf of X_1 conditioned on the measurement sequence Y_1 is

$$f_{X_1|Y_1}(x_1|\boldsymbol{y}_1) = \frac{f_{Y_1|X_1}(\boldsymbol{y}_1|x_1) f_{X_1}(x_1)}{f_{Y_1}(\boldsymbol{y}_1)} \tag{2.11}$$

where $\boldsymbol{y}_1 = \{z_1\}$. It is simpler to consider the conditional pdf without the normalization factor, $f_{Y_1}(\boldsymbol{y}_1)$. Therefore, let us define the unnormalized conditional pdf (ucpdf) as

$$\begin{aligned} \bar{f}_{X_1|Y_1}(x_1|\boldsymbol{y}_1) &= f_{Y_1|X_1}(\boldsymbol{y}_1|x_1) f_{X_1}(x_1) \\ &= f_{Z_1|X_1}(z_1|x_1) f_{X_1}(x_1) \\ &= f_{X_1,Y_1}(x_1, \boldsymbol{y}_1) \end{aligned} \tag{2.12}$$

Using the measurement equation in (2.1) and f_V in (2.4),

$$\begin{aligned} f_{Z_1|X_1}(z_1|x_1) &= f_V(z_1 - Hx_1) \\ &= \frac{1}{2\gamma} \exp\left(-\frac{|H|}{\gamma} \left| \frac{z_1}{H} - x_1 \right| \right). \end{aligned} \tag{2.13}$$

Combining (2.13) and (2.2), we re-write the ucpdf as

$$\begin{aligned} \bar{f}_{X_1|Y_1}(x_1|\boldsymbol{y}_1) &= f_V(z_1 - Hx_1) f_{X_1}(x_1) \\ &= \frac{1}{4\alpha\gamma} \exp\left(-\frac{|H|}{\gamma} \left| \frac{z_1}{H} - x_1 \right| - \frac{1}{\alpha} |x_1| \right). \end{aligned} \tag{2.14}$$

Figure 2.1 shows $\bar{f}_{X_1|Y_1}$ for $\alpha = 1.0$ and $\gamma = 0.25$. Notice the peak at $x_1 = 1.23$ due to the measurement $z_1 = 1.23$, the slight kink at $x_1 = 0$ from the zero initial conditions, and the asymmetry of the curve.

2.2.2 Propagation from $k = 1$ to $k = 2$

For the propagation, we need to determine the conditional density $\bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1)$. Recall that

$$f_{A,B} = f_{A|B} f_B \implies f_{A,B|C} = f_{A|B,C} f_{B|C} = f_{B|C} f_{A|B,C} \tag{2.15}$$

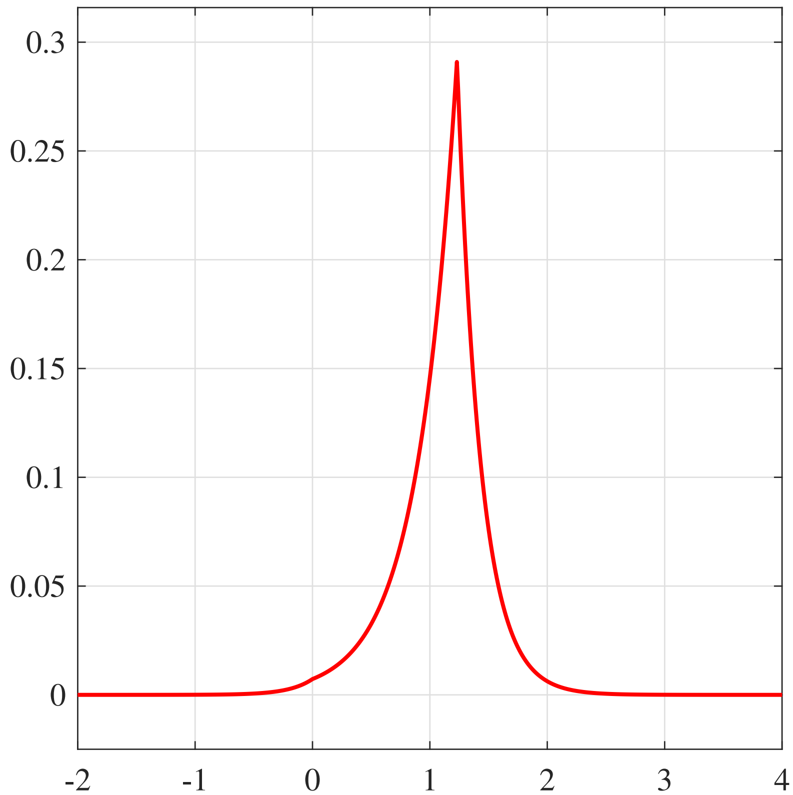


Figure 2.1: $\bar{f}_{X_1|Y_1}$: *a posteriori* unnormalized conditional pdf at $k = 1$ for $z_1 = 1.23$

We can therefore construct the joint density

$$\begin{aligned}\bar{f}_{X_1,X_2|Y_1}(x_2, x_1|\boldsymbol{y}_1) &= \bar{f}_{X_1|Y_1}(x_1|\boldsymbol{y}_1) f_{X_2|X_1,Y_1}(x_2|x_1, \boldsymbol{y}_1) \\ &= \bar{f}_{X_1|Y_1}(x_1|\boldsymbol{y}_1) f_{X_2|X_1}(x_2|x_1).\end{aligned} \tag{2.16}$$

Using the dynamical equation in (2.1) and f_W in (2.3),

$$\begin{aligned}f_{X_2|X_1}(x_2|x_1) &= f_W(x_2 - \Phi x_1) \\ &= \frac{1}{2\beta} \exp\left(\frac{|\Phi|}{\beta}\left|\frac{x_2}{\Phi} - x_1\right|\right),\end{aligned} \tag{2.17}$$

and we can re-write (2.16) as

$$\begin{aligned}\bar{f}_{X_1,X_2|Y_1}(x_2, x_1|\boldsymbol{y}_1) &= \bar{f}_{X_1|Y_1}(x_1|\boldsymbol{y}_1) f_W(x_2 - \Phi x_1) \\ &= \frac{1}{8\alpha\beta\gamma} \exp\left(-\frac{1}{\alpha}|x_1| - \frac{|H|}{\gamma}\left|\frac{z_1}{H} - x_1\right| - \frac{|\Phi|}{\beta}\left|\frac{x_2}{\Phi} - x_1\right|\right)\end{aligned} \tag{2.18}$$

To obtain $\bar{f}_{X_2|Y_1}$, we then integrate (2.16) over x_1

$$\bar{f}_{X_2|Y_1}(x_2, \boldsymbol{y}_1) = \frac{1}{8\alpha\beta\gamma} \int_{-\infty}^{\infty} \exp\left(-\frac{|H|}{\gamma} \left| \frac{z_1}{H} - x_1 \right| - \frac{1}{\alpha} |-x_1| - \frac{|\Phi|}{\beta} \left| \frac{x_2}{\Phi} - x_1 \right| \right) dx_1 \tag{2.19}$$

The solution to the integral in (2.19) was shown in Appendix B of [IS14] and is re-stated in (7.1) of the Appendix for convenience. Using that result, (2.19) is evaluated as

$$\begin{aligned} \bar{f}_{X_2|Y_1}(x_2, \boldsymbol{y}_1) = {} & g_1^{2|1} \cdot \exp\left(-\frac{1}{\alpha} \left| \frac{z_1}{H} \right| - \frac{1}{\beta} \left| \frac{\Phi z_1}{H} - x_2 \right| \right) \\ & + g_2^{2|1} \cdot \exp\left(-\frac{1}{\gamma} |z_1| - \frac{1}{\beta} |-x_2| \right) \\ & + g_3^{2|1} \cdot \exp\left(-\frac{|H|}{\gamma |\Phi|} \left| \frac{\Phi z_1}{H} - x_2 \right| - \frac{1}{\alpha |\Phi|} |-x_2| \right) \end{aligned} \tag{2.20}$$

where,

$$g_i^{2|1} = \frac{1}{8\alpha\beta\gamma} \left[\frac{1}{\rho_i + \sum\limits_{\substack{l=1 \\ l \neq i}}^{3} \rho_l \operatorname{sgn}(\xi_l - \xi_i)} - \frac{1}{-\rho_i + \sum\limits_{\substack{l=1 \\ l \neq i}}^{3} \rho_l \operatorname{sgn}(\xi_l - \xi_i)} \right] \tag{2.21}$$

and

$$\begin{aligned} \rho_1 &= \tfrac{|H|}{\gamma} & \xi_1 &= \tfrac{z_1}{H} \\ \rho_2 &= \tfrac{1}{\alpha} & \xi_2 &= 0 \\ \rho_3 &= \tfrac{|\Phi|}{\beta} & \xi_3 &= \tfrac{x_2}{\Phi} \end{aligned} \tag{2.22}$$

Since the exponentials without x_2 are constant, we can collapse them into the coefficients to get

$$\begin{aligned} \bar{f}_{X_2|Y_1}(x_2, \boldsymbol{y}_1) = {} & \bar{g}_1^{2|1} \cdot \exp\left(-\frac{1}{\beta} \left| \frac{\Phi z_1}{H} - x_2 \right| \right) \\ & + \bar{g}_2^{2|1} \cdot \exp\left(-\frac{1}{\beta} |x_2| \right) \\ & + \bar{g}_3^{2|1} \cdot \exp\left(-\frac{|H|}{\gamma |\Phi|} \left| \frac{\Phi z_1}{H} - x_2 \right| - \frac{1}{\alpha |\Phi|} |-x_2| \right) \end{aligned} \tag{2.23}$$

The coefficient terms $\bar{g}_i^{2|1}$ are constant except for one or two step changes. For example, $\bar{g}_1^{2|1}$ has a step change at $\frac{\Phi z_1}{H}$, while $\bar{g}_2^{2|1}$ has a step change at 0. $\bar{g}_3^{2|1}$ has step changes at both $\frac{\Phi z_1}{H}$ and 0, which can be observed numerically by plotting the equations with respect to

x_2. Following these observations, we state the following theorem, whose proof is shown in Section 7.2 of the Appendix.

Theorem 2.2.1. *For*

$$\mathcal{A} = \{A_0, A_1, \cdots, A_n\} \triangleq \{(-\infty, \xi_1), [\xi_1, \xi_2), \cdots, [\xi_n, +\infty)\} \tag{2.24}$$

with $\xi_1 < \cdots < \xi_n \in \mathbb{R}$*, any* $g : \mathbb{R} \to \mathbb{R}$ *constant on* $A_i \in \mathcal{A}$*, can be simplified to*

$$g(x) = \rho_0 + \sum_{i=1}^{n} \rho_i sgn\left(\xi_i - x\right) \tag{2.25}$$

where, for $x_i \in A_i$*,*

$$\begin{aligned} \rho_0 &= \frac{g(x_0) + g(x_n)}{2} \\ \rho_i &= \frac{g(x_{i-1}) - g(x_i)}{2} \end{aligned} \tag{2.26}$$

Therefore, the complicated expressions for $\bar{g}_i^{2|1}$ can be simplified to

$$\mathcal{G}_i^{2|1} = \rho_{i0}^{2|1} + \sum_{l=1}^{N_i^{2|1}} \rho_{il}^{2|1} \text{sgn}\left(\xi_{il}^{2|1} - x_2\right) \tag{2.27}$$

where $N_i^{2|1}$ is the ith element of

$$N^{2|1} = \begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^T. \tag{2.28}$$

This simplification is key to developing the recursive structure of the Laplace estimator. The number of sign functions in each term corresponds to the number of x_2 in the associated exponential part. We refer to these as "elements" (not to be confused with elements in a vector). Therefore, $N_i^{2|1}$ refers to the number of sign functions of term i.

Figure 2.2 shows $\bar{f}_{X_2|Y_1}$, propagated from $\bar{f}_{X_1|Y_1}$ in Figure 2.1, with $\beta = 0.33$. Observe how the convolution with f_W smooths out the sharp features from $\bar{f}_{X_1|Y_1}$.

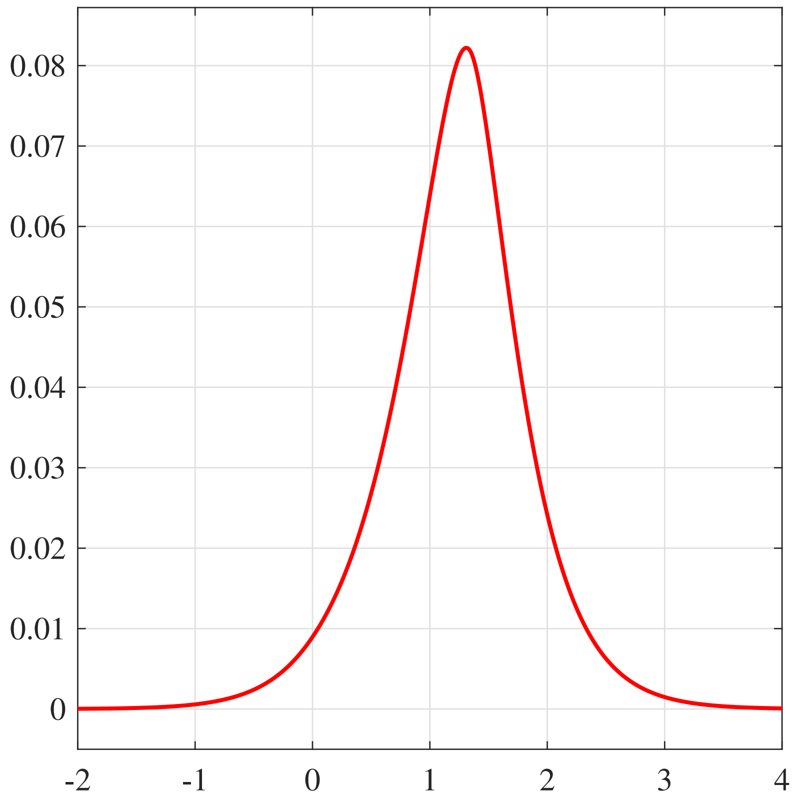


Figure 2.2: $\bar{f}_{X_2|Y_1}$: *a priori* unnormalized conditional pdf at $k = 2$

2.2.3 Update at $k = 2$

The update of the unnormalized conditional pdf $\bar{f}_{X_2|Y_2}(x_2|\boldsymbol{y}_2)$ is given by

$$\begin{aligned}\bar{f}_{X_2|Y_2}(x_2|\boldsymbol{y}_2) &= \bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1) f_{Z_2|X_2,Y_1}(z_2|x_2, \boldsymbol{y}_1) \\ &= \bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1) f_{Z_2|X_2}(z_2|x_2) \\ &= \bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1) f_V\left(z_2 - Hx_2\right)\end{aligned} \tag{2.29}$$

where, similar to (2.13),

$$f_V = \frac{1}{2\gamma} \exp\left(-\frac{|H|}{\gamma}\left|\frac{z_2}{H} - x_2\right|\right) \tag{2.30}$$

and $\bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1)$ is given in (2.20).

Figure 2.3 shows $\bar{f}_{X_2|Y_2}$ updated from $\bar{f}_{X_2|Y_1}$ in Figure 2.2 with measurement $z_2 = 1.09$.

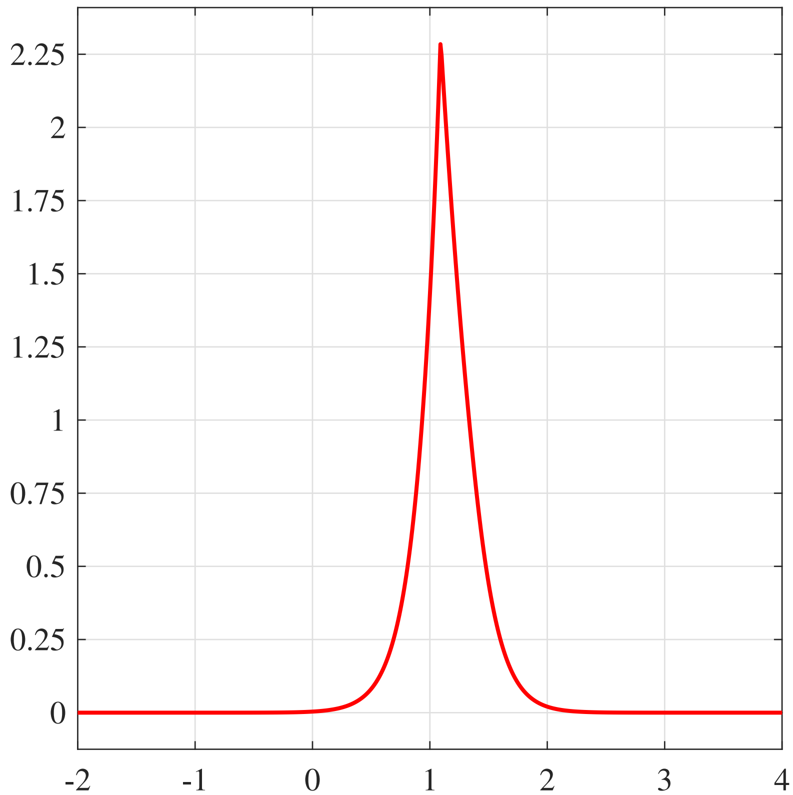


Figure 2.3: $\bar{f}_{X_2|Y_2}$: *a posteriori* unnormalized conditional pdf at $k = 2$ for $z_2 = 1.09$

2.2.4 Propagation to $k = 3$

To propagate from $k = 2$ to $k = 3$, we first obtain the joint conditional density

$$\bar{f}_{X_3,X_2|Y_2}(x_3, x_2|\boldsymbol{y}_2) = \bar{f}_{X_2|Y_2}(x_2|\boldsymbol{y}_2) f_{X_3|X_2}(x_3|x_2) \tag{2.31}$$

where, as in (2.17),

$$\begin{aligned} f_{X_3|X_2}(x_3|x_2) &= f_W(x_3 - \Phi x_2) \\ &= \frac{1}{2\beta} \exp\left(-\frac{|\Phi|}{\beta} \left| \frac{x_3}{\Phi} - x_2 \right| \right) \end{aligned} \tag{2.32}$$

and $\bar{f}_{X_2|Y_2}$ is given in (2.29). We can then compute the marginal conditional density

$$\begin{aligned}
\bar{f}_{X_3|Y_2}(x_3|\boldsymbol{y}_2) &= \frac{1}{4\beta\gamma}\int_{-\infty}^{\infty} \bar{f}_{X_2|Y_2}(x_2|\boldsymbol{y}_2) f_{X_3|X_2}(x_3|x_2)\, dx_2 \\
&= \frac{1}{4\beta\gamma}\int_{-\infty}^{\infty} \bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1) f_V(z_2 - Hx_2) f_W(x_3 - \Phi x_2)\, dx_2 \qquad (2.33) \\
&= \frac{1}{4\beta\gamma}\int_{-\infty}^{\infty} \left(\mathcal{G}_1^{2|1}\cdot\epsilon_1^{2|1} + \mathcal{G}_2^{2|1}\cdot\epsilon_2^{2|1} + \mathcal{G}_3^{2|1}\cdot\epsilon_3^{2|1}\right) dx_2
\end{aligned}$$

where

$$\begin{aligned}
\epsilon_1^{2|1} &= \exp\left(-\frac{|H|}{\gamma}\left|\frac{z_2}{H} - x_2\right| - \frac{1}{\beta}\left|\frac{\Phi z_1}{H} - x_2\right| - \frac{|\Phi|}{\beta}\left|\frac{x_3}{\Phi} - x_2\right|\right) \\
\epsilon_2^{2|1} &= \exp\left(-\frac{|H|}{\gamma}\left|\frac{z_2}{H} - x_2\right| - \frac{1}{\beta}\left|-x_2\right| - \frac{|\Phi|}{\beta}\left|\frac{x_3}{\Phi} - x_2\right|\right) \qquad (2.34) \\
\epsilon_3^{2|1} &= \exp\left(-\frac{|H|}{\gamma}\left|\frac{z_2}{H} - x_2\right| - \frac{|H|}{\gamma\,|\Phi|}\left|\frac{\Phi z_1}{H} - x_2\right| - \frac{1}{\alpha\,|\Phi|}\left|-x_2\right| - \frac{|\Phi|}{\beta}\left|\frac{x_3}{\Phi} - x_2\right|\right)
\end{aligned}$$

Using the integral formula (7.3) of the Appendix for each term in the integral, we get

$$\bar{f}_{X_3|Y_2}(x_3|\boldsymbol{y}_2) = \sum_{i=1}^{3}\sum_{j=0}^{N_i^{2|1}+1} G_{ij}^{3|2}\exp\left(\sum_{l=0}^{N_i^{2|1}+1}\eta_{jl}\left|\xi_{jl} - x_3\right|\right) \qquad (2.35)$$

where ρ_{ij}, ξ_{ij} and ρ_{ij} are

$$
\begin{array}{lll}
\rho_{10} = 0 & \xi_{10} = \frac{z_2}{H} & \eta_{10} = \frac{|H|}{\gamma} \\
\rho_{11} = a_{11}^{2|1} & \xi_{11} = \frac{\Phi z_1}{H} & \eta_{11} = \frac{1}{\beta} \\
\rho_{12} = 0 & \xi_{12} = \frac{x_3}{\Phi} & \eta_{12} = \frac{|\Phi|}{\beta} \\
\hline
\rho_{20} = 0 & \xi_{20} = \frac{z_2}{H} & \eta_{20} = \frac{|H|}{\gamma} \\
\rho_{21} = a_{21}^{2|1} & \xi_{21} = 0 & \eta_{21} = \frac{1}{\beta} \\
\rho_{22} = 0 & \xi_{22} = \frac{x_3}{\Phi} & \eta_{22} = \frac{|\Phi|}{\beta} \\
\hline
\rho_{30} = 0 & \xi_{30} = \frac{z_2}{H} & \eta_{30} = \frac{|H|}{\gamma} \\
\rho_{31} = a_{31}^{2|1} & \xi_{31} = \frac{\Phi z_1}{H} & \eta_{31} = \frac{|H|}{\gamma|\Phi|} \\
\rho_{32} = a_{32}^{2|1} & \xi_{32} = 0 & \eta_{32} = \frac{1}{\alpha|\Phi|} \\
\rho_{33} = 0 & \xi_{33} = \frac{x_3}{\Phi} & \eta_{33} = \frac{|\Phi|}{\beta}
\end{array}
\tag{2.36}
$$

and, for

$$
\begin{aligned}
\bar{\delta}_{ij} &= \sum_{\substack{l=0 \\ l\neq j}}^{N_i^{2|1}+1} \rho_{il} \mathrm{sgn}\left(\xi_{il} - \xi_{ij}\right) \\
\delta_{ij} &= \sum_{\substack{l=0 \\ l\neq j}}^{N_i^{2|1}+1} \eta_{il} \mathrm{sgn}\left(\xi_{il} - \xi_{ij}\right)
\end{aligned}
\tag{2.37}
$$

$$
G_{ij}^{3|2} = \frac{\left(\rho_{i0}^{2|1} + \bar{\delta}_{ij}\right)}{\eta_{ij} + \delta_{ij}} - \frac{\left(-\rho_{i0}^{2|1} + \bar{\delta}_{ij}\right)}{-\eta_{ij} + \delta_{ij}} \tag{2.38}
$$

Since the exponentials without x_3 are constant, we can collapse them into the coefficients and then combine terms with the same exponents to get

$$\begin{aligned}\bar{f}_{X_3|Y_2}(x_3|\boldsymbol{y}_2) = \bar{G}_1^{3|2} &\cdot \exp\left(-\frac{1}{\beta}\left|\frac{\Phi z_2}{H} - x_3\right|\right) \\ + \bar{G}_2^{3|2} &\cdot \exp\left(-\frac{1}{\beta}\left|\frac{\Phi^2 z_1}{H} - x_3\right|\right) \\ + \bar{G}_3^{3|2} &\cdot \exp\left(-\frac{1}{\beta}\left|-x_3\right|\right) \\ + \bar{G}_4^{3|2} &\cdot \exp\left(-\frac{|H|}{\gamma|\Phi|}\left|\frac{\Phi z_2}{H} - x_3\right| - \frac{1}{\beta|\Phi|}\left|-x_3\right|\right) \\ + \bar{G}_5^{3|2} &\cdot \exp\left(-\frac{1}{\beta}\left|\frac{\Phi z_2}{H} - x_3\right| - \frac{1}{\beta|\Phi|}\left|\frac{\Phi^2 z_1}{H} - x_3\right|\right) \\ + \bar{G}_6^{3|2} &\cdot \exp\left(-\frac{|H|}{\gamma|\Phi|}\left|\frac{\Phi z_2}{H} - x_3\right| - \frac{|H|}{\gamma|\Phi|^2}\left|\frac{\Phi^2 z_1}{H} - x_3\right| - \frac{1}{\alpha|\Phi|^2}\left|-x_3\right|\right)\end{aligned} \tag{2.39}$$

Note that (2.39) combines the indices i and j from (2.35) to a single index. As in step $k = 2$, $\bar{G}_i^{3|2}$ can be simplified, using Section 7.2.1 of the Appendix, to the form

$$\mathcal{G}_i^{3|2} = \rho_{i0}^{3|2} + \sum_{l=1}^{N_i^{3|2}} \rho_{il}^{3|2} \text{sgn}\left(\xi_{il}^{3|2} - x_3\right) \tag{2.40}$$

where $N_i^{3|2}$ is the ith element of

$$N^{3|2} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 3 \end{bmatrix}^T. \tag{2.41}$$

Note that terms with one element, known as “single-element” terms, combine with other single-element terms with the same exponentials. For example,

$$\begin{aligned}&\bar{G}_1^{3|2} \exp\left(-\beta\left|-x_3\right|\right) \\ &\quad = G_1^{3|2} \exp\left(-\beta\left|-x_3\right|\right) + G_5^{3|2} \exp\left(-\beta\left|-x_3\right|\right)\end{aligned} \tag{2.42}$$

After combining all the terms with identical exponential parts, we order the terms so that the ones with the fewest elements are listed first. As implied in (2.41), a general pattern emerges such that for step k, there are k terms with one element, followed by $k-1$ terms with two elements, and $k-2$ terms with three, or k, elements.

Figure 2.4 shows $\bar{f}_{X_3|Y_2}$, propagated from the $\bar{f}_{X_2|Y_2}$ in Figure 2.3.

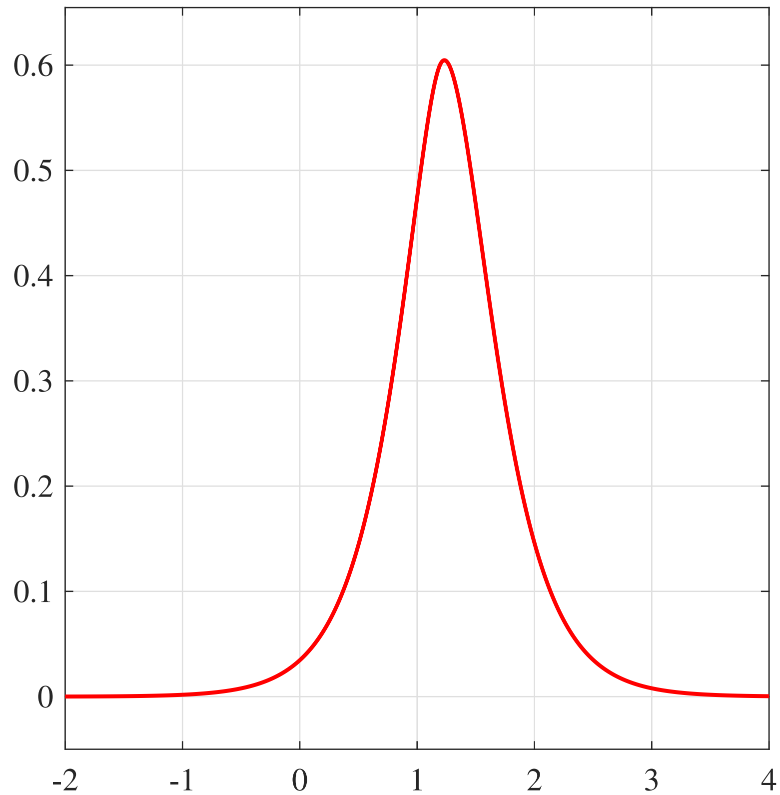


Figure 2.4: $\bar{f}_{X_3|Y_2}$: *a priori* unnormalized conditional pdf at $k = 3$

2.3 Proof of general recursive form in $\mathbb{R}$ by induction

We will use induction to show that (2.43) is the general form of the ucpdf which is preserved under measurement update and propagation.

2.3.1 Induction hypothesis

From the first few steps presented above, we hypothesize that the general *a priori* unnormalized conditional pdf at step $k|k-1$ is

$$\bar{f}_{X_k|Y_{k-1}}(x_k|\boldsymbol{y}_{k-1}) = \sum_{i=1}^{n^{k|k-1}} \mathcal{G}_i^{k|k-1}\mathcal{E}_i^{k|k-1} \tag{2.43}$$

where

$$\begin{aligned}\mathcal{G}_i^{k|k-1} &= \rho_{i0}^{k|k-1} + \sum_{l=1}^{N_i^{k|k-1}} \rho_{il}^{k|k-1} \operatorname{sgn}\left(\xi_{il}^{k|k-1} - x_k\right) \\ \mathcal{E}_i^{k|k-1} &= \exp\left(-\sum_{l=1}^{N_i^{k|k-1}} \eta_{il}^{k|k-1} \left|\xi_{il}^{k|k-1} - x_k\right|\right)\end{aligned} \tag{2.44}$$

2.3.2 Base case

By construction, we know that the ucpdf at 2|1 and 3|2 from section 2.2 have the form shown in (2.43). Next, we assume $\bar{f}_{X_k|Y_{k-1}}$ (2.43) and show that $\bar{f}_{X_{k+1}|Y_k}$ has the same form.

2.3.3 Update from $k|k-1$ to $k|k$

First update to step k by

$$\begin{aligned}\bar{f}_{X_k|Y_k} &= \bar{f}_{X_k|Y_{k-1}} f_{Z_k|X_k} \\ &= \bar{f}_{X_k|Y_{k-1}} f_V\left(\frac{z_k}{H} - x_k\right)\end{aligned} \tag{2.45}$$

This step is straight-forward, and it is clear that the form in (2.43) is preserved.

2.3.4 Propagate from $k|k$ to $k+1|k$

Then we propagate to step $k+1$ by forming the joint density $\bar{f}_{X_{k+1},X_k|Y_k}$ and evaluating the convolution integral

$$\begin{aligned}\bar{f}_{X_{k+1}|Y_k}(x_{k+1}|\boldsymbol{y}_k) &= \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k} f_{X_{k+1}|X_k}\, dx_k \\ &= \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k} f_W\left(\frac{x_{k+1}}{\Phi} - x_k\right) dx_k\end{aligned} \tag{2.46}$$

Each term i in $\bar{f}_{X_{k+1}|Y_k}$ must be evaluated using the integral formula (7.3) of the Appendix by choosing the appropriate parameters ρ, η and ξ, as we had done for step $k = 3$ (2.36). We set each offset of a particular term in the exponent of $\mathcal{E}_i^{k|k-1}$ as ξ and its corresponding coefficient as η. If there is a corresponding sign function in $\mathcal{G}_i^{k|k-1}$, then its coefficient is ρ. When we include f_V and f_W, there is no corresponding sign function, so we set $\rho = 0$. These were the first and last rows for each set of integration parameters for step $k = 3$ in (2.36).

In this way, we generate the *a priori* ucpdf at $k+1$ using the *a priori* ucpdf at k in (2.43). Based on the patterns for 2|1 and 3|2, we determine that the parameters at $k+1|k$ are

$$\begin{array}{lll}
\rho_{i0} = 0 & \xi_{i0} = \frac{z_k}{H} & \eta_{i0} = \frac{|H|}{\gamma} \\
\rho_{i1} = a_{i1}^{k|k-1} & \xi_{i1} = \xi_{i1}^{k|k-1} & \eta_{i1} = \eta_{i1}^{k|k-1} \\
\rho_{i2} = a_{i2}^{k|k-1} & \xi_{i2} = \xi_{i2}^{k|k-1} & \eta_{i2} = \eta_{i2}^{k|k-1} \\
\rho_{i3} = a_{i3}^{k|k-1} & \xi_{i3} = \xi_{i3}^{k|k-1} & \eta_{i3} = \eta_{i3}^{k|k-1} \\
 & \vdots & \\
\rho_{i,\tilde{N}_i} = a_{i,\tilde{N}_i}^{k|k-1} & \xi_{i,\tilde{N}_i} = \xi_{i,\tilde{N}_i}^{k|k-1} & \eta_{i,\tilde{N}_i} = \eta_{i,\tilde{N}_i}^{k|k-1} \\
\rho_{i,\tilde{N}_i+1} = 0 & \xi_{i,\tilde{N}_i+1} = \frac{x_{k+1}}{\Phi} & \eta_{i,\tilde{N}_i+1} = \frac{|\Phi|}{\beta}
\end{array} \tag{2.47}$$

where $\tilde{N}_i = N_i^{k|k-1}$ for short. The first and last rows of (2.47) correspond to the measurement update and propagation, respectively, and are the same for any term. The newest measurements being added to the top is done to preserve a consistent structure. The middle rows are determined by the parent term at step k, so ρ is simply the coefficients of the parent's sign functions, ξ are from the arguments, and η are from the coefficients in the exponential. The solution to the integral is determined using the integral formula (7.3) of the Appendix.

2.3.5 Isolate x_{k+1} and factor out constant terms

After performing the integral, the arguments of the sign functions are of the form $\xi_l - \xi_i$, only some of which involve $\xi_{i,\tilde{N}_i+1}$ which contains x_{k+1}. For example,

$$\begin{aligned}
\eta_j \mathrm{sgn}\left(\frac{x_{k+1}}{\Phi} - \frac{z_k}{H}\right) &= -\eta_j \mathrm{sgn}\left(\Phi\right) \mathrm{sgn}\left(\frac{\Phi z_k}{H} - x_{k+1}\right) \\
\eta_j \left|\frac{x_{k+1}}{\Phi} - \frac{z_k}{H}\right| &= \frac{\eta_j}{|\Phi|} \mathrm{sgn}\left(\frac{\Phi z_k}{H} - x_{k+1}\right)
\end{aligned} \tag{2.48}$$

Every term not involving x_{k+1} is constant, so the exponentials are factored out as a scalar multiple. In the coefficient functions, the constant sign terms are moved into the ρ_{i0} and η_{i0} in the numerator and denominator of (7.4), respectively.

2.3.6 Simplify coefficient function and combine terms

The complicated coefficient (G) functions are simplified using the method described in Section 7.2.1 of the Appendix. This allows for the terms with identical exponential parts to be combined. Finally, the expression is re-written in the general form given in (2.43), with the indices $k-1$ and k incremented to k and $k+1$, respectively, as

$$\bar{f}_{X_{k+1}|Y_k}(x_{k+1}|\boldsymbol{y}_k) = \sum_{i=1}^{n^{k+1|k}} \mathcal{G}_i^{k+1|k} \mathcal{E}_i^{k+1|k} \tag{2.49}$$

where

$$\begin{aligned} \mathcal{G}_i^{k+1|k} &= \rho_{i0}^{k+1|k} + \sum_{l=1}^{N_i^{k+1|k}} \rho_{il}^{k+1|k} \mathrm{sgn}\left(\xi_{il}^{k+1|k} - x_{k+1}\right) \\ \mathcal{E}_i^{k+1|k} &= \exp\left(-\sum_{l=1}^{N_i^{k+1|k}} \eta_{il}^{k+1|k} \left|\xi_{il}^{k+1|k} - x_{k+1}\right|\right) \end{aligned} \tag{2.50}$$

Thus, we have shown by induction that the update and propagation algorithm is recursive for $\bar{f}_{X_k|Y_{k-1}}$ in (2.43).

2.4 Properties of the ucpdf

By observing the first few ucpdfs through $k = 4$, we can deduce a pattern in the number of terms as well as identify which terms contain which elements, allowing us to know *a priori* which terms will combine. The first few terms counts are shown in Table 2.1 and it is clear that the number of terms is

$$n^{k|k-1} = \sum_{i=1}^{k} i = \frac{k(k+1)}{2}. \tag{2.51}$$

k	Terms	Elements
1	1	2
2	3	1+2
3	6	1+2+3
4	10	1+2+3+4

Table 2.1: Term count for $\mathbb{R}$

The number of elements (or sign functions) are deduced in a similar way to be

$$N^{k|k-1} = \begin{bmatrix} \mathbf{1}^k & \times 1 \\ \mathbf{1}^{k-1} & \times 2 \\ \mathbf{1}^{k-2} & \times 3 \\ \vdots & \\ \mathbf{1}^1 & \times k \end{bmatrix}, \tag{2.52}$$

where $\mathbf{1}^k$ is a column vector of length k composed of 1's. There are k terms with one element, $k-1$ terms with two elements, $k-2$ terms with three elements, $\cdots$, and one term with k elements.

An important property of the conditional pdf is that it is log-concave and unimodal. Consider the joint density function of initial condition and measurements, which are then convolved with the process noise. Since the measurement and process noise densities are log-concave, and log-concavity is preserved under multiplication and convolution, it follows that the conditional density function at time step k is also log-concave. Furthermore, since the conditional pdf is a marginalization of the log-concave joint density function, then it is also log-concave and, therefore, unimodal. This property allows for the use a standard tools to determine the maximum *a posteriori* (MAP) estimate as an alternative to computing the conditional mean.

2.5 Mean and Variance

The mean, μ and variance σ^2 are given as

$$\begin{aligned}
\mu_{X_k|Y_k} &= E[X_k|Y_k] \\
\sigma^2_{X_k|Y_k} &= E[(X_k - \mu_{X_k|Y_k})^2|Y_k] \\
&= E[X_k^2|Y_k] - 2\mu_{X_k|Y_k}E[X_k|Y_k] + \mu^2_{X_k|Y_k} \\
&= E[X_k^2|Y_k] - \mu^2_{X_k|Y_k}
\end{aligned} \tag{2.53}$$

where $E[X_k|Y_k]$ and $E[X_k^2|Y_k]$ are the first and second moments of the normalized conditional pdf $f_{X_k|Y_k}$, respectively.

Therefore, we first need to normalize the *a posteriori* ucpdf $\bar{f}_{X_k|Y_k}$, which is given by (2.43) and (2.45), or more explicitly,

$$\bar{f}_{X_k|Y_k}(x_k|\boldsymbol{y}_k) = \left(\sum_{i=1}^{n^{k|k-1}+1} \mathcal{G}_i^{k|k-1}\mathcal{E}_i^{k|k-1}\right) \cdot \frac{1}{2\gamma}\exp\left(-\frac{|H|}{\gamma}\left|\frac{z_k}{H} - x_k\right|\right) \tag{2.54}$$

where $\mathcal{G}_i^{k|k-1}$ and $\mathcal{E}_i^{k|k-1}$ are given in (2.44). We can ignore the $\frac{1}{2\gamma}$, since we will normalize anyway, and re-index the terms to get

$$\bar{f}_{X_k|Y_k}(x_k|\boldsymbol{y}_k) = \sum_{i=1}^{n^{k|k}} \mathcal{G}_i^{k|k}\mathcal{E}_i^{k|k} \tag{2.55}$$

where

$$\begin{aligned}
\mathcal{G}_i^{k|k} &= \rho_{i0}^{k|k} + \sum_{l=1}^{N_i^{k|k}} \rho_{il}^{k|k}\operatorname{sgn}\left(\xi_{il}^{k|k} - x_k\right) \\
\mathcal{E}_i^{k|k} &= \exp\left(-\sum_{l=1}^{N_i^{k|k-1}} \eta_{il}^{k|k-1}\left|\xi_{il}^{k|k-1} - x_k\right| - \frac{|H|}{\gamma}\left|\frac{z_k}{H} - x_k\right|\right) \\
&\triangleq \exp\left(-\sum_{l=1}^{N_i^{k|k}} \eta_{il}^{k|k}\left|\xi_{il}^{k|k} - x_k\right|\right)
\end{aligned} \tag{2.56}$$

Note that the only substantial difference between $\bar{f}_{X_k|Y_k}$ and $\bar{f}_{X_k|Y_{k-1}}$ is the additional element from f_V.

2.5.1 Characteristic function

Instead of integrating for them directly, the normalization and moments can be obtained by using the characteristic function of the ucpdf, which is the Fourier transform of $\bar{f}_{X|Y}$. For $\nu \in \mathbb{C}, x \in \mathbb{R}$, the characteristic function is

$$\bar{\phi}_{X|Y}(\nu) = \int_{\mathbb{R}} e^{j\nu x} \bar{f}_{X|Y}(x|\boldsymbol{y})\, dx \tag{2.57}$$

The derivation of the characteristic function can be found in Section 7.3 of the Appendix, but suffice it to say that it simply is an application of the integral formula (7.3) that was used in propagation. Using (7.16) for each term of $\bar{f}^i_{X|Y} = g_i \epsilon_i$, the corresponding term in the characteristic function is

$$\bar{\phi}^i_{X|Y} = \mathcal{G}_i(\nu)\mathcal{E}_i(\nu) \tag{2.58}$$

where

$$\begin{aligned}
\mathcal{G}_i(\nu) &= \frac{g_i^\dagger(\xi_i)}{j\nu + \eta_i + \sum\limits_{\substack{l=1\\l\neq i}}^{m} \eta_i \text{sgn}(\xi_l - \xi_i)} - \frac{g_i(\xi_i)}{j\nu - \eta_i + \sum\limits_{\substack{l=1\\l\neq i}}^{m} \eta_i \text{sgn}(\xi_l - \xi_i)} \\
&\triangleq \frac{a_{i1}}{j\nu + a_{i2}} - \frac{b_{i1}}{j\nu + b_{i2}} \\
\mathcal{E}_i(\nu) &= \exp\left(-\sum_{\substack{l=1\\l\neq i}}^{m} \eta_l |\xi_l - \xi_i| + j\nu\xi_i \right) \\
&\triangleq \exp(c_{i1} + c_{i2} j\nu)
\end{aligned} \tag{2.59}$$

where $g_i^\dagger$ indicates that we use the convention $\text{sgn}(0) = 1$ instead of the usual -1 stated in (1.3). The reason for this comes from the derivation of the integral formula, which makes the indexing line up.

2.5.2 Normalization

From (7.24), the normalization factor f_Y is simply

$$f_Y = \bar{\phi}_{X|Y}(0). \tag{2.60}$$

Therefore, using (2.58), we get

$$f_Y = \sum_{i=1}^{m} \left(\frac{a_{i1}}{a_{i2}} - \frac{b_{i1}}{b_{i2}} \right) \exp\left(c_{i1}\right) \tag{2.61}$$

2.5.3 First moment and second moments at step k

The mean and variance is given by

$$\begin{aligned} \hat{x} &= E\left[X|Y\right] \\ \mathrm{Var}(x) &= E\left[X^2|Y\right] - E\left[X|Y\right]^2 \end{aligned} \tag{2.62}$$

where

$$\begin{aligned} E\left[X|Y\right] &= \left[\frac{1}{j}\frac{\partial \phi_{X|Y}(\nu)}{\partial \nu}\right]_{\nu=0} \\ E\left[X^2|Y\right] &= \left[-\frac{\partial^2 \phi_{X|Y}(\nu)}{\partial \nu^2}\right]_{\nu=0}. \end{aligned} \tag{2.63}$$

Note that the moments are derived form the *normalized* characteristic function $\phi_{X|Y}$, which is simply $\bar{\phi}_{X|Y}$ divided by f_Y found in (2.61). The proof of (2.63) is detailed in Section 7.4 of the Appendix. Using the formulas in Section 7.5 of the Appendix, the first and second moments of term i are given by

$$\begin{aligned} E[X|Y] &= \sum_{i=1}^{m} \frac{1}{j}\left[\frac{\partial G_i}{\partial \nu}\mathcal{E}_i + G_i \frac{\partial \mathcal{E}_i}{\partial \nu}\right] \\ E_i[X^2|Y] &= -\sum_{i=1}^{m}\left[\frac{\partial^2 G_i}{\partial \nu^2}\epsilon_i + 2\frac{\partial G_i}{\partial \nu}\cdot\frac{\partial \epsilon_i}{\partial \nu} + G_i\frac{\partial^2 \epsilon_i}{\partial \nu^2}\right] \end{aligned} \tag{2.64}$$

where, for brevity, the partial derivatives are shown in (7.26) and (7.28) of the Appendix.

2.5.4 Example: *a posteriori* mean of $f_{X_1|Y_1}$

Let's do generic example of the mean for the cpdf after the first update. Consider the $\bar{f}_{X_1|Y_1}$. For simplicity, let $H = 1$. From (2.13),

$$\bar{f}_{X_1|Y_1} = \frac{1}{4\alpha\gamma}\exp\left(-\frac{1}{\gamma}\left|z_1 - x_1\right| - \frac{1}{\alpha}\left|x_1\right|\right) \tag{2.65}$$

Then, from (2.61), the characteristic function is

$$\begin{aligned}
\bar{\phi}_{X_1|Y_1}(\nu) &= \left(\frac{1}{j\nu + \frac{1}{\gamma} + \frac{1}{\alpha}\mathrm{sgn}\left(-z_1\right)} - \frac{1}{j\nu - \frac{1}{\gamma} + \frac{1}{\alpha}\mathrm{sgn}\left(-z_1\right)}\right)\exp\left(-\frac{1}{\alpha}\left|z_1\right| + j\nu z_1\right) \\
&+ \left(\frac{1}{j\nu + \frac{1}{\alpha} + \frac{1}{\gamma}\mathrm{sgn}\left(z_1\right)} - \frac{1}{j\nu - \frac{1}{\alpha} + \frac{1}{\gamma}\mathrm{sgn}\left(z_1\right)}\right)\exp\left(-\frac{1}{\alpha}\left|z_1\right|\right) \\
&\triangleq \left(\frac{a_{11}}{j\nu + a_{12}} - \frac{b_{11}}{j\nu + b_{12}}\right)\exp\left(c_{11} + c_{12}\cdot j\nu\right) \\
&+ \left(\frac{a_{21}}{j\nu + a_{22}} - \frac{b_{21}}{j\nu + b_{22}}\right)\exp\left(c_{21}\right)
\end{aligned} \tag{2.66}$$

The normalization factor becomes

$$\begin{aligned}
f_{Y_1} &= \left(\frac{a_{11}}{a_{12}} - \frac{b_{11}}{b_{12}}\right)\exp\left(c_{11}\right) + \left(\frac{a_{21}}{a_{22}} - \frac{b_{21}}{b_{22}}\right)\exp\left(c_{21}\right) \\
&= \frac{2\alpha^2\gamma}{\alpha^2 - \gamma^2}\exp\left(-\frac{1}{\alpha}\left|z_1\right|\right) - \frac{2\alpha\gamma^2}{\alpha^2 - \gamma^2}\exp\left(-\frac{1}{\gamma}\left|z_1\right|\right)
\end{aligned} \tag{2.67}$$

Using (2.67) and (2.64), the first moment becomes

$$\begin{aligned}
\hat{x}_1 \triangleq E[X_1|Y_1] &= \frac{\frac{2\alpha^2\gamma^2\mathrm{sgn}(z_1)}{\alpha^2-\gamma^2}\left(e^{-\frac{1}{\gamma}|z_1|} - e^{-\frac{1}{\alpha}|z_1|}\right) + z_1\alpha e^{-\frac{1}{\alpha}|z_1|}}{\alpha e^{-\frac{1}{\alpha}|z_1|} - \gamma e^{-\frac{1}{\gamma}|z_1|}} \\
&= \frac{\frac{2\alpha^2\gamma^2\mathrm{sgn}(z_1)}{\alpha^2-\gamma^2}e^{-\frac{1}{\gamma}|z_1|}}{\alpha e^{-\frac{1}{\alpha}|z_1|} - \gamma e^{-\frac{1}{\gamma}|z_1|}} + \frac{\left(z_1\alpha - \frac{2\alpha^2\gamma^2\mathrm{sgn}(z_1)}{\alpha^2-\gamma^2}\right)e^{-\frac{1}{\alpha}|z_1|}}{\alpha e^{-\frac{1}{\alpha}|z_1|} - \gamma e^{-\frac{1}{\gamma}|z_1|}}
\end{aligned} \tag{2.68}$$

Let's consider a few extremes.

- $\alpha > \gamma$ and $z_1 \gg 0$

  In this case, we can see the that $e^{-\frac{1}{\gamma}|z_1|}$ vanishes faster than $e^{-\frac{1}{\alpha}|z_1|}$, so $e^{-\frac{1}{\alpha}|z_1|}$ and the right term of (2.68) dominate for $z_1 \gg 0$. Furthermore, the linear part of that term, $z_1\alpha$, dominates the constant part, so

$$\begin{aligned}
\hat{x}_1 &\approx \frac{\left(z_1\alpha - \frac{2\alpha^2\gamma^2\mathrm{sgn}(z_1)}{\alpha^2-\gamma^2}\right)e^{-\frac{1}{\alpha}|z_1|}}{\alpha e^{-\frac{1}{\alpha}|z_1|}} \\
&\approx \frac{z_1\alpha e^{-\frac{1}{\alpha}|z_1|}}{\alpha e^{-\frac{1}{\alpha}|z_1|}} = z_1
\end{aligned} \tag{2.69}$$

- $\alpha < \gamma$ and $z_1 \gg 0$

  In this case, we can see the that $e^{-\frac{1}{\alpha}|z_1|}$ vanishes much faster than $e^{-\frac{1}{\gamma}|z_1|}$, so $e^{-\frac{1}{\gamma}|z_1|}$ and the left term of (2.68) dominate. So

$$\begin{aligned}\hat{x}_1 &\approx \frac{\frac{2\alpha^2\gamma^2 \text{sgn}(z_1)}{\alpha^2-\gamma^2} e^{-\frac{1}{\gamma}|z_1|}}{-\gamma e^{-\frac{1}{\gamma}|z_1|}} \\ &= \frac{2\alpha^2\gamma \text{sgn}\,(z_1)}{\gamma^2 - \alpha^2}\end{aligned} \tag{2.70}$$

  is constant!

- $\alpha = \gamma$

  When $\alpha = \gamma$, it can be shown that $\hat{x}_1 = \frac{z_1}{2}$.

We can confirm these observations with a numerical example. Figure 2.5 shows x_1 as a function of z_1 for $(\alpha, \gamma) \in \{(1, 0.33), (1, 0.67), (1, 1), (0.67, 1), (0.33, 1)\}$. We can see the linear behavior when z_1 is small and when $\gamma < \alpha$, but for $\gamma > \alpha$ the conditional estimate seems to saturate for large z_1. Also, we can see that the slope is $\frac{1}{2}$ when $\alpha = \gamma$. This is consistent with our intuition that the estimate should favor the measurement when the measurement noise is lower than the prior noise. Conversely, when the measurement noise is larger, it effects should be attenuated. When they are equal, it splits the difference.

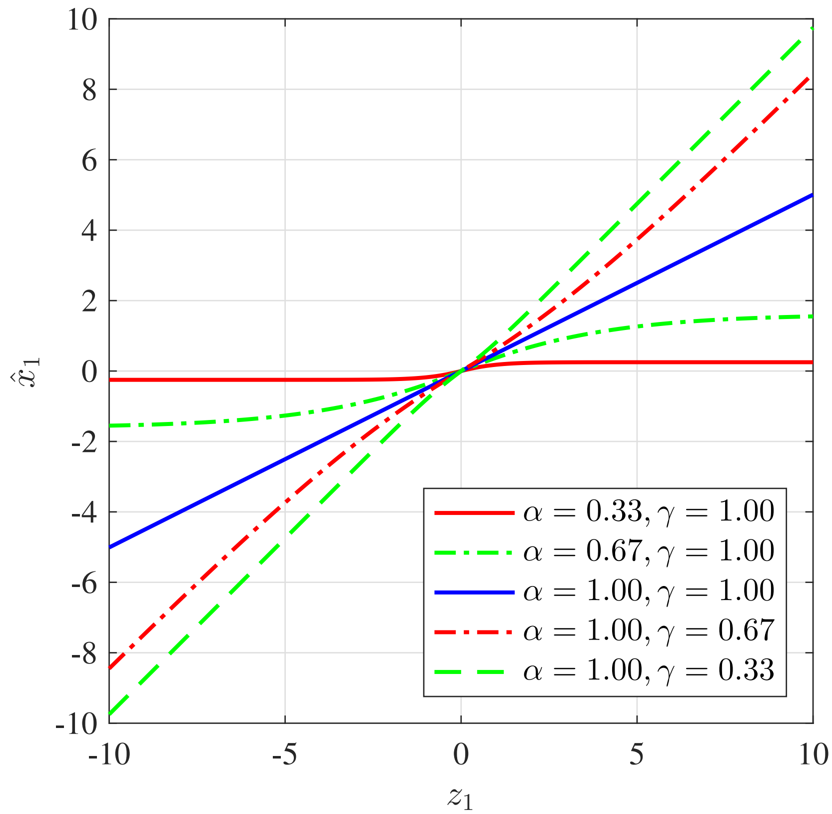


Figure 2.5: Conditional mean $\hat{x}_1$ versus measurement z_1

2.5.4.1 Example: *a posteriori* variance at $k = 1$

Let's do a generic example of the variance for the same cpdf. Again, let $H = 1$ for simplicity. Then, using (2.67) and (2.64), the second moment is then given by

$$
\begin{aligned}
E[X_1^2|Y_1] = &\left[\frac{2z_1^2\alpha^2\gamma}{(\alpha^2-\gamma^2)} - \frac{8\alpha^3\gamma^3|z_1|}{(\alpha^2-\gamma^2)^2} + \frac{4\alpha^4\gamma^3(\alpha^2+3\gamma^2)}{(\alpha^2-\gamma^2)^3}\right] \times \frac{(\alpha^2-\gamma^2)\exp\left(-\frac{1}{\alpha}|z_1|\right)}{2\alpha\gamma\left(\alpha e^{-\frac{1}{\alpha}|z_1|} - \gamma e^{-\frac{1}{\gamma}|z_1|}\right)} \\
&- \frac{4\alpha^3\gamma^4(\gamma^2+3\alpha^2)}{(\alpha^2-\gamma^2)^3} \times \frac{(\alpha^2-\gamma^2)\exp\left(-\frac{1}{\gamma}|z_1|\right)}{2\alpha\gamma\left(\alpha\exp\left(-\frac{1}{\alpha}|z_1|\right) - \gamma\exp\left(-\frac{1}{\gamma}|z_1|\right)\right)} \\
= &\left[z_1^2\alpha - \frac{4\alpha^2\gamma^2|z_1|}{\alpha^2-\gamma^2} + \frac{2\alpha^3\gamma^2(\alpha^2+3\gamma^2)}{(\alpha^2-\gamma^2)^2}\right] \times \frac{\exp\left(-\frac{1}{\alpha}|z_1|\right)}{\alpha\exp\left(-\frac{1}{\alpha}|z_1|\right) - \gamma\exp\left(-\frac{1}{\gamma}|z_1|\right)} \\
&- \frac{2\alpha^2\gamma^3(\gamma^2+3\alpha^2)}{(\alpha^2-\gamma^2)^2} \times \frac{\exp\left(-\frac{1}{\gamma}|z_1|\right)}{\alpha\exp\left(-\frac{1}{\alpha}|z_1|\right) - \gamma\exp\left(-\frac{1}{\gamma}|z_1|\right)}
\end{aligned}
\tag{2.71}
$$

Figure 2.6 shows the error variance of x_1 as a function of measurement z_1 for $(\alpha, \gamma) \in \{(1, 1), (0.5, 1), (0.33, 1)\}$. For small z_1, the error variance looks quadratic, but for large z_1, the error variance goes to a constant. It appears that while the variance is a function of the measurements, the effect of large measurements on the variance is bounded, except when the relative variances are equal. Note that, while not shown in the figure, the conditional variance is the same for equal values of $\alpha\gamma$.

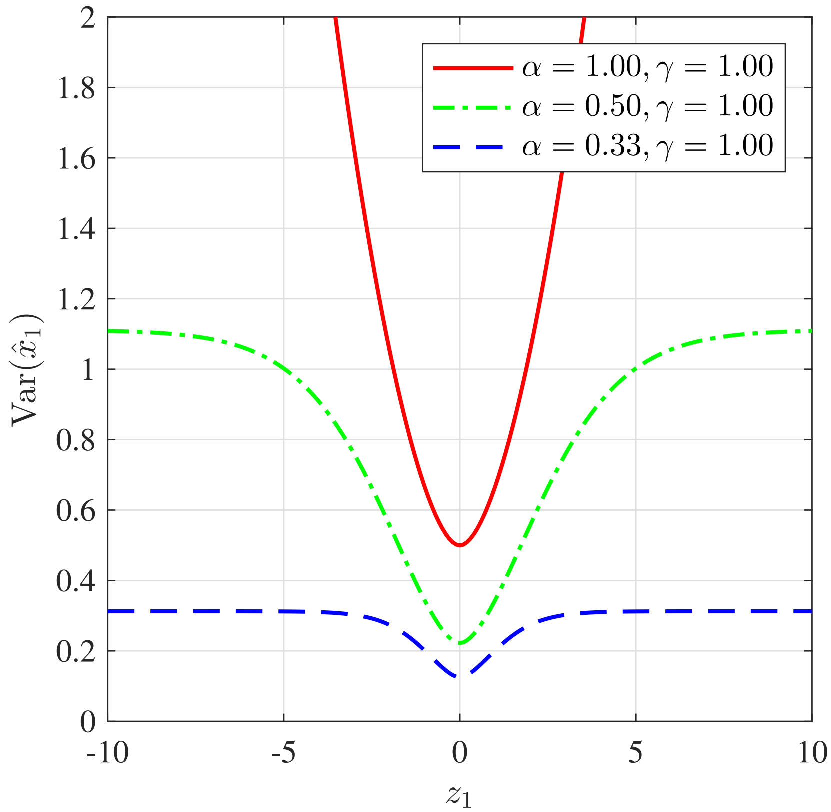


Figure 2.6: Variance versus measurement z_1

Figures 2.7 and 2.8 show the error variance of x_1 as a function of measurement z_1 for the Kalman filter, Laplace estimator, and Cauchy estimator. Using a least-squares fit between Laplace, Cauchy and Gaussian pdfs, the relationship between the noise parameters of the

Laplace distribution (γ), Cauchy distribution (ν), and Gaussian distribution (σ) are

$$\begin{aligned} \sigma &= 1.02\gamma \\ \nu &= 1.4\sigma = 1.43\gamma, \end{aligned} \tag{2.72}$$

where the coefficients were determined numerically [IS12]. We can see that the behavior of the Laplace variance approaches that of the Kalman filter as the difference between α and γ become large, while it approaches a parabolic shape similar to that of the Cauchy estimator when α approaches γ (the coefficient for the z^2 term of the Cauchy variance is $\frac{1}{4}$ [IS14]).

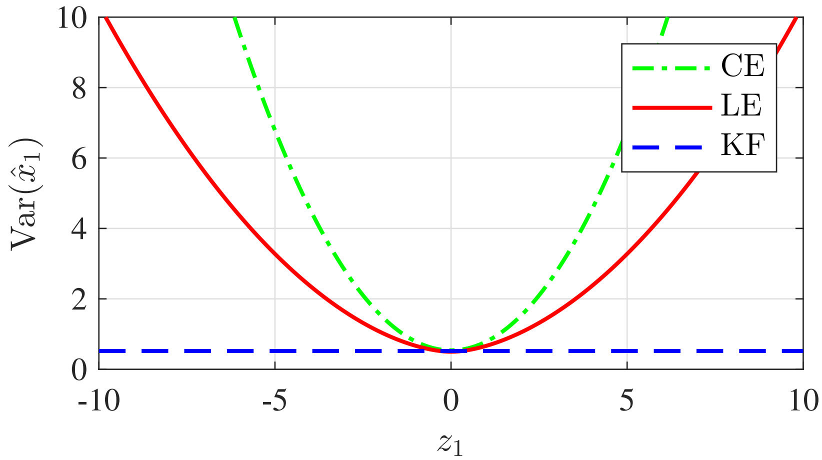


Figure 2.7: Variance versus measurement z_1 for $\alpha = 1, \gamma = 1$

2.6 Example: 50-step simulation

We recursively and analytically computed the conditional pdf as well as the first and second moments for 50 steps using the parameters $\Phi = 0.9, H = 1, \alpha = 1/5, \beta = 1/4, \gamma = 1/3$. Following the motor example in [ABL17], the measurements were generated using a Laplace random number generator, and two spikes of magnitude 10 were added to the measurements at $k = 16$ and $k = 33$ to simulate anomalies.

Figure 2.9 shows the *a posteriori* cpdf at step $k = 17$, just after the first spike, along with the conditional mean $\hat{x}$ and maximum *a posteriori* estimate x_{MAP}. We see that it is

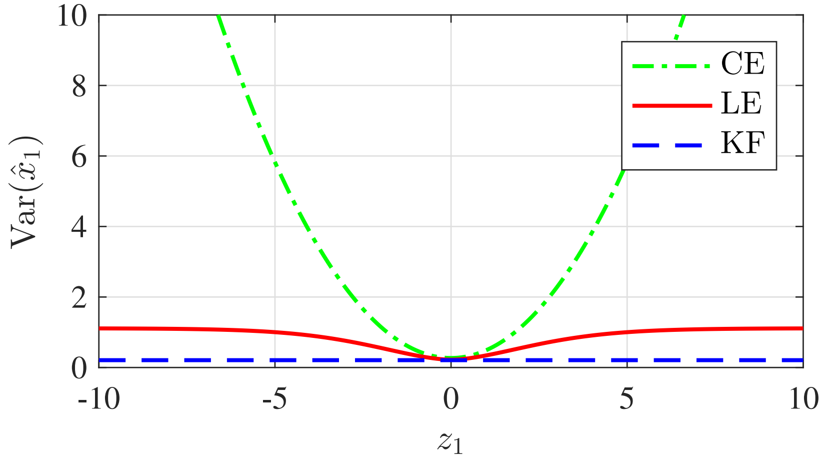


Figure 2.8: Variance versus measurement z_1 for $\alpha = 0.5, \gamma = 1$

unimodal, though $\hat{x}$ is quite different from x_{MAP} due to the asymmetry.

Figure 2.10 shows the measurement noise, maximum *a posteriori* estimation error, and conditional mean estimation error bounded by estimated standard deviations for the Laplace estimator and Kalman filter. The red dashed lines depict the computed estimation standard deviation. As in Figures 2.7 and 2.8, the Kalman filter noise parameters were determined from the Laplace noise parameters using (2.72). The Kalman estimate reacts strongly to the spikes, while its variance remains constant. In contrast, the Laplace estimate appears to attenuate the measurement spikes and the estimated standard deviation increases to account for the increased uncertainty.

Qualitatively, the maximum *a posteriori* (MAP) estimate in the second subplot of Figure 2.10 appears to track the conditional mean reasonably well. Since the conditional probability density function is unimodal, there are many convex optimization methods available to determine its maximum. This supports the work in [ABL17] and [ABB11], where MAP estimates are used as a reasonable alternative to the conditional mean.

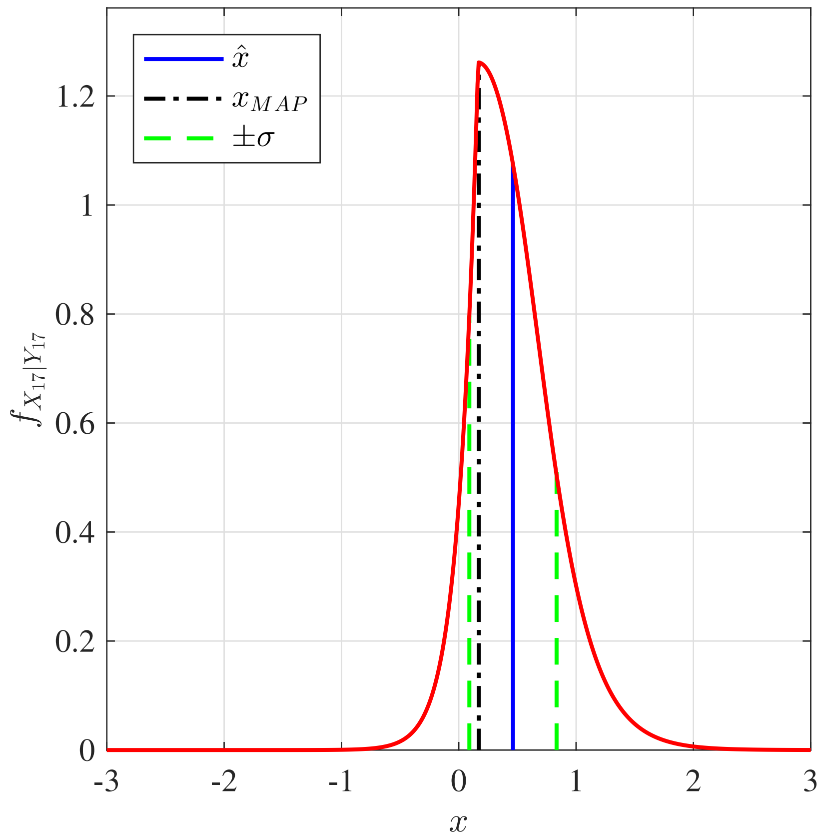


Figure 2.9: *A posteriori* conditional pdf at $k = 17$

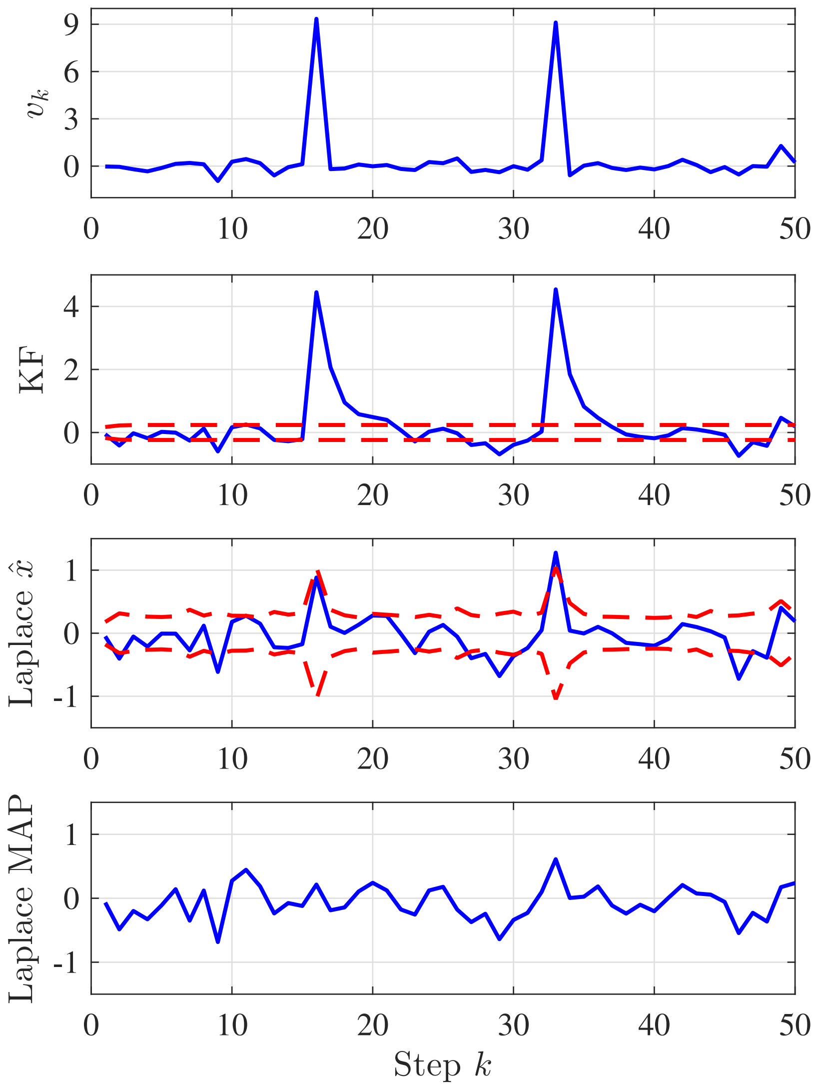


Figure 2.10: Estimation error for Laplace MAP and conditional mean, and Kalman filter

CHAPTER 3

Vector Laplace Conditional PDF

We extend the problem to the vector case with additive scalar measurement and process noise. While it is not presented here, generalizing to vector noises is straight-forward, particularly if the elements are independent. In contrast to the chapter on scalar estimation, we will save the mean and variance for the next chapter, because it is quite an involved topic on its own.

In Section 3.1, we define the general discrete linear system in $\mathbb{R}^n$. In Section 3.2, we formally show the equations used to perform the measurement and propagation steps in $\mathbb{R}^n$. In Section 3.3, we derive the first update and propagation steps in $\mathbb{R}^2$ because it becomes quite complicated to show for $\mathbb{R}^n$. During the first propagation, we extend the concept of simplifying the coefficient function to 2-dimensions. Using the lessons learned in Section 3.3, we hypothesize the general form of the ucpdf at step $k|k-1$ and prove that it is preserved under update and propagation by induction in Section 3.4. Finally, in Section 3.6, we discuss the computational complexity and then extend it to $\mathbb{R}^n$. It turns out that while the computational complexity increases, conceptually extending from $\mathbb{R}^2$ to $\mathbb{R}^n$ is rather straight-forward.

3.1 Vector dynamical system

Let the discrete linear system with state $\boldsymbol{x}_k \in \mathbb{R}^n$, scalar measurement $z_k \in \mathbb{R}$, independent scalar measurement noise v_k, and scalar process noise w_k be defined as

$$\begin{aligned}\tilde{\boldsymbol{x}}_{k+1} &= \Phi\tilde{\boldsymbol{x}}_k + \Gamma w_k \\ \tilde{z}_k &= H\tilde{\boldsymbol{x}}_k + v_k\end{aligned} \tag{3.1}$$

with $\Phi \in \mathbb{R}^{n\times n}$, $\Gamma \in \mathbb{R}^{n\times 1}$, and $H \in \mathbb{R}^{1\times n}$ and where $\boldsymbol{x}_1, w_k$ and v_k are all Laplace distributed as

$$f_{\tilde{X}_1}(\tilde{\boldsymbol{x}}_1) = \prod_{i=1}^{n} \frac{1}{2\alpha} e^{-\frac{1}{\alpha}|\tilde{x}_i - \bar{x}_i|} = \prod_{i=1}^{n} \frac{1}{2\alpha} e^{-\frac{1}{\alpha}|-\bar{x}_i + E_i\tilde{\boldsymbol{x}}|} \tag{3.2}$$

$$f_W(w_k) = \frac{1}{2\beta} e^{-\frac{1}{\beta}|w_k|} \tag{3.3}$$

$$f_V(v_k) = \frac{1}{2\gamma} e^{-\frac{1}{\gamma}|v_k|}, \tag{3.4}$$

the elements of $\tilde{\boldsymbol{x}}_1$ are mutually independent and $E_i \in \mathbb{R}^{1\times n}$ have elements 1 at i and 0 elsewhere.

For convenience, we decompose the system into deterministic and stochastic parts, so that

$$\tilde{\boldsymbol{x}}_k = \bar{\boldsymbol{x}}_k + \boldsymbol{x}_k, \tag{3.5}$$

$$\begin{aligned}\bar{\boldsymbol{x}}_{k+1} &= \Phi\bar{\boldsymbol{x}}_k \\ \bar{z}_k &= H\bar{\boldsymbol{x}}_k,\end{aligned} \tag{3.6}$$

with initial conditions $\bar{\boldsymbol{x}}_1 = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_n \end{bmatrix}^T$ and

$$\begin{aligned}\boldsymbol{x}_{k+1} &= \Phi\boldsymbol{x}_k + \Gamma w_k \\ z_k &= H\boldsymbol{x}_k + v_k\end{aligned} \tag{3.7}$$

For the remainder of the derivation, we will only consider only the stochastic part.

3.2 Laplace conditional PDF update and propagation

Given the *a priori* ucpdf $\bar{f}_{X_k|Y_{k-1}}$, the *a posteriori* pdf conditioned on the measurement sequence Y_k is, using Bayes' Theorem, given by

$$\bar{f}_{X_k|Y_k}(\boldsymbol{x}_k|\boldsymbol{y}_k) = \frac{f_{X_k|Y_{k-1}}(\boldsymbol{x}_k|\boldsymbol{y}_{k-1})f_{Z_k|X_k}(z_k|\boldsymbol{x}_k)}{f_{Y_k}(\boldsymbol{y}_k)}, \tag{3.8}$$

where

$$\boldsymbol{y}_k = \{z_1, z_2, \cdots, z_k\}. \tag{3.9}$$

Since f_{Y_k} is just a normalization factor that can be computed at any time, we will ignore it for now and, going forward, only consider the unnormalized conditional pdf

$$\begin{aligned} \bar{f}_{X_k|Y_k}(\boldsymbol{x}_k|\boldsymbol{y}_k) &= \bar{f}_{X_k|Y_{k-1}}(\boldsymbol{x}_1|\boldsymbol{y}_{k-1})f_{Z_k|X_k}(z_k|\boldsymbol{x}_k) \\ &= \bar{f}_{X_k|Y_{k-1}}(\boldsymbol{x}_k|\boldsymbol{y}_{k-1})f_V(z_k - H\boldsymbol{x}_k). \end{aligned} \tag{3.10}$$

To determine the *a priori* unnormalized conditional density $\bar{f}_{X_{k+1}|Y_k}(\boldsymbol{x}_{k+1}|\boldsymbol{y}_k)$, we first construct the joint density $f_{X_{k+1},W|Y_k}$ from

$$\bar{f}_{X_k,W_k|Y_k}(\boldsymbol{x}_k, w_k|\boldsymbol{y}_k) = \bar{f}_{X_k|Y_k}(\boldsymbol{x}_k|\boldsymbol{y}_k)f_W(w_k). \tag{3.11}$$

Then, we make the change of variables from $\boldsymbol{x}_k$ to $(\boldsymbol{x}_{k+1}, w_k)$ using

$$\begin{aligned} \begin{bmatrix} \boldsymbol{x}_{k+1} \\ w_k \end{bmatrix} &= \begin{bmatrix} \Phi & \Gamma \\ 0 & I \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_k \\ w_k \end{bmatrix} \triangleq A \begin{bmatrix} \boldsymbol{x}_k \\ w_k \end{bmatrix} \\ \implies \begin{bmatrix} \boldsymbol{x}_k \\ w_k \end{bmatrix} &= \begin{bmatrix} \Phi^{-1} & -\Phi^{-1}\Gamma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{k+1} \\ w_k \end{bmatrix} \triangleq A^{-1} \begin{bmatrix} \boldsymbol{x}_{k+1} \\ w_k \end{bmatrix} \end{aligned} \tag{3.12}$$

[SC08, p. 51] and integrate with respect to w_k

$$\begin{aligned} \bar{f}_{X_{k+1}|Y_k}(\boldsymbol{x}_{k+1}|\boldsymbol{y}_k) = \left|\Phi^{-1}\right| \times \\ \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k}\left(\Phi^{-1}\boldsymbol{x}_{k+1} - \Phi^{-1}\Gamma w_k|\boldsymbol{y}_k\right) f_W(w_k)\, dw_k, \end{aligned} \tag{3.13}$$

where $\det\left(A^{-1}\right) = \det\left(\Phi^{-1}\right) \triangleq |\Phi^{-1}|$. Using these general steps, we determine the recursive algorithm for generating the *a priori* ucpdf given the parameters of the *a priori* ucdpf from the previous step.

3.3 First update and propagation in $\mathbb{R}^2$ to motivate general form

We explicitly compute the first update and propagation steps for the 2-dimensional case to motivate the general structure of the unnormalized conditional pdf (ucpdf) as well as the recursive algorithm to generate successive *a posteriori* and *a priori* ucpdfs at any step k for the general n-dimensional case.

3.3.1 Initial conditions for $\boldsymbol{x} \in \mathbb{R}^2$

Let the elements of initial condition be described by independent Laplace distributions with means 0 and with spread parameter α so that

$$\begin{aligned} f_{X_1}(\boldsymbol{x}_1) &= \frac{1}{2\alpha} \exp\left(-\frac{1}{\alpha}|x_1|\right) \cdot \frac{1}{2\alpha} \exp\left(-\frac{1}{\alpha}|x_2|\right) \\ &= \frac{1}{4\alpha^2} \exp\left(-\frac{1}{\alpha}|E_1\boldsymbol{x}_1| - \frac{1}{\alpha}|E_2\boldsymbol{x}_1|\right) \end{aligned} \tag{3.14}$$

where

$$E_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}. \tag{3.15}$$

3.3.2 Update at $k = 1$ for $\boldsymbol{x} \in \mathbb{R}^2$

Let the measurement at $k = 1$ be z_1 so that, using (3.4) and (3.7), the density of Z_1 conditioned on X_1 is

$$f_{Z_1|X_1} = \frac{1}{2\gamma} \exp\left(-\frac{1}{\gamma}|z_1 - H\boldsymbol{x}_1|\right) = f_V\left(-\frac{1}{\gamma}|z_1 - H\boldsymbol{x}_1|\right) \tag{3.16}$$

The unnormalized pdf of X_1 conditioned on Z_1 is the joint density

$$\begin{aligned} \bar{f}_{X_1|Y_1} &= f_{X_1}(\boldsymbol{x}_1) f_V(z_1 - H\boldsymbol{x}_1) \\ &= \frac{1}{8\alpha^2\gamma} \exp\left(-\frac{1}{\alpha}|E_1\boldsymbol{x}_1| - \frac{1}{\alpha}|E_2\boldsymbol{x}_1| - \frac{1}{\gamma}|z_1 - H\boldsymbol{x}_1|\right) \end{aligned} \tag{3.17}$$

which is a product since V is independent of X_1. Since we can normalize at any time, we will carry forward the joint density function, or unnormalized conditional density function (ucpdf), without the coefficient $\frac{1}{8\alpha^2\gamma}$.

Figure 3.1 shows $\bar{f}_{X_1|Y_1}$ with $\alpha = 0.5$ and $\gamma = 0.4$. Notice the asymmetry and kink on the left side, which is the peak of $f_V(z_1 - Hx_2)$.

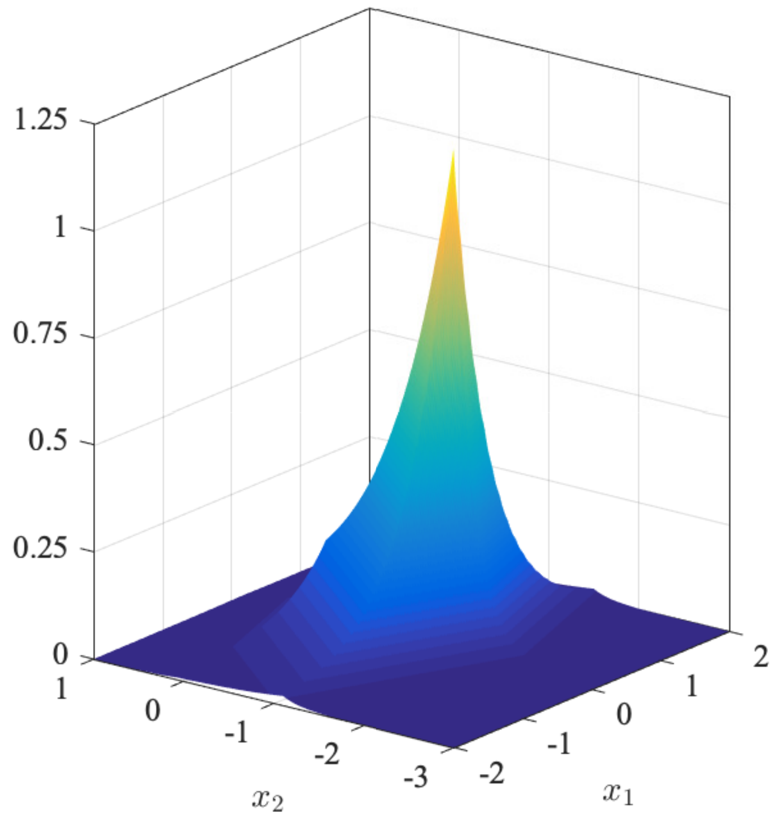


Figure 3.1: $\bar{f}_{X_1|Y_1}$: *a posteriori* unnormalized conditional pdf at $k = 1$

3.3.3 Propagation from $k = 1$ to $k = 2$ for $x \in \mathbb{R}^2$

To determine the ucpdf $\bar{f}_{X_2|Y_1}(\boldsymbol{x}_2|\boldsymbol{y}_1)$, we first construct the joint density $f_{X_2,W|Y_1}$ from

$$f_{X_1,W|Y_1} = f_{X_1|Y_1} f_W, \tag{3.18}$$

which is again a product due to the independence of W, where f_W is given by (3.3). Then, we make the change of variables from $\boldsymbol{x}_1$ to $(\boldsymbol{x}_2, w_1)$ and then integrate with respect to w_1.

For the initial conditions, the exponents become

$$\begin{aligned} -\frac{1}{\alpha}\left|\bar{x}_{1i} - E_i \boldsymbol{x}_1\right| &= -\frac{1}{\alpha}\left|E_i\left(\Phi^{-1}\boldsymbol{x}_2\right) - E_i\Phi^{-1}\Gamma w_1\right| \\ &= -\frac{\left|E_i\Phi^{-1}\Gamma\right|}{\alpha}\left|\frac{E_i\Phi^{-1}\boldsymbol{x}_2}{E_i\Phi^{-1}\Gamma} - w_1\right| \end{aligned} \tag{3.19}$$

where $i = 1, 2$, while the exponent for the measurement becomes

$$\begin{aligned} -\frac{1}{\gamma}\left|z_1 - H\boldsymbol{x}_1\right| &= -\frac{1}{\alpha}\left|-\bar{z}_1 + H\left(\Phi^{-1}\boldsymbol{x}_2\right) - H\Phi^{-1}\Gamma w_1\right| \\ &= -\frac{\left|H\Phi^{-1}\Gamma\right|}{\gamma}\left|\frac{-z_1}{H\Phi^{-1}\Gamma} + \frac{H\Phi^{-1}\boldsymbol{x}_2}{H\Phi^{-1}\Gamma} - w_1\right| \end{aligned} \tag{3.20}$$

Using (3.3) and (3.17), (3.13) becomes explicitly in terms of the Laplace densities

$$\begin{aligned} \bar{f}_{X_2|Y_1}(\boldsymbol{x}_2|\boldsymbol{y}_1) = \frac{|\Phi|^{-1}}{2\beta}\int_{-\infty}^{\infty}\exp\Bigg(&-\frac{\left|E_1\Phi^{-1}\Gamma\right|}{\alpha}\left|\frac{E_1\Phi^{-1}\boldsymbol{x}_2}{E_1\Phi^{-1}\Gamma} - w_1\right| \\ &-\frac{\left|E_2\Phi^{-1}\Gamma\right|}{\alpha}\left|\frac{E_2\Phi^{-1}\boldsymbol{x}_2}{E_2\Phi^{-1}\Gamma} - w_1\right| \\ &-\frac{\left|H\Phi^{-1}\Gamma\right|}{\gamma}\left|\frac{-z_1}{H\Phi^{-1}\Gamma} + \frac{H\Phi^{-1}\boldsymbol{x}_2}{H\Phi^{-1}\Gamma} - w_1\right| - \frac{1}{\beta}\left|w_1\right|\Bigg)\, dw_1, \end{aligned} \tag{3.21}$$

where it is assumed that $E_1\Phi^{-1}\Gamma \neq 0$, $E_2\Phi^{-1}\Gamma \neq 0$, and $H\Phi^{-1}\Gamma \neq 0$. As before, we will ignore the coefficient $\frac{\left|\Phi^{-1}\right|^{-1}}{2\beta}$. Using the integral formula (7.59), and defining pre-integration parameters $\bar{\rho}_i^{2|1}$ and $\bar{\xi}_i^{2|1}$ as

$$\begin{aligned} \bar{\rho}_1^{2|1} &= \frac{\left|E_1\Phi^{-1}\Gamma\right|}{\gamma}, & \bar{\xi}_1^{2|1} &= \frac{E_1\Phi^{-1}\boldsymbol{x}_2}{E_1\Phi^{-1}\Gamma} \triangleq \bar{\psi}_1^{2|1} + \bar{\theta}_1^{2|1}\boldsymbol{x}_2, \\ \bar{\rho}_2^{2|1} &= \frac{\left|E_2\Phi^{-1}\Gamma\right|}{\gamma}, & \bar{\xi}_2^{2|1} &= \frac{E_2\Phi^{-1}\boldsymbol{x}_2}{E_2\Phi^{-1}\Gamma} \triangleq \bar{\psi}_2^{2|1} + \bar{\theta}_2^{2|1}\boldsymbol{x}_2, \\ \bar{\rho}_3^{2|1} &= \frac{\left|H\Phi^{-1}\Gamma\right|}{\gamma}, & \bar{\xi}_3^{2|1} &= \frac{-z_1}{H\Phi^{-1}\Gamma} + \frac{H\Phi^{-1}\boldsymbol{x}_2}{H\Phi^{-1}\Gamma} \triangleq \bar{\psi}_3^{2|1} + \bar{\theta}_3^{2|1}\boldsymbol{x}_2, \\ \bar{\rho}_4^{2|1} &= \frac{1}{\beta}, & \bar{\xi}_4^{2|1} &= 0 \triangleq \bar{\psi}_4^{2|1} + \bar{\theta}_4^{2|1}\boldsymbol{x}_2, \end{aligned} \tag{3.22}$$

the solution to (3.21) is

$$\begin{aligned}\bar{f}_{X_2|Y_1}(\boldsymbol{x}_2|\boldsymbol{y}_1) &= \frac{|\Phi|^{-1}}{4\beta\gamma}\sum_{j=1}^{4} g_j^{2|1}\exp\left(-\sum_{\substack{l=1\\l\neq j}}^{4}\rho_l^{2|1}\left|\xi_l^{2|1}-\xi_j^{2|1}\right|\right)\\ &\triangleq \frac{|\Phi|^{-1}}{4\beta\gamma}\sum_{j=1}^{4} g_j^{2|1}\epsilon_j^{2|1}\end{aligned} \tag{3.23}$$

where

$$\begin{aligned}g_j^{2|1}(\boldsymbol{x}_2) = & \frac{1}{\bar{\rho}_j^{2|1}+\sum\limits_{\substack{l=1\\l\neq j}}^{4}\bar{\rho}_l^{2|1}\mathrm{sgn}\left(\bar{\xi}_l^{2|1}-\bar{\xi}_j^{2|1}\right)} - \frac{1}{-\bar{\rho}_j^{2|1}+\sum\limits_{\substack{l=1\\l\neq j}}^{4}\rho_l^{2|1}\mathrm{sgn}\left(\bar{\xi}_l^{2|1}-\bar{\xi}_j^{2|1}\right)}\\ &= \frac{1}{\bar{\rho}_j^{2|1}+\sum\limits_{\substack{l=1\\l\neq j}}^{4}\bar{\rho}_l^{2|1}\mathrm{sgn}\left(\left(\bar{\psi}_l^{2|1}-\bar{\psi}_j^{2|1}\right)+\left(\bar{\theta}_l^{2|1}-\bar{\theta}_j^{2|1}\right)\boldsymbol{x}_2\right)}\\ &\quad - \frac{1}{-\bar{\rho}_j^{2|1}+\sum\limits_{\substack{l=1\\l\neq j}}^{4}\rho_l^{2|1}\mathrm{sgn}\left(\left(\psi_l^{2|1}-\psi_j^{2|1}\right)+\left(\theta_l^{2|1}-\theta_j^{2|1}\right)\boldsymbol{x}_2\right)}\\ &\triangleq \frac{1}{\bar{\rho}_j^{2|1}+\sum\limits_{\substack{l=1\\l\neq j}}^{4}\rho_l^{2|1}\mathrm{sgn}\left(\psi_{jl}^{2|1}+\theta_{jl}^{2|1}\boldsymbol{x}_2\right)} - \frac{1}{-\bar{\rho}_j^{2|1}+\sum\limits_{\substack{l=1\\l\neq j}}^{4}\rho_l^{2|1}\mathrm{sgn}\left(\psi_{jl}^{2|1}+\theta_{jl}^{2|1}\boldsymbol{x}_2\right)}\end{aligned} \tag{3.24}$$

and,

$$\begin{aligned}\epsilon_j^{2|1}(\boldsymbol{x}_2) &= \exp\left(-\sum_{\substack{l=1\\l\neq j}}^{4}\bar{\rho}_l^{2|1}\left|\bar{\xi}_l^{2|1}-\bar{\xi}_j^{2|1}\right|\right)\\ &= \exp\left(-\sum_{\substack{l=1\\l\neq j}}^{4}\bar{\rho}_l^{2|1}\left|\left(\bar{\psi}_l^{2|1}-\bar{\psi}_j^{2|1}\right)+\left(\bar{\theta}_l^{2|1}-\bar{\theta}_j^{2|1}\right)\boldsymbol{x}_2\right|\right)\\ &\triangleq \exp\left(-\sum_{\substack{l=1\\l\neq j}}^{4}\bar{\rho}_l^{2|1}\left|\psi_{jl}^{2|1}+\theta_{jl}^{2|1}\boldsymbol{x}_2\right|\right)\end{aligned} \tag{3.25}$$

Note that the arguments of the sign functions are converted into standard form

$$\xi_{jl}^{2|1} = \psi_{jl}^{2|1}+\theta_{jl}^{2|1}\boldsymbol{x}_2 \tag{3.26}$$

3.3.4 Simplify g after propagation for $\boldsymbol{x} \in \mathbb{R}^2$

As in the scalar case, the form of g_j in (3.24) is not suitable for recursion. Furthermore, it is impractical to combine any terms while dealing with such an unwieldy coefficient function. However, we can extend Theorem 2.2.1 to n-dimensions to simplify this.

Theorem 3.3.1. *Let $\mathcal{H}$ be a hyperplane arrangement of m affine hyperplanes $H_1, H_2, \cdots, H_m$, where $H_i = \{\boldsymbol{x} | \psi_i + \theta_i \boldsymbol{x}\}$, $\boldsymbol{x} \in \mathbb{R}^n, \psi_i \in \mathbb{R}$. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a function of $sgn(\psi_i + \theta \boldsymbol{x})$, constant on the faces of $\mathcal{A}$. Then,*

$$g = \rho_0 + \sum_{i=1}^{n} \sum_{j=1}^{\binom{m}{i}} \rho_{ij} \prod_{l \in \sigma_{ij}} sgn(\psi_l + \theta_l \boldsymbol{x}) \tag{3.27}$$

where

$$\sigma_i = \{\{1\}, \cdots, \{m\}, \cdots, \{1,2\}, \cdots, \{m-1, m\}, \cdots, \{m-n+1, \cdots, m\}\} \tag{3.28}$$

is the set of all subsets of $\{1, \cdots, m\}$ with at most n elements. The proof is provided by collaborators from the Technion in Israel and an excerpt of their work is included in Section 7.7 of the Appendix.

Theorem 3.3.1 has a profound effect on the general n-dimensional Laplace estimator. Not only does it make more obvious how we can preserve the structure of the ucpdf under propagation, it provides a straight-forward framework for combining terms simply by adding coefficients. Furthermore, it enables a parallel effort for the n-dimensional Cauchy estimator as well.

Section 7.6 of the Appendix describes a straight-forward algorithm for determining the coefficients of the simplified (or flattened) $g_j(\boldsymbol{x}_2)$. For now, we will use it to flatten our relatively simple coefficient functions. After flattening, each $g_j(\boldsymbol{x}_2)$ assumes the form

$$\begin{aligned}
g_j^{2|1}(\boldsymbol{x}_2) &= \rho_{j0}^{2|1} + \rho_{j1}^{2|1}\operatorname{sgn}\left(\psi_{j1}^{2|1} + \theta_{j1}^{2|1}\boldsymbol{x}_2\right) + \rho_{j2}^{2|1}\operatorname{sgn}\left(\psi_{j2}^{2|1} + \theta_{j2}^{2|1}\boldsymbol{x}_2\right) + \rho_{j3}^{2|1}\operatorname{sgn}\left(\psi_{j3}^{2|1} + \theta_{j3}^{2|1}\boldsymbol{x}_2\right) \\
&\quad + \rho_{j4}^{2|1}\operatorname{sgn}\left(\psi_{j1}^{2|1} + \theta_{j1}^{2|1}\boldsymbol{x}_2\right)\operatorname{sgn}\left(\psi_{j2}^{2|1} + \theta_{j2}^{2|1}\boldsymbol{x}_2\right) \\
&\quad + \rho_{j5}^{2|1}\operatorname{sgn}\left(\psi_{j1}^{2|1} + \theta_{j1}^{2|1}\boldsymbol{x}_2\right)\operatorname{sgn}\left(\psi_{j3}^{2|1} + \theta_{j3}^{2|1}\boldsymbol{x}_2\right) \\
&\quad + \rho_{j6}^{2|1}\operatorname{sgn}\left(\psi_{j2}^{2|1} + \theta_{j2}^{2|1}\boldsymbol{x}_2\right)\operatorname{sgn}\left(\psi_{j3}^{2|1} + \theta_{j3}^{2|1}\boldsymbol{x}_2\right) \\
&= \rho_{j0}^{2|1} + \sum_{i=1}^{6}\rho_{ji}^{2|1}\prod_{l\in\sigma_{ji}}\operatorname{sgn}\left(\psi_{jl}^{2|1} + \theta_{jl}^{2|1}\boldsymbol{x}_2\right)
\end{aligned} \tag{3.29}$$

and

$$\sigma = \{\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}\} \tag{3.30}$$

Figure 3.2 shows $\bar{f}_{X_2|Y_1}$ with $\alpha = 0.5$, $\beta = 0.3$ and $\gamma = 0.4$. Note the smoothness of the pdf after convolution with f_W compared to the sharp point of $\bar{f}_{X_1|Y_1}$ in Figure 3.1.

3.3.5 Generalized integral formula

The simplification of the coefficient function in (3.29) introduces one complication. The integral formula that we've been using the propagate $\bar{f}_{X_k|Y_k}$ is not valid for when g is not a function of sums of signs. Therefore, we re-derive the integral formula to account for the products of sign functions and show examples for both $x \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^2$ in Section 7.8 of the Appendix. The results in the relevant form in (7.59) are restated here for convenience. For $x \in \mathbb{R}$ and

$$g(x) = \rho_0 + \sum_{i=1}^{m}\rho_i\prod_{\ell\in\sigma_i}\operatorname{sgn}(\xi_\ell - x), \tag{3.31}$$

the integral

$$\begin{aligned}
I &= \int_{-\infty}^{\infty} g(x)\exp\left(-\sum_{i=1}^{n}\eta_i\left|\xi_i - x\right| + j\nu x\right)dx \\
&= \sum_{i=1}^{n} G_i\exp\left(-\sum_{\ell=1}^{n}\rho_\ell(\xi_\ell - \xi_i)s_i^\ell + j\nu\xi_i\right)
\end{aligned} \tag{3.32}$$

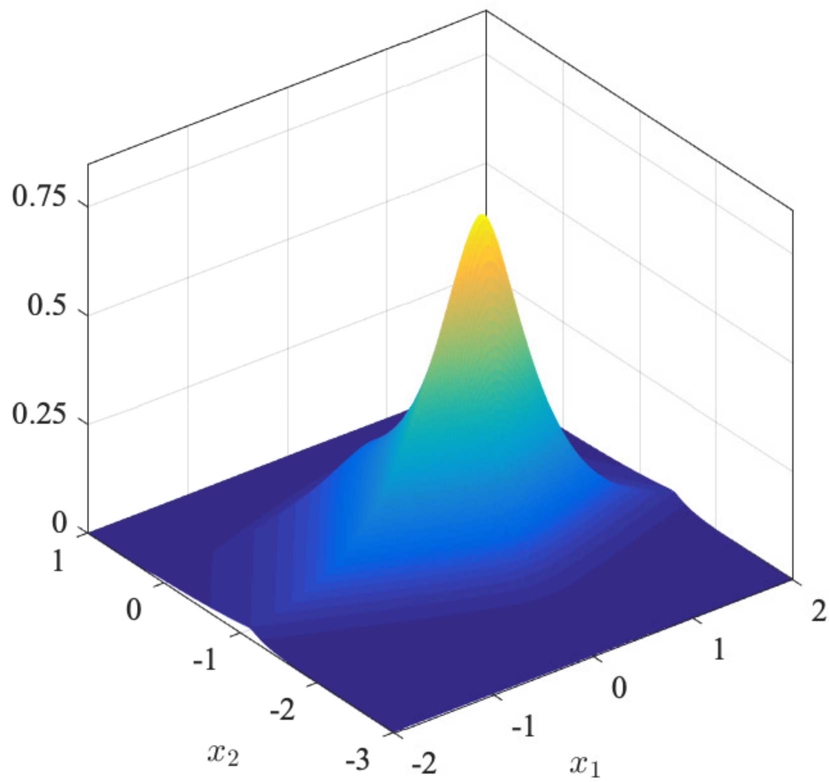


Figure 3.2: $\bar{f}_{X_2|Y_1}$: *a priori* unnormalized conditional pdf at $k = 2$

where

$$G_i = \left[\frac{g(\xi_i)}{j\nu + \sum_{\ell=1}^{n} \eta_\ell \mathrm{sgn}\left(\xi_\ell - \xi_i\right)} \right]^\dagger - \frac{g(\xi_i)}{j\nu + \sum_{\ell=1}^{n} \eta_\ell \mathrm{sgn}\left(\xi_\ell - \xi_i\right)} \tag{3.33}$$

and the † indicates that $\mathrm{sgn}\,(0) = 1$ instead of our usual convention $\mathrm{sgn}\,(0) = -1$. The reason for this comes from the derivation, when we make a shift from ξ_{i+1} to ξ_i for convenience but causes a change in convention at that particular location. This integral formula allows for the propagation in $\mathbb{R}^n$ as well as calculation of the moments in the next chapter.

3.4 Proof of general recursive form for $\mathbb{R}^n$ by induction

We now present the recursive algorithm for determining the *a posteriori* and *a priori* ucpdf at step $k+1|k$ by undergoing a measurement update from step $k|k-1$ to $k|k$, followed by a convolution integral. Finally, we do some bookkeeping to organize it in the standard structure.

3.4.1 Induction hypothesis in $\mathbb{R}^2$

Given the *a priori* ucpdf at $k|k-1$

$$\bar{f}_{X_k|Y_{k-1}} = \sum_{i=1}^{N_i^{k|k-1}} G_i^{k|k-1} \epsilon_i^{k|k-1} \tag{3.34}$$

where

$$G_i^{k|k-1} = \rho_{i0} + \sum_{j=1}^{P_i^{k|k-1}} \rho_{ij}^{k|k-1} \prod_{l \in \sigma_{ij}} \operatorname{sgn}\left(\xi_{il}^{k|k-1}\right) \tag{3.35}$$

and

$$\epsilon_i^{k|k-1} = \exp\left(-\sum_{j=1}^{M_i^{k|k-1}} \eta_{ij}^{k|k-1} \left|\xi_{ij}^{k|k-1}\right|\right) \tag{3.36}$$

where

$$\xi_{ij}^{k|k-1} = \psi_{ij}^{k|k-1} + \theta_{ij}^{k|k-1} x_k, \tag{3.37}$$

$N_i^{k|k-1}$ is the number of terms, $M_i^{k|k-1}$ is the number of elements of term i, and $P_i^{k|k-1}$ is the number of terms in $G_i^{k|k-1}$, and σ_{ij} is the set of indices of sign functions associated with term j of $G_i^{k|k-1}$ at step $k|k-1$.

3.4.2 Base case

By construction, we know that the ucpdf at 2|1 from Section 3.3 have the form shown in (3.34). Next, we assume $\bar{f}_{X_k|Y_{k-1}}$ in (3.34) and show that $\bar{f}_{X_{k+1},Y_k}$ has the same structure.

3.4.3 Measurement update from $k|k-1$ to $k|k$

The measurement update step involved constructing the joint density from the *a priori* ucpdf and

$$f_{Z_k|X_k} = f_V(z_k - H\boldsymbol{x}_k) = \frac{1}{2\gamma} \exp\left(-\frac{1}{\gamma} |z_k - H\boldsymbol{x}_k| \right). \tag{3.38}$$

Since V is independent noise, this is simply the product, and effectively contributes an additional element to the exponential so that

$$\bar{f}_{X_k|Y_k} = \sum_{i=1}^{N_i^{k|k-1}+1} G_i^{k|k} \epsilon_i^{k|k} \tag{3.39}$$

where

$$G_i^{k|k} = \frac{1}{2\gamma} G_i^{k|k-1} \tag{3.40}$$

and

$$\begin{aligned} \epsilon_i^{k|k} &= \exp\left(-\sum_{j=1}^{M_i^{k|k-1}} \eta_{ij}^{k|k-1} \left| \xi_{ij}^{k|k-1} \right| - \frac{1}{\gamma} \left| -z_k + Hx_k \right| \right) \\ &= \exp\left(-\sum_{j=1}^{M_i^{k|k-1}+1} \eta_{ij}^{k|k-1} \left| \xi_{ij}^{k|k-1} \right| \right) \end{aligned} \tag{3.41}$$

with

$$\xi_{ij}^{k|k-1} = \psi_{ij}^{k|k-1} + \theta_{ij}^{k|k-1} x \tag{3.42}$$

This gives us the *a posteriori* ucpdf at $k|k$, with the additional parameters

$$\rho^* = \frac{1}{\gamma}, \qquad \psi^* = -z_k, \qquad \theta^* = H \tag{3.43}$$

3.4.4 Propagation from $k|k$ to $k+1|k$

The propagation for the exponential part is a convolution with f_W and can be done in two stages. In the first stage, the parameters are updated with those of f_W and a change of variables is performed to get it into a form that will go into the second stage, the integration.

Similarly to the exponential part, the coefficient portion also undergoes a two-stage propagation. First, we perform a change of variables. Then, we augment each coefficient function with a pair of zeros for f_V and f_W.

Given (3.39), the *a priori* ucpdf at $k+1$ is

$$\bar{f}_{X_{k+1}|Y_k}(\boldsymbol{x}_{k+1}|\boldsymbol{y}_k) = \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k}(\boldsymbol{x}_k|\boldsymbol{y}_k) f_W(w_k)\, dw_k \tag{3.44}$$

In order to perform this integral, we perform a change of variables from $\boldsymbol{x}_k$ to $\boldsymbol{x}_{k+1}, w_k$ so that for every term of $\bar{f}_{X_k|Y_k}$, the integral in (3.44) becomes

$$\tilde{f}^i_{X_{k+1},W|Y_k}(\boldsymbol{x}_{k+1}, w_k|\boldsymbol{y}_k) = \int_{-\infty}^{\infty} \tilde{G}_i^{k|k}(\boldsymbol{x}_{k+1}, w_k)\tilde{\epsilon}_i^{k|k}(\boldsymbol{x}_{k+1}, w_k)\, dw \tag{3.45}$$

where

$$\begin{aligned}
\tilde{G}_i^{k|k}(\boldsymbol{x}_{k+1}, w_k) &= \tilde{\rho}_{i0}^{k|k} + \sum_{j=1}^{m} \tilde{\rho}_{ij}^{k|k} \prod_{l\in\sigma_j} \operatorname{sgn}\left(\tilde{\xi}_{il}^{k|k}(\boldsymbol{x}_{k+1}) - w_k\right) \\
\tilde{\epsilon}(\boldsymbol{x}_{k+1,w_k}) &= \exp\left(-\sum_{j=1}^{M_i^{k|k-1}+2} \tilde{\eta}_{ij}^{k|k} \left|\tilde{\xi}_{ij}^{k|k}(\boldsymbol{x}_{k+1}) - w_k\right|\right) \\
\tilde{\rho}_{i0}^{k|k} &= \frac{|\Phi|^{-1}}{2\beta}\rho_{i0}^{k|k} \\
\tilde{\rho}_{ij}^{k|k} &= \frac{|\Phi|^{-1}}{2\beta}\rho_{ij}^{k|k} \prod_{l\in\sigma_j} \operatorname{sgn}\left(\theta_{il}^{k|k}\Phi^{-1}\Gamma\right) \\
\tilde{\eta}_{ij}^{k|k} &= \eta_{ij}^{k|k} \cdot \frac{1}{\left|\theta_{ij}^{k|k}\Phi^{-1}\Gamma\right|} \\
\tilde{\xi}_{ij}^{k|k} &= \frac{\psi_{ij}^{k|k}}{\theta_{ij}^{T}\Phi^{-1}\Gamma} + \frac{\theta_{ij}^{k|k}\Phi^{-1}}{\theta_{ij}^{k|k}\Phi^{-1}\Gamma}\boldsymbol{x}_{k+1} \\
&\triangleq \tilde{\psi}_{ij}^{k|k} + \tilde{\theta}_{ij}^{k|k}\boldsymbol{x}_{k+1}
\end{aligned} \tag{3.46}$$

Each term is then integrated using the generalized integral formula (3.32) to get

$$\begin{aligned}
\tilde{f}_{i,X_{k+1}|Y_k}(\boldsymbol{x}_{k+1}|\boldsymbol{y}_k) &= \sum_{j=1}^{m} \tilde{G}_{ijl}^{k+1|k} \exp\left(-\sum_{l=1}^{m} \tilde{\eta}_{il}^{k|k} \left|\tilde{\xi}_{il}^{k|k} - \tilde{\xi}_{ij}^{k|k}\right|\right) \\
&= \sum_{j=1}^{m} \tilde{G}_{ijl}^{k+1|k} \exp\left(-\sum_{l=1}^{m} \tilde{\eta}_{il}^{k|k} \left|\left(\tilde{\psi}_{ij}^{k|k} - \tilde{\psi}_{il}^{k|k}\right) + \left(\tilde{\theta}_{ij}^{k|k} - \tilde{\theta}_{il}^{k|k}\right)\boldsymbol{x}_{k+1}\right|\right) \\
&\triangleq \sum_{j=1}^{m} \tilde{G}_{ijl}^{k+1|k} \exp\left(-\sum_{l=1}^{m} \tilde{\eta}_{ijl}^{k+1|k} \left|\tilde{\psi}_{ijl}^{k+1|k} + \tilde{\theta}_{ijl}^{k|k}\boldsymbol{x}_{k+1}\right|\right) \\
&= \sum_{j=1}^{m} \tilde{G}_{ijl}^{k+1|k} \exp\left(-\sum_{l=1}^{m} \tilde{\eta}_{ijl}^{k+1|k} \left|\tilde{\xi}_{ijl}^{k+1|k}\right|\right)
\end{aligned} \tag{3.47}$$

where

$$\tilde{G}_{ijl}^{k+1|k} = \left[\frac{\tilde{G}_i^{k|k}}{\sum_{l=1}^{m} \tilde{\eta}_{ijl}^{k+1|k} \operatorname{sgn}\left(\tilde{\xi}_{ijl}^{k+1|k}\right)}\right]_{s_j^j=1} - \frac{\tilde{G}_i^{k|k}}{\sum_{l=1}^{m} \tilde{\eta}_{ijl}^{k+1|k} \operatorname{sgn}\left(\tilde{\xi}_{ijl}^{k+1|k}\right)} \tag{3.48}$$

and $s_j^j = 1$ indicates that $\operatorname{sgn}(0) = 0$ instead of the normal convention $\operatorname{sgn}(0) = -1$ in (1.3). Note that each new argument has the form

$$\begin{aligned}
\tilde{\eta}_{ijl}^{k+1|k} &= \tilde{\eta}_{ij}^{k|k} \\
\tilde{\xi}_{ijl}^{k+1|k} &= \tilde{\xi}_{ij}^{k|k} - \tilde{\xi}_{il}^{k|k} \\
&= \left(\tilde{\psi}_{ij}^{k|k} - \tilde{\psi}_{il}^{k|k}\right) + \left(\tilde{\theta}_{ij}^{k|k} - \tilde{\theta}_{il}^{k|k}\right)\boldsymbol{x}_{k+1} \\
&\triangleq \tilde{\psi}_{ijl}^{k+1|k} + \tilde{\theta}_{ijl}^{k|k}\boldsymbol{x}_{k+1}
\end{aligned} \tag{3.49}$$

3.4.5 Simplify coefficient function

The coefficient term in (3.48) is then flattened into the standard form of (3.35) using the procedures outlined in Section 7.6 of the Appendix so that

$$\tilde{G}_{ijl}^{k+1|k} = \tilde{\rho}_{ij0}^{k+1|k} + \sum_{p=1}^{m} \tilde{\rho}_{ijp}^{k+1|k} \prod_{l\in\sigma_{ijp}} \operatorname{sgn}\left(\tilde{\xi}_{ijl}^{k+1|k}\right) \tag{3.50}$$

This flattening accomplished two things. Firstly, it simplifies a complicated nested fractional form of G and makes it easier to carry forward from step to step. Secondly, it allows for terms with identical exponentials to sum, thus eliminating extra terms and reducing future computational burden.

3.4.6 Combine terms and re-index at $k+1|k$

After the coefficient terms have been flattened, the triple index ijl can then be reduced to two new indices, ij. Furthermore, terms with the same exponentials can be combined, though care must be taken to determine if all of the associated $\tilde{\xi}_{ijl}^{k+1|k}$ are the same. Finally, the re-indexed ucpdf can be written without the tilde on the parameters for step $k+1|k$.

3.4.7 Handle collapsed hyperplanes

In many cases, one or more hyperplanes collapse into a constant. That is, for some hyperplane $\xi_{ij}^* = \psi_{ij}^* + \theta_{ij}^* \boldsymbol{x}$, $\theta_{ij}^* = 0$. During the next propagation, the coefficients of ξ_{ij}^* will divide $\theta_{ij}^* \Phi^{-1} \Gamma$, which will not work. It is easier to just look for these cases and factor out the constant $\exp\left(-\rho_{ij}^* \left|\psi_{ij}^*\right|\right)$ from the exponential at this point, rather than modify the propagation to do the same thing.

The coefficient term is a bit more complicated, since all sub-terms with constant ξ_{ij}^* must be found and dealt with, including possibly shifting. Finally, all element indices must be shifted when the element associated with ξ_{ij}^* is removed. An example of this process is shown in Section 7.9 of the Appendix.

3.4.8 Unnormalized cpdf at $k+1|k$

Finally, we state the ucpdf at $k+1|k$ as

$$\bar{f}_{X_{k+1}|Y_k} = \sum_{i=1}^{N_i^{k+1|k}} G_i^{k+1|k} \epsilon_i^{k+1|k} \tag{3.51}$$

where

$$G_i^{k+1|k} = \rho_{i0}^{k+1|k} + \sum_{j=1}^{P_i^{k+1|k}} \rho_{ij}^{k+1|k} \prod_{l\in\sigma_{ij}} \operatorname{sgn}\left(\xi_{il}^{k+1|k}\right) \tag{3.52}$$

and

$$\epsilon_i^{k+1|k} = \exp\left(-\sum_{j=1}^{M_i^{k+1|k}} \eta_{ij}^{k+1|k} \left|\xi_{ij}^{k+1|k}\right|\right) \tag{3.53}$$

where

$$\xi_{ij}^{k+1|k} = \psi_{ij}^{k+1|k} + \theta_{ij}^{k+1|k} \boldsymbol{x}_{k+1}, \tag{3.54}$$

$N_i^{k+1|k}$ is the number of terms, $M_i^{k+1|k}$ is the number of elements of term i, and $P_i^{k+1|k}$ is the number of terms in $G_i^{k+1|k}$, and σ_{ij} is the set of indices of sign functions associated with term j of $G_i^{k+1|k}$ at step $k+1|k$. Thus, we've shown by induction that the update and propagation algorithm is recursive for $\bar{f}_{X_k|Y_{k-1}}$ in (3.34).

3.5 Example of first several *a posteriori* ucpdfs

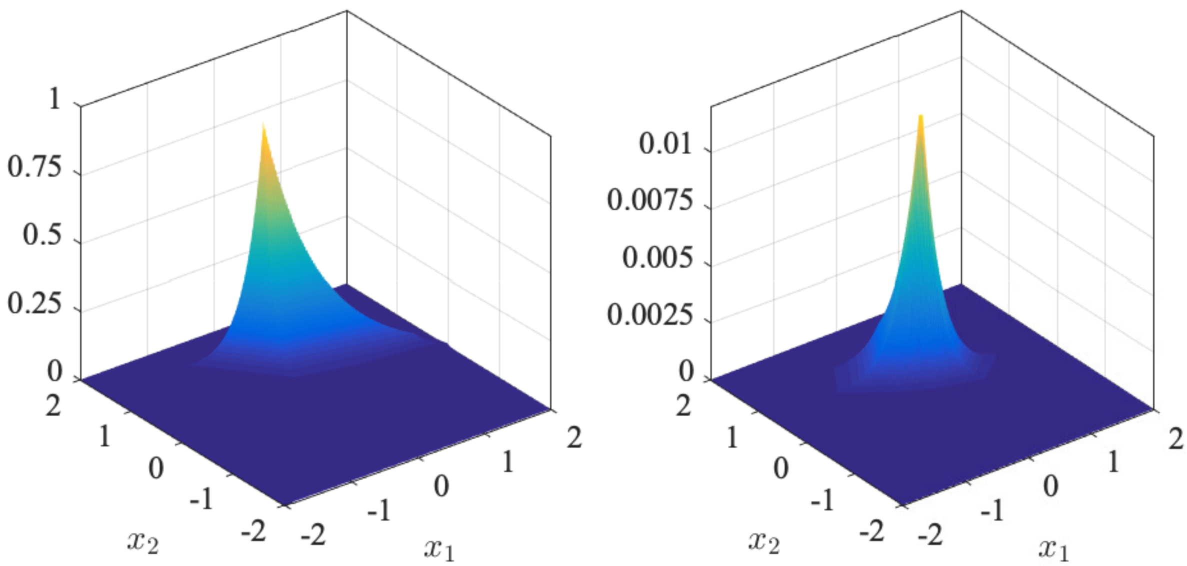
Using the algorithm described Section 3.4, the first several *a posteriori* ucpdfs are computed for system parameters

$$\Phi = \begin{bmatrix} 0.95 & 0.01 \\ -0.1 & 1 \end{bmatrix} \qquad \Gamma = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0.5 \end{bmatrix} \tag{3.55}$$

with initial conditions

$$\bar{\boldsymbol{x}}_1 = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} \tag{3.56}$$

and noise parameters $\alpha = 0.3$, $\beta = 0.01$ and $\gamma = 0.1$. The ucpdfs are presented in Figures 3.3 to 3.5, showing the various shapes that the *a posteriori* ucpdf takes on from step to step.



(a) (b)

Figure 3.3: $\bar{f}_{X_1|Y_1}$ (a) and $\bar{f}_{X_2|Y_2}$ (b)

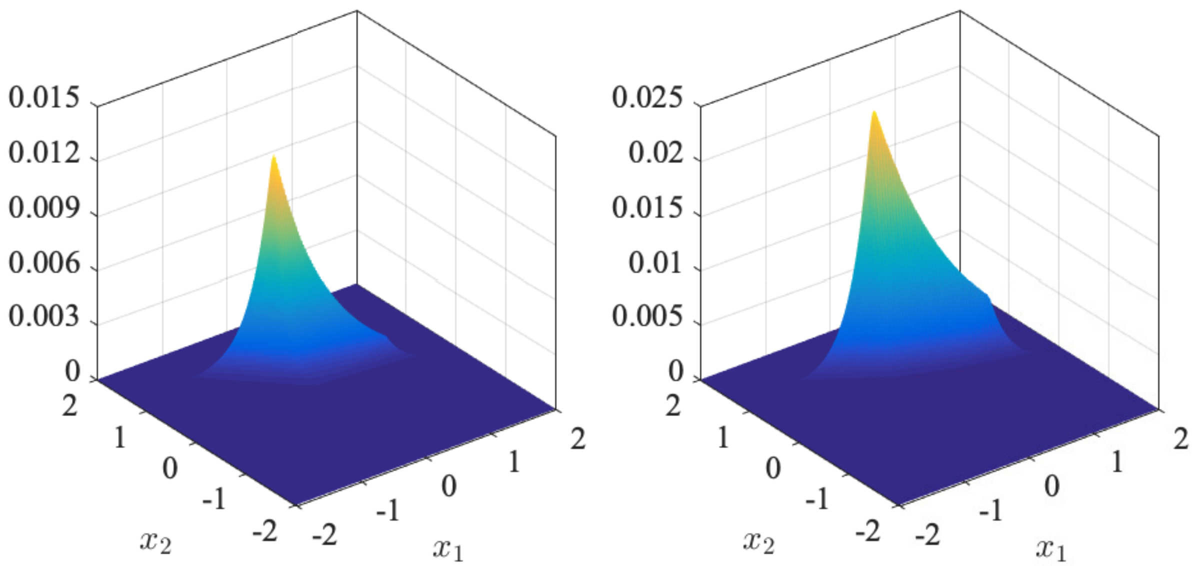
3.6 Properties of the ucpdf

We now discuss some properties of the ucpdf in (3.34). In particular, the computational complexity of the ucpdf in $\mathbb{R}^2$ and $\mathbb{R}^n$ will be characterized.

3.6.1 Computational complexity in $\mathbb{R}^2$

In contrast to the scalar case, the number of elements in each term stays relatively the same, with the exception of some elements collapsing due to their associated hyperplane becoming constant ($\theta \to 0$). For the *a priori* step $k+1|k$, the number of elements is $d+k$, where d accounts for the initial conditions.

Without term combination or element removal, the number of terms for the first few steps are shown in Table 3.1. The number of new terms is the product of current terms and the number of elements per term. Since there is no reduction, all terms have the same



(a) (b) $\bar{f}_{X_2|Y_1}$

Figure 3.4: $\bar{f}_{X_3|Y_3}$ (a) and $\bar{f}_{X_4|Y_4}$ (b)

maximal number of elements.

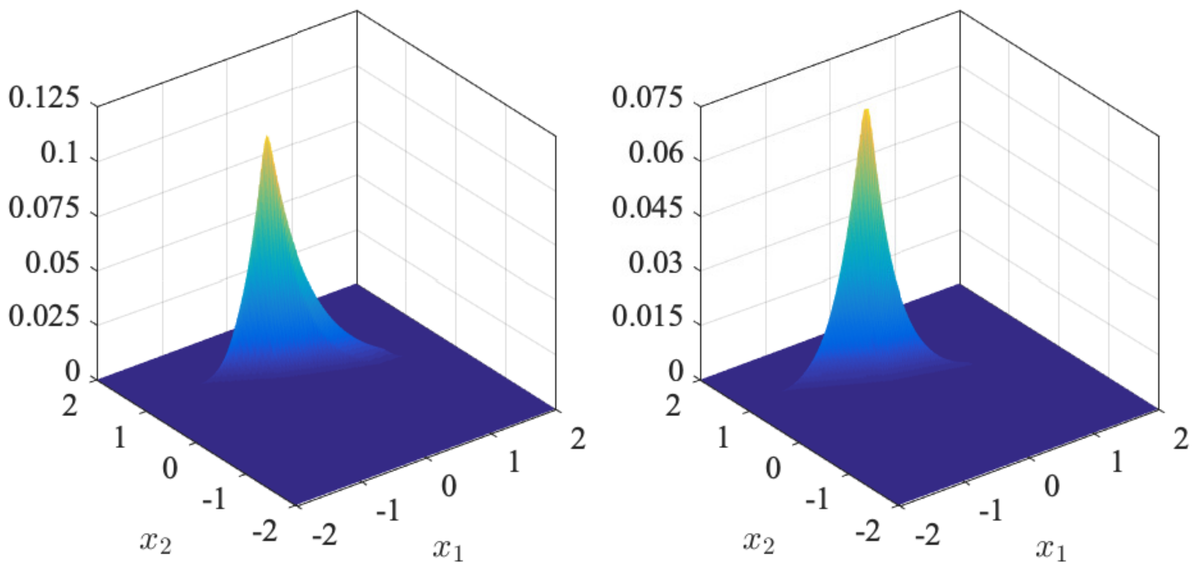
It can be shown that the theoretical worst-case number of terms at $k+1|k$ for the two-dimensional case without term combination or element removal is

$$\max n^{k+1|k} = \frac{(2+k)!}{3!} \tag{3.57}$$

In practice, combining terms and removing hyperplanes which collapse into constants drive down the rate considerably. However, it does appear that rate of increase remains faster than polynomial. Figure 3.6 shows the number of terms for the first six steps. The number of terms are equal for the first three steps, but they start to diverge at step 4.

3.6.2 Extension to $\mathbb{R}^n$

First, notice that the only time we explicitly use the fact that $\boldsymbol{x} \in \mathbb{R}^2$ is in the initial conditions and in simplifying the coefficient function. However, Theorem 3.3.1 applies to



(a) (b) $\bar{f}_{X_2|Y_1}$

Figure 3.5: $\bar{f}_{X_5|Y_5}$ (a) and $\bar{f}_{X_6|Y_6}$ (b)

n-dimensions. Therefore, all we need to do is start with the appropriate f_{X_1} and use the correct coefficient simplifying algorithm and the preceding results extend to n-dimensions as well. However, the term count in (3.57) will change to

$$\max n^{k+1|k} = \frac{(n+k)!}{(d+1)!} \tag{3.58}$$

3.6.3 Unimodality

Similarly to the scalar case, we can make an argument that the vector ucpdf is log-concave and unimodal. This becomes an important point as the number of dimensions increases, because as shown in (3.58), the number of terms increases rapidly. The unimodal nature of the ucpdf offers an opportunity to use the maximum *a posteriori* (MAP) estimate instead of the conditional mean. Since it's unimodal, it has a single local maximum which is also the global maximum, and many numerical methods may be leveraged to find it.

k	Terms	Elements
1	1	4
2	4	5
3	20	6
4	120	7
5	840	8
6	6720	9

Table 3.1: Theoretical worst-case term count for $\mathbb{R}^2$

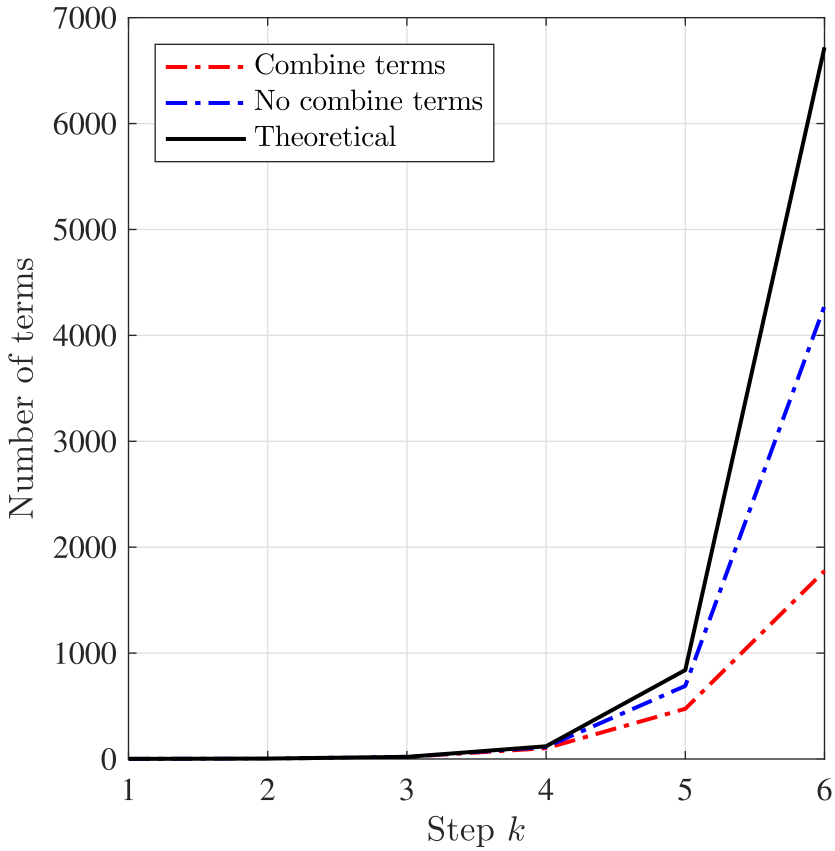


Figure 3.6: Theoretical worst case vs. actual number of terms in $\mathbb{R}^2$

CHAPTER 4

Mean and Variance in $\mathbb{R}^2$

To determine the mean and variance of X given Y, we first normalize $\bar{f}_{X|Y}$ and then compute the first and second moments. For this chapter, $\bar{f}_{X|Y}$ can be either *a priori* or *a posteriori* at any step k, so we will not need to include indices. As in the scalar case, we use the characteristic function of the ucpdf. For $\boldsymbol{\nu} \in \mathbb{C}^n, \boldsymbol{x} \in \mathbb{R}^n$, the characteristic function is

$$\bar{\phi}_{X|Y}(\boldsymbol{\nu}) = \int_{\mathbb{R}^n} e^{j\boldsymbol{\nu}^T\boldsymbol{x}} \bar{f}_{X|Y}(\boldsymbol{x}|\boldsymbol{y})\, d\boldsymbol{x} \tag{4.1}$$

and the mean and variance are given by

$$\begin{aligned} \mu &= E[X|Y] \\ \text{Var}(x) &= E\left[XX^T|Y\right] - E[X|Y]\, E[X|Y]^T \end{aligned} \tag{4.2}$$

where the i-th element of $E[X|Y]$ and the $i\ell$ element of the symmetric $E\left[XX^T|Y\right]$ are

$$\begin{aligned} E[X_i|Y] &= \frac{1}{j}\frac{\partial \phi_{X|Y}(\boldsymbol{\nu})}{\partial \nu_i} \\ E[X_i X_\ell|Y] &= -\frac{\partial^2 \phi_{X|Y}(\boldsymbol{\nu})}{\partial \nu_i \partial \nu_\ell}, \end{aligned} \tag{4.3}$$

respectively. The proof for these expressions follow the same argument as the scalar case in Section 7.4 of the Appendix. Note that we have abused the notation a little bit. Since we are not labeling the step number k, we use the subscript on the random variable X to indicate the element number instead. This should be clear in this context, even if it doesn't agree with the standard use of the subscript throughout this document.

For $\boldsymbol{x} \in \mathbb{R}^n$, this requires n integrations as well as n single and $\sum_i^n i = \frac{n(n+1)}{2}$ double partial differentiations, though the latter can be done *a priori*. As bad as that sounds, it appears to be less complicated than than directly integrating to find the moments.

4.1 Characteristic function of ucpdf

To determine the characteristic function of $f_{X|Y}$ in $\mathbb{R}^n$, we have to integrate over $\mathbb{R}^n$, which means evaluating n successive integrals using the same integral formula that we've used in previous chapters. Consider the form $f_{X|Y}$ to be

$$f_{X|Y} = \sum_{i=1}^{N} g_i(x) \exp\left(-\sum_{l=1}^{m} \rho_l \left|\psi_l + \theta_l \boldsymbol{x}\right|\right) \tag{4.4}$$

where, for m unique $\psi + \theta\boldsymbol{x}$, $\theta \in \mathbb{R}^{1\times n}$,

$$g_i(x) = \varrho_0 + \sum_{q=1}^{M_i} \varrho_q \prod_{l\in\sigma_q} \operatorname{sgn}\left(\psi_l + \theta_l \boldsymbol{x}\right) \tag{4.5}$$

and $\sigma_q \subseteq \{1, \cdots, m\}$ are unique. Note that when $m > n$, the maximum number of elements of σ_k is n. Using the generalized integral formula (3.32), (4.1) can be solved element-wise, and each element can be explicitly written as

$$\begin{aligned}
\phi^i_{X|Y} &= \int_x e^{j\nu^T x} g_i(x) \exp\left(-\sum_{l=1}^{m} \rho_l \left|\psi_l + \theta_l \boldsymbol{x}\right|\right) dx \\
&= \int_{x_1}\int_{x_2}\cdots\int_{x_n} \left(\rho_0 + \sum_{q=1}^{M_i} \rho_q \prod_{l\in\sigma_q} \operatorname{sgn}\left(\psi_l + \theta_l \boldsymbol{x}\right)\right) \\
&\qquad \exp\left(-\sum_{l=1}^{m} \rho_l \left|\psi_l + \theta_l \boldsymbol{x}\right| + j\nu^T x\right) dx_1 dx_2 \cdots dx_n \\
&= \int_{x_1}\int_{x_2}\cdots\int_{x_n} \left(\rho_0 + \sum_{q=1}^{M_i} \rho_q \prod_{l\in\sigma_q} \operatorname{sgn}\left(\psi_l + \theta_1 x_1 + \cdots + \theta_n x_n\right)\right) \\
&\qquad \exp\left(-\sum_{l=1}^{m} \rho_l \left|\psi_l + \theta_{l1} x_1 + \cdots + \theta_{ln} x_n\right| + j\left(\nu_1 x_1 + \cdots + \nu_n x_n\right)\right) dx_1 dx_2 \cdots dx_n
\end{aligned} \tag{4.6}$$

To illustrate this integration, the characteristic function $\bar{\phi}_{X|Y}$ for $\bar{f}_{X|Y}$ in $\mathbb{R}^2$ is derived in closed form in Section 7.11 of the Appendix. In addition, an example of the characteristic function of $\bar{f}_{X_1|Y_1}$ is detailed in Section 7.12 of the Appendix.

4.2 Normalization and moments from CF in $\mathbb{R}^2$

Evaluating $\bar{\phi}_{X|Y}$ at $\boldsymbol{\nu} = \mathbf{0}$ gives

$$\left[\bar{\phi}_{X|Y}\right]_{\boldsymbol{\nu}=\mathbf{0}} = \int_{\mathbb{R}^n} \bar{f}_{X|Y}(\boldsymbol{x}|\boldsymbol{y})\, d\boldsymbol{x} = f_Y \tag{4.7}$$

Then, the normalized cpdf and characteristic functions are

$$\begin{aligned} f_{X|Y} &= \frac{\bar{f}_{X|Y}}{f_Y} \\ \phi_{X|Y} &= \frac{\bar{\phi}_{X|Y}}{f_Y} \end{aligned} \tag{4.8}$$

Applying (4.7) to (7.108), we can evaluate each term of the normalization factor as

$$f_Y^i = \left[\phi_{X|Y}^i\right]_{\boldsymbol{\nu}=\mathbf{0}} = \left(\frac{\frac{a_{i1}}{a_{i2}} - \frac{a_{i4}}{a_{i5}}}{a_{i7}} - \frac{\frac{b_{i1}}{b_{i2}} - \frac{b_{i4}}{b_{i5}}}{b_{i7}} \right) \exp\left(c_{i3}\right) \tag{4.9}$$

Before defining the first and second moments, we simplify the definition of G_i from (7.108) and ϵ_i as

$$\begin{aligned} G_i &\triangleq \frac{N_a}{D_a} - \frac{N_b}{D_b} \\ \epsilon &\triangleq \exp\left(j\nu_1 c_1 + j\nu_2 c_2 + c_3\right) \end{aligned} \tag{4.10}$$

We drop the indices on ϕ, G and ϵ for cleaner notation, but it is assumed that they are components of $\phi_{X|Y}^i$.

The first two moments in $\mathbb{R}^2$ are

$$\begin{aligned} E\left[X\right] &= \begin{bmatrix} \frac{1}{j}\frac{\partial \phi}{\partial \nu_1} \\ \frac{1}{j}\frac{\partial \phi}{\partial \nu_2} \end{bmatrix}_{\boldsymbol{\nu}=\mathbf{0}} \\ E\left[XX^T\right] &= \begin{bmatrix} -\frac{\partial^2 \phi}{\partial \nu_1^2} & -\frac{\partial^2 \phi}{\partial \nu_1 \partial \nu_2} \\ -\frac{\partial^2 \phi}{\partial \nu_1 \partial \nu_2} & -\frac{\partial^2 \phi}{\partial \nu_2^2} \end{bmatrix}_{\boldsymbol{\nu}=\mathbf{0}} \end{aligned} \tag{4.11}$$

For the sake of brevity, the derivation of the partial derivatives for a general term of the characteristic function are detailed in Section 7.13 of the Appendix.

4.3 Example: 2-state conditional pdfs

Figure 4.1 shows the results for a 7-step simulation where the system parameters are given in the previous section in (3.55). Similarly to the scalar case, we can see that the covariance changes as a function fo the measurement, rather than steadily decrease as in the Kalman filter. In addition, the response to the large jump in the measurement at $k = 2$ is rather muted, despite the fact that $z_k = x_1 + 0.5x_2$.

Figure 4.2 shows the results for the same system but with a spike of -10 introduced at $k = 3$. As in the scalar case, the estimator only responds to a certain extent. Due to some numerical limitations discussed later, we were unable to run spikes on longer runs. While some work is still needed to improve the duration and speed of the numerical implementations, these examples demonstrate that the estimator works in two-dimensions and shows resilience to huge noise events.

4.3.1 External algorithms

In practice, we do not use the algorithm in Section 7.6 to simplify the coefficient functions. While it is fast and effective for benign cases, it is slow for large numbers of hyperplanes or misses faces for hyperplane arrangements with small or thin faces. We first experimented with the reverse search algorithm by Avis and Fukada in [AF93], but that had some numerical issues of its own and was particularly slow, even compared to our simple algorithm in Section 7.6. Currently, we use the algorithms developed by Rada and Cerný in [RC18], along with linear program solvers by Gurobi to produce fast, accurate sign vectors.

4.3.2 Numerical sensitivity

The reason why we can only run 7 steps for the 2-dimensional case is because we run into numerical errors around there at the noise parameters we use. When the noise is large, the huge jumps cause the conditional pdf to flatten and the scale of certain coefficients defining it to blow up. As a result, we suspect that the machine error after term cancellation for certain terms results in large errors in the normalization factor and expectations.

This is not a comprehensive assessment of the problem, but it offer evidence that the current definition of the cpdf will not work unless we find a way to effectively remove the need for this cancellation.

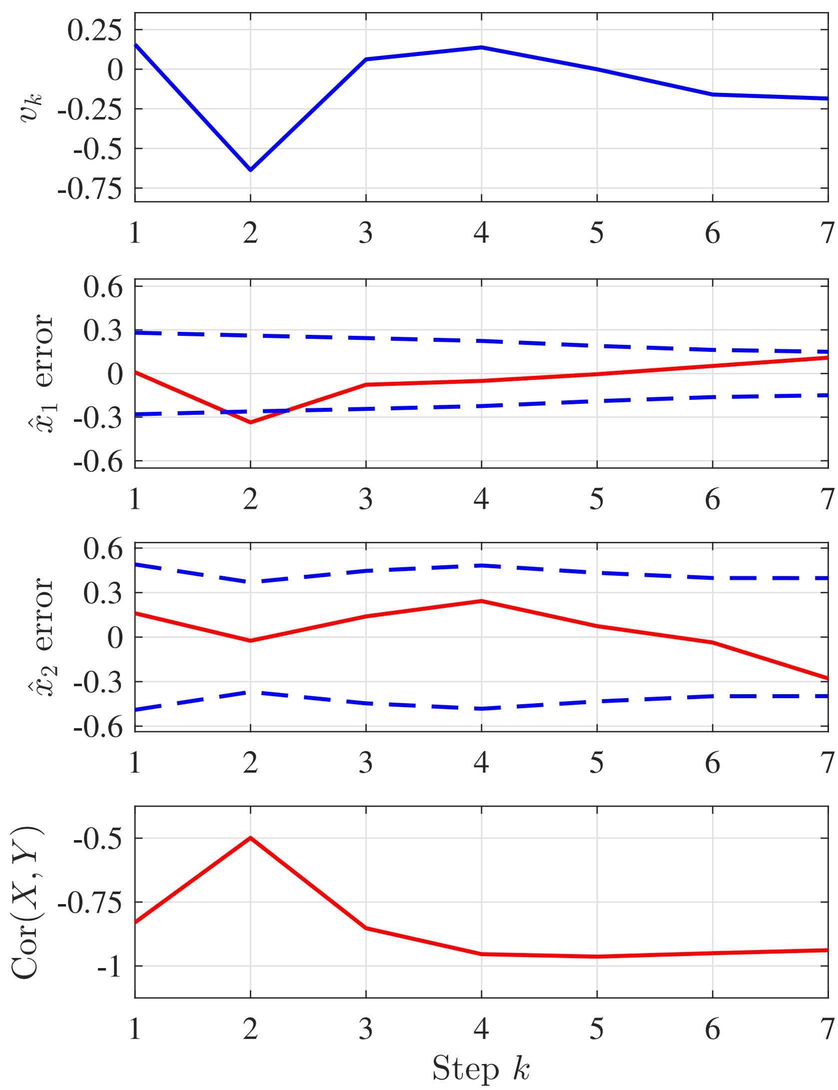


Figure 4.1: Estimation errors for 7 steps in $\mathbb{R}^2$ for $\alpha = 0.3, \beta = 0.01, \gamma = 0.1$

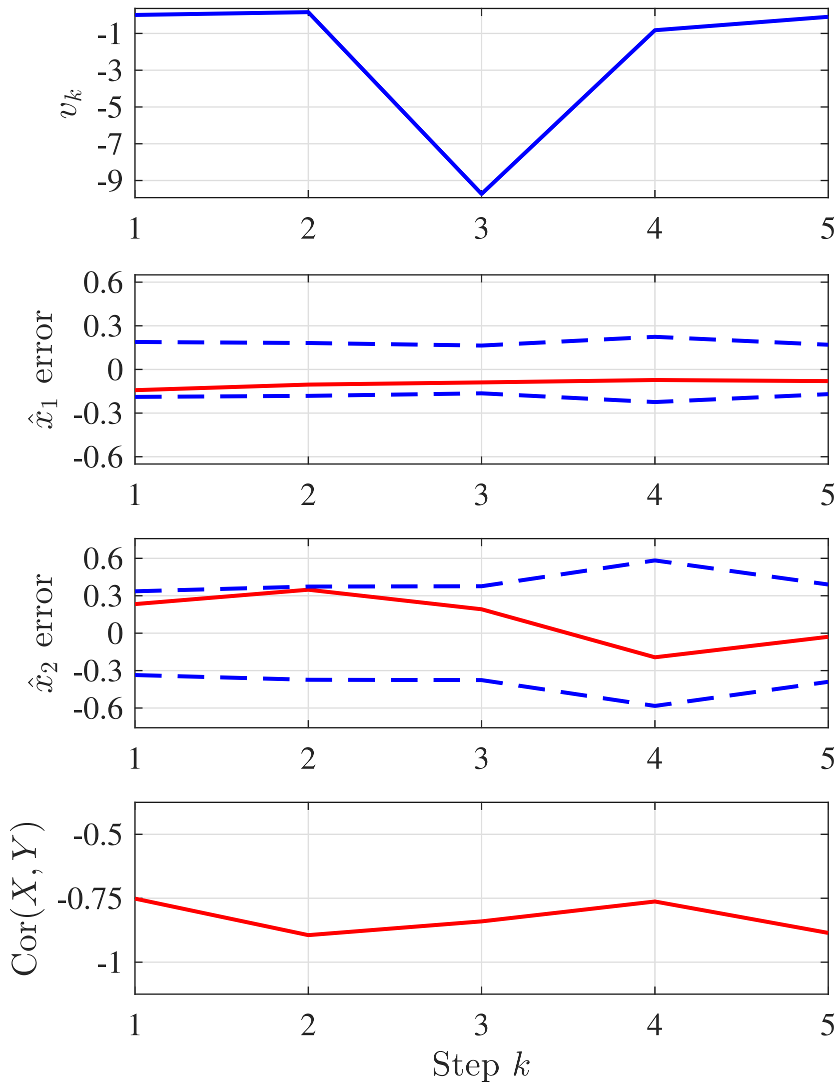


Figure 4.2: Estimation errors for 5 steps with spike at $k = 3$ in $\mathbb{R}^2$ for $\alpha = 0.3, \beta = 0.01, \gamma = 0.1$

CHAPTER 5

Scalar Laplace Controller

In many optimization problems, the L_1-norm is preferable to the typical L_2-norm for the cost function because it more directly measures real quantities, such as fuel consumption [VM06],[Ros06] and cost of treatment of infectious diseases [PIM14] and cancer [LBS04]. For the control problem near steady-state, where the control input is small, the cost is higher for the L_1-norm than it is for the L_2-norm and can be used to suppress control input. Conversely, the L_1-norm has a lower penalty for large errors and is used for applications with many outliers [MKP14]. However, there are no explicit solutions to linear stochastic optimal control problems with L_1 cost functions. In fact, only a few explicit solutions exist for any linear stochastic optimal control problem, *e.g.* the Linear Quadratic Gaussian and Linear Exponential Gaussian [SC08].

In Chapter 2, it is shown that, for linear scalar systems with additive noises described by Laplace densities, the conditional probability density functions of the state conditioned on the measurements is a sum of exponentials of weighted sums of absolute values. This form is log-concave and is therefore unimodal [Sta89], though it is not symmetric. In [SYD18], it is shown that, when addressing the control problem of such a system, a properly-defined cost function of the L_1-norm of the control and state maintains that same log-concave form for the first step. This suggests that the cost function is also unimodal, which greatly simplifies the extremization of the cost. In this chapter, we generalize the results in [SYD18] and present the recursive, analytic solution to the one-step predictive controller for any step k.

In Section 5.1, we define the cost criterion for the one-step predictive controller. In

Section 5.2, we derive the expression of the cost function at step $k = 1$ in a form which motivates a recursive structure. We also examine the optimal control values as a function of the measurement. In Section 5.3, we deduce the cost function at any step k given the conditional probability density function from (2.43). In Section 5.4, we show numerical results for a 50-step simulation.

5.1 One-step model-predictive controller

Consider the discrete scalar linear system defined in Chapter 2. We seek to derive a controller which drives the state x_k to zero while minimizing some cost as a function of the input u_k. In constructing the cost criterion, we consider a measure of the control input and predictive state. We want the controller to drive the state to zero while using minimal control input, so we chose a cost function whose maximum occurs when u_k and x_{k+1} are small. For simplicity, our cost criterion only projects the state one step into the future so that the cost criterion only involves u_k and x_{k+1}. Therefore, we want to find u_k^* such that

$$u_k^* = \underset{u_k}{\operatorname{argmax}}\, J_{Y_k}(u_k) \tag{5.1}$$

where

$$\begin{aligned} J_{Y_k}(u_k) &= E\left[e^{-c|u_k| - S|x_{k+1} + \bar{x}_{k+1}|} | Y_k\right] \\ &= E\left[e^{-c|u_k| - S|x_{k+1} + \Phi \bar{x}_k + u_k|} | Y_k\right], \end{aligned} \tag{5.2}$$

and c and S are weighting parameters on the control and future state, respectively. The expectation in (5.2) is taken with respect to the unnormalized conditional pdf (ucpdf) $\bar{f}_{X_{k+1}|Y_k}(x_{k+1}|y_k)$, which is conditioned on the measurement history y_k. We can use the ucpdf because normalization is not necessary for the optimization. For simplicity, we will assume $\bar{x}_1 = 0$.

5.2 Objective function evaluated at $k = 1$

To motivate a recursive structure of the objective function at step k, we first determine the objective function at $k = 1$. This is similar to the work in [SYD18], but it uses a simplified structure for the ucpdf first shown in [DSY18]. To streamline the presentation, we have moved the more involved derivation to the Appendices.

We first consider the *a priori* ucpdf at step $k = 2$ conditioned on measurements up to $k = 1$. Assuming $\bar{x}_1 = 0$, this was defined in [DSY18] to be

$$\begin{aligned}
&\bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1) \\
&\quad = G_1^{2|1} \cdot \exp\left(-\frac{1}{\beta}\left|\frac{\Phi z_1}{H} - x_2\right|\right) \\
&\quad + G_2^{2|1} \cdot \exp\left(-\frac{1}{\beta}\left|x_2\right|\right) \\
&\quad + G_3^{2|1} \cdot \exp\left(-\frac{|H|}{\gamma\Phi}\left|\frac{\Phi z_1}{H} - x_2\right| - \frac{1}{\alpha\left|\Phi\right|}\left|x_2\right|\right),
\end{aligned} \tag{5.3}$$

where $G_i^{2|1}$ are coefficient functions whose form will be described below. First, define the parameters of (5.3) as

$$\begin{aligned}
\rho_{11}^{2|1} = \tfrac{1}{\beta}, \quad \rho_{21}^{2|1} = \tfrac{1}{\beta}, \quad \rho_{31}^{2|1} &= \tfrac{|H|}{\gamma|\Phi|} \\
\rho_{32}^{2|1} &= \tfrac{\alpha}{|\Phi|}
\end{aligned} \tag{5.4}$$

and

$$\begin{aligned}
\xi_{11}^{2|1} = \tfrac{\Phi z_1}{H}, \quad \xi_{21}^{2|1} = 0, \quad \xi_{31}^{2|1} &= \tfrac{\Phi z_1}{H} \\
\xi_{32}^{2|1} &= 0.
\end{aligned} \tag{5.5}$$

Then, the coefficient and exponential functions in (5.3) can then be written as in terms of the parameters in (5.4) and (5.5) as

$$G_i^{2|1} = a_{i0}^{2|1} + \sum_{l=1}^{N_i^{2|1}} a_{il}^{2|1} \operatorname{sgn}\left(\xi_{il}^{2|1} - x_2\right) \tag{5.6}$$

and

$$\epsilon_i^{2|1} = \exp\left(\sum_{l=1}^{N_i^{2|1}} -\rho_{il}^{2|1}\left|\xi_{il}^{2|1} - x_2\right|\right), \tag{5.7}$$

where $N_i^{2|1}$ is the ith element of

$$N^{2|1} = \begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^T \tag{5.8}$$

and the parameters $a_{i0}^{2|1}$ and $a_{il}^{2|1}$ are developed in Appendix B. The ucpdf in (5.3) can then be written in the compact form

$$\bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1) = \sum_{i=1}^{3} G_i^{2|1} \epsilon_i^{2|1} \tag{5.9}$$

5.2.1 Determining the objective function

Using (5.9), the objective function in (5.2) becomes

$$\begin{aligned} &J_{Y_1}(u_1) \\ &= e^{-c|u_1|} \int_{-\infty}^{\infty} e^{-S|x_2+\Phi\bar{x}_1+u_1|} \bar{f}_{X_2|Y_1}(x_2|\boldsymbol{y}_1)\, dx_2 \\ &= e^{-c|u_1|} \sum_{i=1}^{3} \int_{-\infty}^{\infty} G_i^{2|1} \epsilon_i^{2|1} e^{-S|x_2+\Phi\bar{x}_1+u_1|}\, dx_2 \end{aligned} \tag{5.10}$$

The general solution to (5.10) is derived in Appendix B of [IS14] and is restated in the Appendix for convenience. In order to use the integral formula (7.3), we augment the parameters in (5.10) so that the number of elements in $G_i^{2|1}$ matches the number of elements in $\epsilon_i^{2|1}$. Here, we make a subtle change in notation from the superscript $k+1|k = 2|1$, which refers to the ucpdf, to the superscript $k = 1$. This is done to simplify the task of organizing the parameters to use the integration formula in Appendix A. So, the augmented parameters for the objective function integral of (5.10) are

$$\begin{array}{lll}
\varrho_{10}^{1} = 0 & \xi_{10}^{1} = -\Phi\bar{x}_1 - u_1 & \rho_{10}^{1} = S \\
\varrho_{11}^{1} = a_{11}^{2|1} & \xi_{11}^{1} = \xi_{11}^{2|1} & \rho_{11}^{1} = \rho_{11}^{2|1} \\
\hline
\varrho_{20}^{1} = 0 & \xi_{20}^{1} = -\Phi\bar{x}_1 - u_1 & \rho_{20}^{1} = S \\
\varrho_{21}^{1} = a_{21}^{2|1} & \xi_{21}^{1} = \xi_{21}^{2|1} & \rho_{21}^{1} = \rho_{21}^{2|1} \\
\hline
\varrho_{30}^{1} = 0 & \xi_{30}^{1} = -\Phi\bar{x}_1 - u_1 & \rho_{30}^{1} = S \\
\varrho_{31}^{1} = a_{31}^{2|1} & \xi_{31}^{1} = \xi_{31}^{2|1} & \rho_{31}^{1} = \rho_{31}^{2|1} \\
\varrho_{32}^{1} = a_{32}^{2|1} & \xi_{32}^{1} = \xi_{32}^{2|1} & \rho_{32}^{1} = \rho_{32}^{2|1}
\end{array} \tag{5.11}$$

Using (5.11), the solution to (5.10) becomes

$$J_{Y_1}(u_1) = e^{-c|u_1|} \sum_{i=1}^{3} \sum_{j=0}^{N_i^{2|1}} G_{ij}^{1} \epsilon_{ij}^{1} \tag{5.12}$$

where, for

$$\begin{aligned}
\bar{\delta}_{ij}^{1} &= \sum_{\substack{l=0 \\ l \neq j}}^{N_i^{2|1}} \varrho_{il}^{1} \operatorname{sgn}\left(\xi_{il}^{1} - \xi_{ij}^{1}\right) \\
\delta_{ij}^{1} &= \sum_{\substack{l=0 \\ l \neq j}}^{N_i^{2|1}} \rho_{il}^{1} \operatorname{sgn}\left(\xi_{il}^{1} - \xi_{ij}^{1}\right),
\end{aligned} \tag{5.13}$$

the coefficient and exponential terms are

$$G_{ij}^{1} = \frac{a_{i0}^{2|1} + \bar{\delta}_{ij}^{1}}{\rho_{ij}^{1} + \delta_{ij}^{1}} - \frac{-a_{i0}^{2|1} + \bar{\delta}_{ij}^{1}}{-\rho_{ij}^{1} + \delta_{ij}^{1}} \tag{5.14}$$

and

$$\epsilon_{ij}^{1} = \exp\left(\sum_{\substack{l=0 \\ l \neq j}}^{N_i^{2|1}} \rho_{jl}^{1} \left| \xi_{il}^{1} - \xi_{ij}^{1} \right| \right). \tag{5.15}$$

Again, G_{ij}^1 and ϵ_{ij}^1 belong to the objective function and are distinct from $G_i^{2|1}$ and $\epsilon_{ij}^{2|1}$ of the ucpdf. While G_{ij}^1 is known explicitly, it is a very complicated function of u_1. Furthermore, several of the terms in (5.12) can combine, but the complexity of G_{ij}^1 makes adding them cumbersome. Fortunately, a method for simplifying G_{ij}^1 into a sum of sign functions was presented in Appendix A of [DSY18], which is restated in Appendix B for convenience. Using this method, the coefficient functions become

$$\mathcal{G}_{ij}^1 = a_{i0}^1 + \sum_{\substack{l=0 \\ l \neq j}}^{N_i^{2|1}} a_{il}^1 \operatorname{sgn}\left(\xi_{il}^1 - \Phi\bar{x}_1 - u_1\right) \tag{5.16}$$

where the ρ_{ij}^1 have been transformed into a_{ij}^1 using Appendix B. More importantly, a_{ij}^1 is constant, so $\mathcal{G}_{ij}^1$ is a sum of simple sign functions of u_1.

Expanding the exponential terms explicitly, we get

$$\begin{aligned}
\epsilon_{10}^1 &= \exp\left(-\frac{1}{\beta}\left|-\frac{\Phi z_1}{H} - \Phi\bar{x}_1 - u_1\right|\right) \\
\epsilon_{11}^1 &= \exp\left(-S\left|-\frac{\Phi z_1}{H} - \Phi\bar{x}_1 - u_1\right|\right) \\
\epsilon_{20}^1 &= \exp\left(-\frac{1}{\beta}\left|-\Phi\bar{x}_1 - u_1\right|\right) \\
\epsilon_{21}^1 &= \exp\left(-S\left|-\Phi\bar{x}_1 - u_1\right|\right) \\
\epsilon_{30}^1 &= \exp\left(-\frac{|H|}{\gamma|\Phi|}\left|-\frac{\Phi z_1}{H} - \Phi\bar{x}_1 - u_1\right|\right. \\
&\qquad\left. - \frac{1}{\alpha|\Phi|}\left|-\Phi\bar{x}_1 - u_1\right|\right) \\
\epsilon_{31}^1 &= \exp\left(-S\left|-\frac{\Phi z_1}{H} - \Phi\bar{x}_1 - u_1\right| - \frac{1}{\alpha|\Phi|}\left|-\frac{\Phi z_1}{H}\right|\right) \\
\epsilon_{32}^1 &= \exp\left(-S\left|-\Phi\bar{x}_1 - u_1\right| - \frac{|H|}{\gamma|\Phi|}\left|\frac{\Phi z_1}{H}\right|\right)
\end{aligned} \tag{5.17}$$

We can see from the integral formula (7.3) of the Appendix that the arguments in the sign functions are the same as the arguments in the absolute values of the exponentials. Therefore, the offsets from u_1 in the sign functions in $\mathcal{G}_{ij}^1$ are the same as the offsets in the terms of (5.17). Furthermore, they are the same as the offsets in $\bar{f}_{X_2|Y_1}$, except for a sign change and shift by $\Phi\bar{x}_1$.

The constant elements in the exponentials can be absorbed into their associated $\mathcal{G}_{ij}^1$ functions. Furthermore, the terms with the same exponentials can be combined by summing up their associated $\mathcal{G}_{ij}^1$ functions. In this way, we collapse the double indices ij to just i, and the coefficients become

$$\bar{G}_i^1 = \bar{a}_{i0}^1 + \sum_{l=1}^{N_i^{2|1}} \bar{a}_{il}^1 \text{sgn}\left(-\xi_{il}^1 - \Phi\bar{x}_1 - u_1\right) \tag{5.18}$$

$$\tilde{G}_i^1 = \tilde{a}_{i0}^1 + \tilde{a}_{i1}^1 \text{sgn}\left(-\xi_{i1}^1 - \Phi\bar{x}_1 - u_1\right). \tag{5.19}$$

We can now replace the parameters ξ_{il}^k used for the integration with the original parameters $\xi_{il}^{k+1|k}$ from the ucpdf to get

$$\bar{G}_i^1 = \bar{a}_{i0}^1 + \sum_{l=1}^{N_i^{2|1}} \bar{a}_{il}^1 \text{sgn}\left(-\xi_{il}^{2|1} - \Phi\bar{x}_1 - u_1\right) \tag{5.20}$$

$$\tilde{G}_i^1 = \tilde{a}_{i0}^1 + \tilde{a}_{i1}^1 \text{sgn}\left(-\xi_{i1}^{2|1} - \Phi\bar{x}_1 - u_1\right) \tag{5.21}$$

Note that we separate the coefficient terms associated with $\rho_{ij}^1 = S$ from the other terms and designate them with $\tilde{G}_i^1$. In this way, $\bar{G}_i^1$ has the same structure as $G_{i,N_i^{2|1}}^{2|1}$, which greatly simplifies the indexing. The number of elements are described by the same vector $N^{2|1}$. The additional terms, $\tilde{G}_i^1$, only has one sign function and has the same structure as $G_{i1}^{2|1}$. After this simplification, we can then write the objective function as

$$J_{Y_1}(u_1) = e^{-c|u_1|}\left(\sum_{i=1}^{3} \bar{G}_i^1 \bar{\epsilon}_i^1 + \sum_{i=1}^{2} \tilde{G}_i^1 \tilde{\epsilon}_i^1\right) \tag{5.22}$$

where

$$\bar{\epsilon}_i^1 = \exp\left(-\sum_{l=1}^{N_i^{2|1}} \rho_{il}^1 \left|-\xi_{il}^1 - \Phi\bar{x}_1 - u_1\right|\right) \tag{5.23}$$

$$\tilde{\epsilon}_i^1 = \exp\left(-\rho_{i0}^1 \left|-\xi_{i1}^1 - \Phi\bar{x}_1 - u_1\right|\right). \tag{5.24}$$

Again, we can replace the temporary parameters ρ_{il^k} used for the integration with the original parameters $\rho_{il}^{k+1|k}$ from the ucpdf to get

$$\bar{\epsilon}_i^1(u_1) = \exp\left(-\sum_{l=1}^{N_i^{2|1}} \rho_{il}^{2|1} \left|-\xi_{il}^{2|1} - \Phi\bar{x}_1 - u_1\right|\right) \tag{5.25}$$

$$\tilde{\epsilon}_i^1(u_1) = \exp\left(-S\left|-\xi_{i1}^{2|1} - \Phi\bar{x}_1 - u_1\right|\right). \tag{5.26}$$

The solution is divided into two parts to highlight the structure inherited from the ucpdf. In particular, the coefficient function $\bar{G}_i^1$ has the same form as $G_i^{2|1}$ in (5.3), and the associated exponential function $\bar{\epsilon}_i^1$ has the same form as $\epsilon_i^{2|1}$. The only new terms come from the control cost $\exp(-c|u_1|)$, as well as $\tilde{G}_i^1$ and the associated exponential function $\tilde{\epsilon}_i^1$, which contains the weighting S.

Now that we have the objective function in closed form, we can then maximize $J_{Y_1}(u_1)$ to obtain u_1^*. Unfortunately, it is a complicated function of u_1, so the optimization must be performed numerically. However, since it is also unimodal, there are a wide variety of tools available to do so efficiently. The specific methods will not be discussed in this paper.

5.2.2 Numerical Example: u^* at $k = 1$

Figure 5.1 shows the relationship between the optimal control u_1^* and the measurement z_1 at step $k = 1$ for $\Phi = 1.1, H = 1, c = 1, S = 3, \alpha = 1, \beta = 0.25, \gamma = 0.33$. This is similar to results shown in [SYD18]. The control goes to zero for a certain set of estimates around zero, suggesting that the controller takes into account the uncertainty of x_1 and hedges to avoid unnecessarily increasing the cost. This dead zone of zero control about small measurements helps explain the control performance in the general control case.

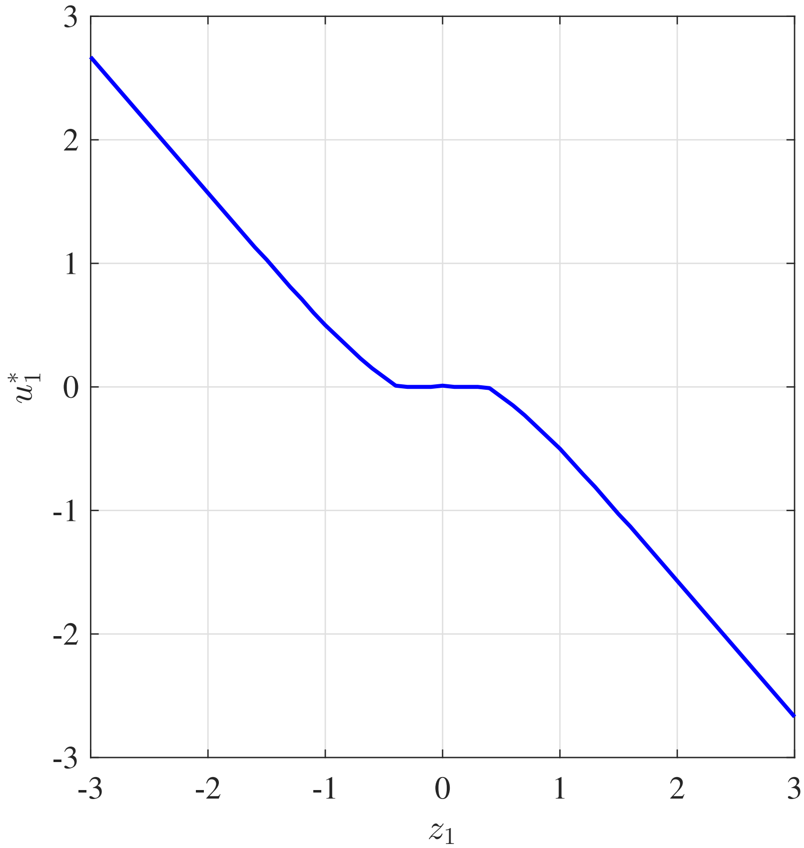


Figure 5.1: Optimal control u_1^* versus measurement z_1 for $c = 1, S = 3, \alpha = 1, \beta = 0.25, \gamma = 0.33$

5.3 Objective Function at Step k

After repeating the process used to obtain $J_{Y_1}(u_1)$ for several steps, a pattern emerges such that we can deduce the objective function for any k. The objective function for step k is

$$J_{Y_k}(u_k) = e^{-c|u_k|} \int_{-\infty}^{\infty} e^{-S|x_{k+1}+\Phi\bar{x}_k+u_k|} \times \bar{f}_{X_{k+1}|Y_k}(x_{k+1}|\boldsymbol{y}_k)\, dx_{k+1} \tag{5.27}$$

From [DSY18], the *a priori* ucpdf at step $k+1$ is given by

$$\bar{f}_{X_{k+1}|Y_k}(x_{k+1}|\boldsymbol{y}_k) = \sum_{i=1}^{n^{k+1|k}} G_i^{k+1|k} \epsilon_i^{k+1|k} \tag{5.28}$$

where

$$
\begin{aligned}
G_i^{k+1|k} &= a_{i0}^{k+1|k} + \sum_{l=1}^{N_i^{k+1|k}} a_{il}^{k+1|k} \operatorname{sgn}\left(\xi_{il}^{k+1|k} - x_{k+1}\right) \\
\epsilon_i^{k+1|k} &= \exp\left(-\sum_{l=1}^{N_i^{k+1|k}} \rho_{il}^{k+1|k} \left|\xi_{il}^{k+1|k} - x_{k+1}\right|\right),
\end{aligned}
\tag{5.29}
$$

$a_{i0}^{k+1|k}, a_{il}^{k+1|k}$, $\xi_{il}^{k+1|k}$ and $\rho_{il}^{k+1|k}$ are known, the number of terms

$$
n^{k+1|k} = \frac{(k+1)(k+2)}{2}, \tag{5.30}
$$

and the number of elements for each term are the elements of the vector

$$
N^{k+1|k} = \begin{bmatrix} \mathbf{1}^{k+1} \times 1 \\ \mathbf{1}^{k} \quad \times 2 \\ \mathbf{1}^{k-1} \times 3 \\ \vdots \\ \mathbf{1}^{1} \times (k+1) \end{bmatrix}. \tag{5.31}
$$

To obtain the solution to (5.27), we first re-write the integrand in (5.27) by merging $e^{-c|u_k|}$ with $\epsilon_i^{k+1|k}$ and augmenting $G_i^{k+1|k}$ so that the coefficient and exponential parts have the same number of elements. For each term i, they become

$$
\begin{aligned}
\bar{G}_i^{k+1|k} = a_{i0}^{k+1|k} &+ \sum_{l=1}^{N_i^{k+1|k}} a_{il}^{k+1|k} \operatorname{sgn}\left(\xi_{il}^{k+1|k} - x_{k+1}\right) \\
&+ 0 \cdot \operatorname{sgn}\left(-\Phi \bar{x}_k - u_k - x_{k+1}\right),
\end{aligned}
\tag{5.32}
$$

where the 0 is a placeholder for $\operatorname{sgn}\left(-\Phi \bar{x}_k - u_k - x_{k+1}\right)$ and

$$
\begin{aligned}
\bar{\epsilon}_i^{k+1|k} = \exp &\left(-\sum_{l=1}^{N_i^{k+1|k}} \rho_{ij}^{k+1|k} \left|\xi_{ij}^{k+1|k} - x_{k+1}\right|\right) \\
&\times \exp\left(S\left|-\Phi \bar{x}_k - u_k - x_{k+1}\right|\right).
\end{aligned}
\tag{5.33}
$$

The parameters of (5.32) and (5.33) are summarized by

$$\begin{array}{lll} \varrho_{i0}^{k} = 0 & \xi_{i0}^{k} = -\Phi\bar{x}_k - u_k & \rho_{i0}^{k} = S \\ \varrho_{i1}^{k} = a_{i1}^{k+1|k} & \xi_{i1}^{k} = \xi_{i1}^{k+1|k} & \rho_{i1}^{k} = \rho_{i1}^{k+1|k} \\ \varrho_{i2}^{k} = a_{i2}^{k+1|k} & \xi_{i2}^{k} = \xi_{i2}^{k+1|k} & \rho_{i2}^{k} = \rho_{i2}^{k+1|k} \\ \vdots & \vdots & \vdots \\ \varrho_{i,\tilde{N}_i}^{k} = a_{i,\tilde{N}_i}^{k+1|k} & \xi_{i,\tilde{N}_i}^{k} = \xi_{i,\tilde{N}_i}^{k+1|k} & \rho_{i,\tilde{N}_i}^{k} = \rho_{i,\tilde{N}_i}^{k+1|k} \end{array} \tag{5.34}$$

where $\tilde{N}_i = N_i^{k+1|k}$ for short. We can then use Appendix B of [IS14] to solve the integral of (5.27) to obtain $J_{Y_k}(u_k)$, which can be written as

$$\begin{aligned} J_{Y_k}(u_k) = \exp\left(-c\left|u_k\right|\right) \sum_{i=1}^{n^{k+1|k}} \bar{G}_i^k \bar{\epsilon}_i^k(u_k) \\ + \exp\left(-c\left|u_k\right|\right) \sum_{i=1}^{k+1} \tilde{G}_i^k \tilde{\epsilon}_i^k(u_k) \end{aligned} \tag{5.35}$$

where

$$\bar{G}_i^k = \bar{a}_{i0}^k + \sum_{l=1}^{N_i^k} \bar{a}_{il}^k \operatorname{sgn}\left(-\xi_{il}^{k+1|k} - \Phi\bar{x}_k - u_k\right) \tag{5.36}$$

$$\tilde{G}_i^k = \tilde{a}_{i0}^k + \tilde{a}_{i1}^k \operatorname{sgn}\left(-\xi_{i1}^{k+1|k} - \Phi\bar{x}_k - u_k\right). \tag{5.37}$$

and

$$\bar{\epsilon}_i^k(u_k) = \exp\left(\sum_{l=1}^{N_i^{k+1|k}} \rho_{il}^{k+1|k} \left|-\xi_{il}^{k+1|k} - \Phi\bar{x}_k - u_k\right|\right) \tag{5.38}$$

$$\tilde{\epsilon}_i^k(u_k) = \exp\left(-S\left|-\xi_{i1}^{k+1|k} - \Phi\bar{x}_k - u_k\right|\right) \tag{5.39}$$

5.4 Numerical Example for $k = 1, \cdots, 50$

We recursively and analytically computed and optimized the cost function at each step for the first 50 steps using the parameters $\Phi = 1.1, H = 1, \alpha = 1, \beta = 0.25, \gamma = 0.33, c =$

$0.5, S = 2.99$. The measurement and process noise were generated using a Laplace random number generator, and a spike was added to the measurement at $k = 5, 6, 20$ and 30 to test the resilience of the controller. We initiate the state at $x_1 = 3$ to clearly see the system driven to zero. Furthermore, we simulate control saturation so that $-1 \leq u_k \leq 1$.

The cost function at $k = 40, 41$ and 42 is shown in Figure 5.2, where the progression of u^* to zero is seen. We can also see that they are unimodal and have a cusp at $u_k = 0$ corresponding to the control penalty term. The measurement noise v_k, optimal control u_k^*, and the resultant state $x_k = \bar{x}_k + x_k$ are shown in Figure 5.3. As expected, the control maximized to drive the state to zero and then settles quickly. Furthermore, when the state is close to zero, the control takes on a value of zero. This seems to be due to the uncertainty in the state in combination with the absolute value penalty on the control. This behavior does not occur in the LQG controller, because the cost of the quadratic penalty on the control is negligible around zero.

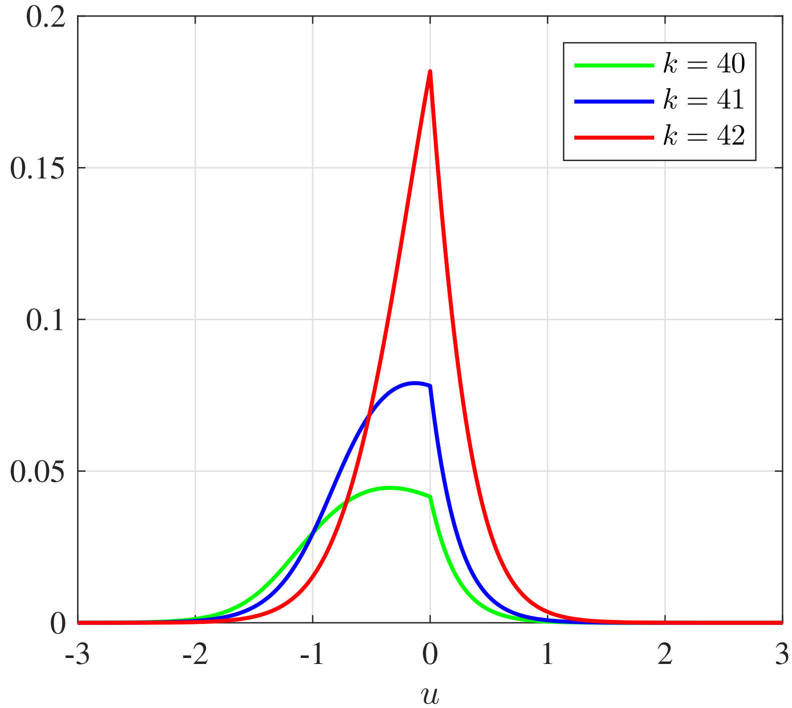


Figure 5.2: Cost function $J_{Y_k}(u_k)$ versus control u_k for $s = 2, S = 3, \alpha = 1.0, \beta = 0.25, \gamma = 0.33$

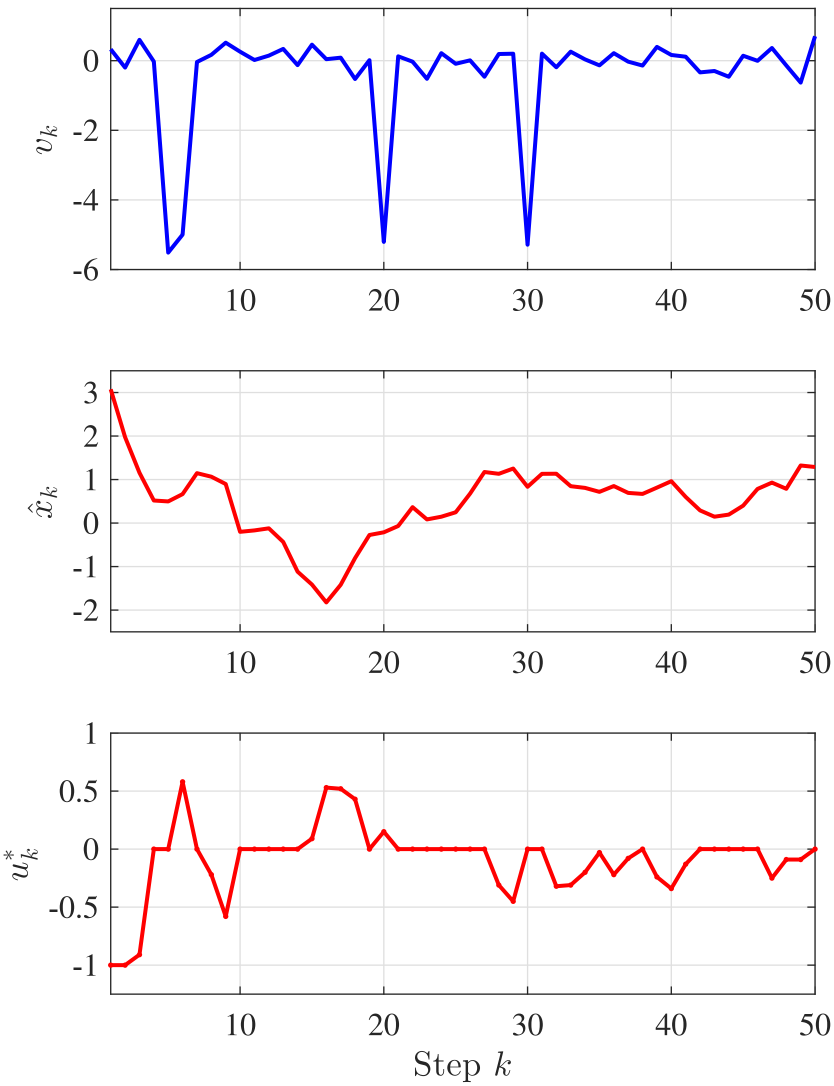


Figure 5.3: Measurement noise v_k, full state x_k, and control u_k with initial $\bar{x}_1 = 3$ and $c = 2, S = 3, \alpha = 1, \beta = 0.25, \gamma = 0.33$

CHAPTER 6

Conclusions

The main contribution of this work is the theoretical development of an analytic recursive algorithm for state estimation of a discrete linear system with scalar additive Laplace process and measurement noises. Proof-of-concept numerical implementations were demonstrated for the scalar and 2-dimensional cases, including explicit closed-form expressions for the conditional mean and error variance. Using the scalar estimation results, a one-step predictive controller was also derived and implemented numerically.

In addition, two supporting contributions have allowed for significant progress in the vector Cauchy estimator. First, the representation of the coefficient function using a basis of sign functions has greatly reduced the complexity of defining the coefficient functions. The cascading, nested fractions of sums of sign functions was turned into a sum of products of sign functions, making it straight-forward to add two terms. Second, the key integral derived in Appendix B of [IS14] was generalized to account for all functions of sign functions, rather than only sums of sign functions. Together, these two contributions made it possible to combine, update and propagate terms for both the Laplace and Cauchy estimators in a tractable manner.

6.1 Next steps

This work primarily presents the theoretical background in Laplace estimation, and the proof-of-concept code in MATLAB has demonstrated that the algorithms are correct. In addition, interesting ideas have been raised that offer a glimpse into the possibilities afforded by Laplace-based estimation and control. The following topics represent key steps to address in order to implement the Laplace estimator in real-world situations or new directions to take Laplace-based research.

6.1.1 Predictive term combination

Currently, the algorithms used to combine terms in $\mathbb{R}^2$ involve normalizing and sorting the hyperplanes across all terms and then applying a metric to determine if they are equal. This process is very computationally expensive, taking a similar amount of time as the propagation itself. However, in the scalar case, there was a fairly straight-forward way of determining *a priori* which terms will combine, and it is likely possible for cases in higher dimensions as well.

6.1.2 Term reduction or windowing

Even in the 2-dimensional case, the number of terms appear to grow faster rate than a polynomial function of the step k. The goal of a real-time estimator is to keep the size of the expressions constant. Therefore, a windowing technique similar to the one used for the 2-state Cauchy estimator in [SIF14] could be implemented to do this.

6.1.3 Re-mapping hyperplanes

Because the number of terms in subsequent steps come from the elements in the prior step, it is necessary to reduce the number of elements to help control the growth of terms. In the 2-dimensional case, some elements go away when the θ component of their associated hyperplanes go to zero. It may be possible to locate elements whose θ terms are near zero to force them to zero, particularly if they are far from the relevant parts of the pdf.

In a preliminary investigation, a pdf of the general form in (2.43) with $\mathcal{G} = 1$ was fitted to a scalar cpdf at step $k = 25$. However, it used only one-tenth as many elements, distributed evenly a few standard deviations around the mean. Surprisingly, this rather crude approximation had an error of less than 1 percent. It may be possible to improve upon this with less trivial coefficient functions as well. The immediate impact is that we only use one-tenth number of elements, thereby reducing the number of future terms as well.

6.1.4 Limit hyperplane density

While the number of hyperplanes grows with each update, at some point, the increase in hyperplane density will have segmented the relevant regions of the conditional pdf to such a fine mesh that adding more hyperplanes may not significantly change the relevant statistics. Therefore, one could consider a scheme where the number of hyperplanes were fixed and possibly manipulated in such a way that does not change the mean and variance. This will contribute to the effort to hold fixed the amount of memory it takes to define the conditional pdf after a certain number of steps have been reached.

6.1.5 Switch to indicator functions

One of the drawbacks of using sign functions is that it is difficult to identify which terms are important based on the scale of their coefficients. Indeed, it is often the case that individual terms cannot be singled out because they cancel another term at certain points. Such is the nature of sign functions. It may be possible to re-derive the equations using indicator functions instead. With indicator functions, one may be able to reduce either terms or elements by isolating individual faces.

6.1.6 Comparison to particle filters and other methods

As stated in the introduction, particle filters have been employed to deal with Laplace densities. The relative accuracy and computational requirements should be compared to see if one should be used over the other in practice. Certainly, the Laplace estimator can be used as a check against all simulation Monte-Carlo methods involving linear systems with Laplace noise.

6.1.7 Robustness

Further theoretical investigations into the robustness of the Laplace estimator are needed, due to errors in initial condition or noise characterizations. In particular, will two estimators initialized with different α converge asymptotically? Will the estimation errors remain bounded for an incorrect β or γ, and what is the size of that bound? These results would be relevant to the viability of the Laplace estimator as an alternative to the Kalman filter and other methods.

6.1.8 Laplace controller in two or more dimensions

The one-step scalar Laplace controller demonstrated some interesting properties not usually seen in a typical linear controller. This area is rich for exploration, including extension to greater number of steps or to higher dimensions. Furthermore, optimization algorithms for finding the optimal control may need to be surveyed.

CHAPTER 7

Appendix

7.1 Key integral formula

A key integral formula was derived by Idan and Speyer in Appendix B of [IS14] and is restated here for convenience. It comes in a simple form,

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n} \eta_i \left|\xi_i - x\right| + jzx\right) dx \\ &= \sum_{i=1}^{n} g_i \left(\sum_{\substack{l=1 \\ l \neq i}}^{n} \eta_l \mathrm{sgn}\left(\xi_l - \xi_i\right)\right) \exp\left(-\sum_{\substack{l=1 \\ l \neq i}}^{n} \eta_l \mathrm{sgn}\left(\xi_l - \xi_i\right)\right) \end{aligned} \tag{7.1}$$

where

$$g_i = \frac{1}{jz + \eta_i + \sum\limits_{\substack{l=1 \\ l \neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right)} - \frac{1}{jz - \eta_i + \sum\limits_{\substack{l=1 \\ l \neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right)} \tag{7.2}$$

as well as a more general form,

$$\begin{aligned} I &= \int_{-\infty}^{\infty} g\left(\sum_{i=1}^{n} \rho_i \mathrm{sgn}\left(\xi_i - x\right)\right) \exp\left(-\sum_{i=1}^{n} \eta_i \left|\xi_i - x\right| + jzx\right) dx \\ &= \sum_{i=1}^{n} G_i \left(\sum_{\substack{l=1 \\ l \neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right), \sum_{\substack{l=1 \\ l \neq i}}^{n} \eta_l \mathrm{sgn}\left(\xi_l - \xi_i\right)\right) \exp\left(-\sum_{\substack{l=1 \\ l \neq i}}^{n} \eta_l \mathrm{sgn}\left(\xi_l - \xi_i\right) + jz\xi_i\right) \end{aligned} \tag{7.3}$$

where

$$G_i = \frac{\rho_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \text{sgn}\,(\xi_l - \xi_i)}{jz + \eta_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \text{sgn}\,(\xi_l - \xi_i)} - \frac{-\rho_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \text{sgn}\,(\xi_l - \xi_i)}{jz - \eta_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \text{sgn}\,(\xi_l - \xi_i)} \tag{7.4}$$

7.2 Scalar piecewise constant functions as sum of signs

Theorem 7.2.1. *For*

$$\mathcal{A} = \{A_0, A_1, \cdots, A_n\} \triangleq \{(-\infty, \xi_1), [\xi_1, \xi_2), \cdots, [\xi_n, +\infty)\} \tag{7.5}$$

with $\xi_1 < \cdots < \xi_n \in \mathbb{R}$, *any* $g : \mathbb{R} \to \mathbb{R}$ *constant on* $A_i \in \mathcal{A}$, *can be simplified to*

$$g(x) = \rho_0 + \sum_{i=1}^{n} \rho_i sgn\,(\xi_i - x) \tag{7.6}$$

where, for $x_i \in A_i$,

$$\begin{aligned} \rho_0 &= \frac{g(x_0) + g(x_n)}{2} \\ \rho_i &= \frac{g(x_{i-1}) - g(x_i)}{2} \end{aligned} \tag{7.7}$$

Proof. By observation, any function s_i constant on $A_i \in \mathcal{A}$ and zero elsewhere can be written as a sum of two unit step function. For example, s with value 1 for $1 \leq x < 2$ and 0 elsewhere can can be expressed as $\text{u}\,(x-1) - \text{u}\,(x-2)$. Therefore, $g(x)$ can be expressed as the sum of s_i. Since the step function

$$\text{u}\,(x - \xi) = \frac{1 - \text{sgn}\,(\xi - x)}{2}, \tag{7.8}$$

it is clear that $g(x)$ is also a sum of sign functions. To find ρ_i, we first construct $g(x)$ in terms of $u(x - \xi_i)$ in order from most negative to positive x, so that

$$g(x) = g(x_0) + \sum_{i=1}^{n} [g(x_i) - g(x_{i-1})]\,\text{u}\,(x - \xi_i) \tag{7.9}$$

for $x_i \in A_i$. We can then covert $g(x)$ to a sum of sign functions using the conversion (7.8) to get

$$
\begin{aligned}
g(x) &= g(x_0) + \sum_{i=1}^{n} \left(g(x_i) - g(x_{i-1})\right) \left[\frac{1 - \operatorname{sgn}(\xi_i - x)}{2}\right] \\
&= g(x_0) + \sum_{i=1}^{n} \frac{g(x_i) - g(x_{i-1})}{2} + \sum_{i=1}^{n} \frac{g(x_{i-1}) - g(x_i)}{2} \cdot \operatorname{sgn}(\xi_i - x) \\
&= \frac{g(x_0) + g(x_n)}{2} + \sum_{i=1}^{n} \frac{g(x_{i-1}) - g(x_i)}{2} \cdot \operatorname{sgn}(\xi_i - x) \\
&\triangleq \rho_0 + \sum_{i=1}^{n} \rho_i \operatorname{sgn}(\xi_i - x)
\end{aligned}
\tag{7.10}
$$

□

7.3 Scalar characteristic function

Given the form of the general scalar ucpdf in (2.43), the Fourier transform is

$$
\begin{aligned}
\bar{\phi}_{X|Y} &= \int_{\mathbb{R}} e^{j\nu x} \bar{f}_{X|Y}(x|y)\, dx \\
&= \int_{\mathbb{R}} e^{j\nu x} \sum_{i=1}^{n} g_i(x) \epsilon_i(x)\, dx \\
&= \sum_{i=1}^{n} \int_{\mathbb{R}} e^{j\nu x} g_i(x) \epsilon_i(x)\, dx
\end{aligned}
\tag{7.11}
$$

where

$$
\begin{aligned}
g_i(x) &= \rho_0 + \sum_{l=1}^{m} \rho_l \operatorname{sgn}(\xi_l - x) \\
\epsilon_i(x) &= \exp\left(-\sum_{l=1}^{m} \eta_l \left|\xi_l - x\right|\right)
\end{aligned}
\tag{7.12}
$$

Hence, $\bar{\phi}_{X|Y}$ can be computed term-wise as

$$
\begin{aligned}
\bar{\phi}^{i}_{X|Y} &= \int_{\mathbb{R}} e^{j\nu x} g_i(x) \epsilon_i(x)\, dx \\
&= \int_{\mathbb{R}} g_i(x) \exp\left(-\sum_{l=1}^{m} \eta_l \left|\xi_l - x\right| + j\nu x\right)\, dx
\end{aligned}
\tag{7.13}
$$

The solution to this integral is given by (7.3), or

$$\bar{\phi}^{i}_{X|Y} = \sum_{i=1}^{n} G_i(\nu) \exp\left(-\sum_{\substack{l=1\\l\neq i}}^{n} \eta_l \mathrm{sgn}\left(\xi_l - \xi_i\right) + j\nu\xi_i \right) \tag{7.14}$$

where

$$G_i = \frac{\rho_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right)}{j\nu + \eta_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right)} - \frac{-\rho_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right)}{j\nu - \eta_i + \sum\limits_{\substack{l=1\\l\neq i}}^{n} \rho_l \mathrm{sgn}\left(\xi_l - \xi_i\right)} \tag{7.15}$$

However, since all the variables are constant with the exception of ν, $G_i(\nu)$ and $\epsilon_i(\nu)$ have the form

$$\begin{aligned} G_i(\nu) &= \frac{a_{i1}}{j\nu + a_{i2}} - \frac{b_{i1}}{j\nu + b_{i2}} \\ \epsilon_i(\nu) &= \exp\left(c_{i1} + c_{i2} \cdot j\nu\right) \end{aligned} \tag{7.16}$$

7.4 Proofs of moments from characteristic function

The normalization factor of $\bar{f}_{X|Y}(x)$ and moments of $f_{X|Y}(x)$ given its characteristic function $\phi_{X|Y}(\nu)$ are

$$\begin{aligned} f_Y &= \bar{\phi}_{X|Y}(0) \\ E[X|Y] &= \left[\frac{1}{j}\frac{\partial \phi_{X|Y}(\nu)}{\partial \nu}\right]_{\nu=0} \\ E[X^2|Y] &= \left[-\frac{\partial^2 \phi_{X|Y}(\nu)}{\partial \nu^2}\right]_{\nu=0} \end{aligned} \tag{7.17}$$

The proofs of these expressions is fairly straight-forward. The characteristic function is given by

$$\bar{\phi}_{X|Y}(\nu) = \int_{\mathbb{R}} \bar{f}_{X|Y}(x) \exp(j\nu x)\, dx \tag{7.18}$$

Evaluated at $\nu = 0$,

$$\begin{aligned}\bar{\phi}_{X|Y}(0) &= \int_{\mathbb{R}} \bar{f}_{X|Y}(x)\, dx \\ &= f_Y.\end{aligned} \tag{7.19}$$

To determine the moments, first consider the partial derivative of $\phi_{X|Y}$ with respect of ν, which is

$$\begin{aligned}\frac{1}{j}\frac{\partial \phi_{X|Y}(\nu)}{\partial \nu} &= \frac{1}{j}\frac{\partial}{\partial \nu}\int_{\mathbb{R}} f_{X|Y}(x)\exp(j\nu x)\ dx \\ &= \frac{1}{j}\int_{\mathbb{R}} \frac{\partial}{\partial \nu} f_{X|Y}(x)\exp(j\nu x)\ dx \\ &= \frac{1}{j}\int_{\mathbb{R}} jx f_{X|Y}(x)\exp(j\nu x)\ dx \\ &= \int_{\mathbb{R}} x f_{X|Y}(x)\exp(j\nu x)\ dx\end{aligned} \tag{7.20}$$

since ν is independent of x. When evaluated at $\nu = 0$,

$$\begin{aligned}\left[\frac{1}{j}\frac{\partial \phi_{X|Y}(\nu)}{\partial \nu}\right]_{\nu=0} &= \int_{\mathbb{R}} x f_{X|Y}(x)\ dx \\ &= E[X|Y]\end{aligned} \tag{7.21}$$

Similarly,

$$\begin{aligned}-\frac{\partial^2 \phi_{X|Y}(\nu)}{\partial \nu^2} &= -\frac{\partial^2}{\partial \nu_i \partial \nu_\ell}\int_{\mathbb{R}} f_{X|Y}(x)\exp(j\nu x)\ dx \\ &= -\int_{\mathbb{R}} \frac{\partial^2}{\partial \nu^2} f_{X|Y}(x)\exp(j\nu x)\ dx \\ &= -\int_{\mathbb{R}} -x^2 f_{X|Y}(x)\exp(j\nu x)\ dx \\ &= \int_{\mathbb{R}} x^2 f_{X|Y}(x)\exp(j\nu x)\ dx\end{aligned} \tag{7.22}$$

When evaluated at $\nu = 0$

$$\begin{aligned}\left[-\frac{\partial^2 \phi_{X|Y}(\nu)}{\partial \nu^2}\right]_{\nu=0} &= \int_{\mathbb{R}} x^2 f_{X|Y}(x)\ dx \\ &= E\left[X^2|Y\right]\end{aligned} \tag{7.23}$$

7.5 Scalar moments from characteristic function

Using the generic form for each term as well as the definitions for the moments from Section 7.4, we present the explicit form of the moments.

7.5.1 Normalization

The ith term of the normalization of $\bar{f}_{X|Y}$ is

$$\begin{aligned} f_{Y,i} &= \left[\phi^i_{X|Y}(0)\right] \\ &= \left(\frac{a_{i1}}{a_{i2}} - \frac{b_{i1}}{b_{i2}}\right) \exp\left(c_{i1}\right) \end{aligned} \tag{7.24}$$

7.5.2 First moment

The ith term of the first moment is given by

$$\begin{aligned} E_i[X|Y] &= \frac{1}{j}\frac{\partial \phi_{X|Y}(\nu)}{\partial \nu} \\ &= \frac{1}{j}\left[\frac{\partial G_i}{\partial \nu}\epsilon_i + G_i\frac{\partial \epsilon_i}{\partial \nu}\right]_{\nu=0} \end{aligned} \tag{7.25}$$

where

$$\begin{aligned} \frac{\partial G_i}{\partial \nu} &= -\frac{a_{i1}\cdot j}{\left(j\nu + a_{i2}\right)^2} + \frac{b_{i1}\cdot j}{\left(j\nu + b_{i2}\right)^2} \\ &\to -\frac{a_{i1}\cdot j}{\left(a_{i2}\right)^2} + \frac{b_{i1}\cdot j}{\left(b_{i2}\right)^2} \\ \frac{\partial \epsilon_i}{\partial \nu} &= jc_{i2}\exp\left(c_{i1} + c_{i2}\cdot j\nu\right) \\ &\to jc_{i2}\exp\left(c_{i1}\right) \end{aligned} \tag{7.26}$$

7.5.3 Second moment

The ith term of the second moment is given by

$$\begin{aligned}E_i[X^2|Y] &= -\frac{\partial^2 \phi_{X|Y}(\nu)}{\partial \nu^2} \\ &= -\left[\frac{\partial^2 G_i}{\partial \nu^2}\epsilon_i + 2\frac{\partial G_i}{\partial \nu}\cdot\frac{\partial \epsilon_i}{\partial \nu} + G_i\frac{\partial^2 \epsilon_i}{\partial \nu^2}\right]_{\nu=0}\end{aligned} \tag{7.27}$$

where

$$\begin{aligned}\frac{\partial^2 G_i}{\partial \nu^2} &= -\frac{2a_{i1}}{(j\nu + a_{i2})^3} + \frac{2b_{i1}}{(j\nu + b_{i2})^3} \\ &\to -\frac{2a_{i1}}{(a_{i2})^3} + \frac{2b_{i1}}{(b_{i2})^3} \\ \frac{\partial^2 \epsilon_i}{\partial \nu^2} &= -c_{i2}^2 \exp\left(c_{i1} + c_{i2}\cdot j\nu\right) \\ &\to -c_{i2}^2 \exp\left(c_{i1}\right)\end{aligned} \tag{7.28}$$

7.6 Algorithm for finding coefficients in $\mathbb{R}^2$

The algebra involved is messy and difficult to capture algorithmically, but we can take advantage of the fact that G_j is piecewise constant. The jump discontinuities in the sign functions, defined by $\eta_j = \psi_j + \theta_j^T x = 0$, divides $\mathbb{R}^n$ into regions of constant values. By sampling these regions, we can use them as constraints to solve for the coefficients of (3.50).

The number of regions to sample is given by Steiner in (ref). For m sign functions in two dimensions, the maximum number of constant regions is given by

$$N_{r,max} = 1 + \sum_{i=1}^{m} i = 1 + \frac{m(m+1)}{2} = \frac{m^2+m+2}{2}. \tag{7.29}$$

In $\mathbb{R}^3$, it is given by

$$N_{r,max} = \frac{m^3+5m+6}{6}. \tag{7.30}$$

Even if we know how many regions there are, it may not easy to identify these regions. However, if we sample around every intersection point for the maximum number of intersections,

$N_{i,max}$, we will be able to locate at least a sufficient number of samples. Furthermore, it is straightforward to identify these intersections.

For m sign functions in $\mathbb{R}^n$,

$$N_{i,max} = \binom{m}{n}. \tag{7.31}$$

This assumes that every intersection point is unique and there are no parallel η_j. For example, for three sign functions in $\mathbb{R}^2$, the $N_{i,max} = 3$, which can be verified by hand. At each of these intersections, the space is divided into 2^n regions. In $\mathbb{R}^2$, there are four region, and in $\mathbb{R}^3$, there are eight. This is also verifiable by hand. Therefore, the number of sample points is bounded from above by

$$N_{s,max} = \binom{m}{n} \times 2^n \geq N_{r,max} \tag{7.32}$$

For three sign functions in $\mathbb{R}^2$, this becomes 12. If you draw the three lines by hand, you will see that there is a maximum of 7 regions, so the upper bound is not very tight.

Let us evaluate G_j at each sample point and arrange the values into the vector

$$y = \begin{bmatrix} G_j(x_1) \\ G_j(x_2) \\ \vdots \\ G_j(x_{N_{i,max}}) \end{bmatrix} \tag{7.33}$$

Clearly, many of these samples are redundant, but that is okay. We then evaluate the basis functions at those sample points and arrange them into the matrix

$$B = \begin{bmatrix} 1 & \operatorname{sgn}(\eta_1(x_1)) & \operatorname{sgn}(\eta_2(x_1)) & \cdots & \prod_{j=1}^{m-1} \operatorname{sgn}(\eta_j(x_1)) \\ 1 & \operatorname{sgn}(\eta_1(x_2)) & \operatorname{sgn}(\eta_2(x_2)) & \cdots & \prod_{j=1}^{m-1} \operatorname{sgn}(\eta_j(x_2)) \\ & & \vdots & & \\ 1 & \operatorname{sgn}\left(\eta_1(x_{N_{i,max}})\right) & \operatorname{sgn}\left(\eta_2(x_{N_{i,max}})\right) & \cdots & \prod_{j=1}^{m-1} \operatorname{sgn}\left(\eta_j(x_{N_{i,max}})\right) \end{bmatrix} \tag{7.34}$$

The samples (7.33) and basis (7.34) are related as

$$y = Ba, \tag{7.35}$$

where

$$a = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{N_b} \end{bmatrix}^T, \tag{7.36}$$

where N_b is the number of distinct polytopes in $\mathbb{R}^n$ generated by the m sign functions. Note that (7.35) is actually an under-defined problem. There are fewer linearly independent rows than there are columns, so the solution to (7.35) is not unique. However, we can compute the least-norm solution by first removing redundant rows to get $\hat{B}$, where

$$\text{rank}\left(\hat{B}\right) = N_{r,max} \tag{7.37}$$

and is full row-rank. We can use MATLAB and `unique` to eliminate redundant rows in (7.35) to get $\hat{B}$ and the corresponding $\hat{y}$. The least-norm solution becomes

$$\hat{a}_{\text{ln}} = \hat{B}^T \left(\hat{B}\hat{B}^T\right)^{-1} \hat{y}, \tag{7.38}$$

which is what we get when we use the MATLAB `pinv` function. Note that the using the left matrix divide, or backslash (`A\b`), to solve this equation yields the solution with the greatest number of zero elements instead of the least-norm.

An example g function defined by

$$g(x) = \left(\frac{1+s_1}{2+2.5s_2} - \frac{-1+s_1}{-2+2.5s_2}\right) \cdot \frac{1}{3+s_3} \tag{7.39}$$

where

$$\begin{aligned} s_1 &= \text{sgn}\left(\psi_1 + \theta_1^T x\right) \\ s_2 &= \text{sgn}\left(\psi_2 + \theta_2^T x\right) \\ s_3 &= \text{sgn}\left(\psi_3 + \theta_3^T x\right) \end{aligned} \tag{7.40}$$

and

$$\begin{aligned} \psi_1 &= 4, & \theta_1 &= \begin{bmatrix} 2 & -1 \end{bmatrix} \\ \psi_2 &= 6, & \theta_2 &= \begin{bmatrix} 3 & 5 \end{bmatrix} \\ \psi_3 &= -30, & \theta_3 &= \begin{bmatrix} 1 & -20 \end{bmatrix} \end{aligned} \tag{7.41}$$

It is then converted to the sum-product form using basis functions as

$$\hat{g}(x) = s^T a \tag{7.42}$$

where

$$s = \begin{bmatrix} 1 & s_1 & s_2 & s_3 & s_1 s_2 & s_1 s_3 & s_2 s_3 \end{bmatrix}^T. \tag{7.43}$$

The samples were manually chosen at points

$$\left[\begin{array}{c|ccccccc} x_1 & -8 & -6 & -1 & 0 & 0 & 6 & 8 \\ \hline x_2 & -5 & 0 & -1 & 6 & -6 & 0 & -3 \end{array}\right] \tag{7.44}$$

as shown in Figure 7.1. Solving for a yields

$$a^T = \begin{bmatrix} 0 & -0.6667 & 0.8333 & 0 & 0 & 0.2222 & -0.2778 \end{bmatrix} \tag{7.45}$$

7.7 Proof of Theorem 3.3.1

Theorem 3.3.1 was conjectured by the author, in conjunction with Dr. Jason Speyer. However, the proof of Theorem 3.3.1 was completed by collaborators Dr. Rom Pinchasi and Dr. Moshe Idan from the Technion in Israel. Their paper containing the proof has yet to be published, but an excerpt, copied here with their permission, immediately follows.

Theorem 7.7.1. *Let $\mathcal{A}$ be a hyper-plane arrangement of n affine hyper-planes $H1, \cdots, H_n$ in $\mathbb{R}^d$ defined by $H_i = \{x | \langle x, v_i \rangle = c_i\}$, where $x \in \mathbb{R}^d, v_i \in \mathbb{R}^d$ is normal to H_i, and $c_i \in \mathbb{R}$. For every $1 \leq i \leq n$ let σ_i denote the indicator function of the open half-space $\{x | \langle x, v_i \rangle > c_i\}$ bounded by H_i. Let g be any function that is constant in the interior of every d-dimensional face in $\mathcal{A}$. Then there is a linear combination of products of d or less of the functions σ_i that is equal to g at any point that is not in $\cup_{i=1}^n H_i$.*

We start with a preliminary result that will be used to prove the main theorem presented next. When stated separately, not within the problem addressed in this note, its statement and proof can be greatly simplified, without hampering its generalization.

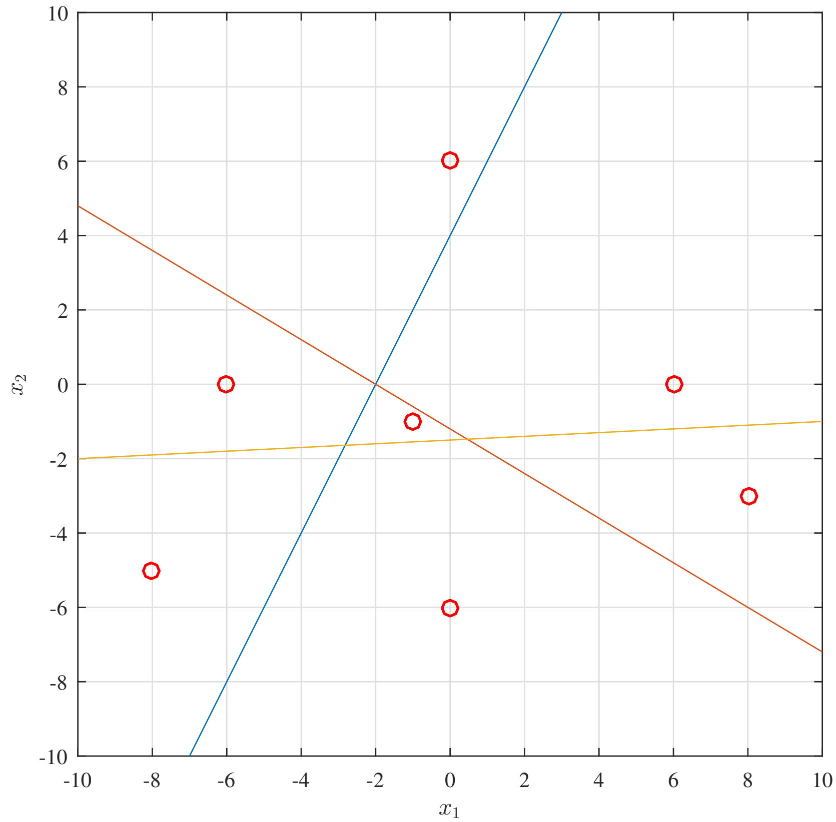


Figure 7.1: Manual selection of sectors of example g function

Lemma 7.7.2. *Let Δ_d be a d-simplex in $\mathbb{R}^d$. Let $H_1, \cdots, H_{d+1}$ be the $d+1$ affine hyperplanes supporting the facets of Δ_d. For $i = 1, \cdots, d+1$, let σ_i be the indicator function of the cosed half-space bounded by H_i and containing Δ_d. Then*

$$\prod_{i=1}^{d+1} (1 - \sigma_i) = 0. \tag{7.46}$$

Proof. Assume, without loss of generality, that $0 \in \Delta_d$. For $i = 1, \cdots, d+1$ we write H_i as $H_i = \{x | \langle x, v_i \rangle = c_i\}$, where $v_i \in \mathbb{R}^d$ (orthogonal to H_i) is chosen such that $c_i > 0$. Then $\Delta_d = \{x | \forall 1 \leq i \leq d+1, \langle x, v_i \rangle \leq c_i\}$. Observe that the statement of the Lemma is equivalent to saying that there is no vector $u \in \mathbb{R}^d$ such that $\langle u, v_i \rangle > c_i$ for every

$1 \leq i \leq d+1$. Assume to the contrary that there is such a vector u. Then for every $\alpha > 0$ and every $1 \leq i \leq d+1$ we have $\langle -\alpha u, v_i \rangle = -\alpha \langle u, v_i \rangle < -\alpha c_i < 0 < c_i$. In other words,$-\alpha u \in \Delta_d$ for every $\alpha > 0$. This is impossible as Δ_d is bounded. □

We can now proceed to the proof of Theorem 7.7.1. Observe that in order to prove Theorem 7.7.1 it is enough to consider functions g that are indicator functions of d-dimensional faces in $\mathcal{A}$.

Proof. Let F be a d-dimensional face in $\mathcal{A}$ and let g be the indicator function of F. Let $I \subseteq \{1, \cdots, n\}$ denote the set of indices i such that H_i supports F at a face of dimension $d-1$. Let I_a and I_b be a partition of I into to parts such that if $i \in I_a$, then $F \subset \{x | \langle x, v_i \rangle > c_i\}$ and if $i \in I_b$, then $F \subset \{x | \langle x, v_i \rangle < c_i\}$.

Observe that F is equal to the intersection of all open half-spaces containing F that are bounded by some hyper-plane H_i where $i \in I$. Therefore, the function $\tilde{g} = \prod_{i \in I_a} \sigma_i \cdot \prod_{i \in Ib} (1 - \sigma_i)$ is equal to g at any point not in $\cup_{i=1}^{n} H_i$.

Hence, if the cardinality of I is smaller than or equal to d we are done because $\tilde{g}$ can clearly be written as a linear combination of products of $|I|$ or less of the indicator functions $\sigma_1, \cdots, \sigma_n$.

If the cardinality of I is larger than d, then $\tilde{g}$ can still be written as a linear combination of products of the indicator functions $\sigma_1, \cdots, \sigma_n$, however the number of terms in each product may exceed d. Therefore, Theorem 7.7.1 will follow if we can show that the product of every $d+1$ of the indicator functions $\sigma_1, \cdots, \sigma_n$ is equal, on $\mathbb{R}^d \setminus \cup_{i=1}^{n} H_i$, to a linear combination of products of d or less of the indicator functions $\sigma_1, \cdots, \sigma_n$.

We prove this by induction of d. The basis of the induction is the case $d = 1$. In this case we have two indicator functions say σ_1 and σ_2. We would like to consider the function $\sigma_1 \sigma_2$ and express it as a linear combination of zero or one of the functions σ_1 and σ_2.

This could easily be left to the reader, but for completeness we bring the simple analysis here. For $i = 1, 2$ there exists x_i such that the functions σ_i is either the indicator function of

$\{x|x < x_i\}$ or of $\{x|x > x_i\}$. Without loss of generality assume that $x_1 \leq x_2$. We consider four possible cases.

Case 1. σ_1 is the indicator function of $\{x|x < x_1\}$ and σ_2 is the indicator function of $\{x|x < x_2\}$. In this case $\sigma_1\sigma_2 = \sigma_1$.

Case 2. σ_1 is the indicator function of $\{x|x < x_1\}$ and σ_2 is the indicator function of $\{x|x > x_2\}$. In this case $\sigma_1\sigma_2 = 0$.

Case 3. σ_1 is the indicator function of $\{x|x > x_1\}$ and σ_2 is the indicator function of $\{x|x < x_2\}$. In this case $\sigma_1\sigma_2$ is equal to the function $\sigma_1 + \sigma_2 - 1$.

Case 4. σ_1 is the indicator function of $\{x|x > x_1\}$ and σ_2 is the indicator function of $\{x|x > x_2\}$. In this case $\sigma_1\sigma_2$ is equal to the function σ_2.

This concludes the case $d = 1$ being the basis of induction.

For $d > 1$ we consider two possible cases:

Case 1. $k+1$ of the vectors $v_1, \cdots, v_{k+1}$ are linearly independent for some $1 \leq k \leq d$. Without loss of generality, assume that $v_!, \cdots, v_{k+1}$ are linearly independent. By a possible rotation of $\mathbb{R}^d$, we can assume that span $\{v_1, \cdots, v_{k+1}\} \subseteq$ span $\{e_1, \cdots, e_k\}$, where $e_1, \cdots, e_k$ are the first k elements of the standard basis of $\mathbb{R}^d$. Let $P : \mathbb{R}^d \to \mathbb{R}^k$ be the projection on the first k elements of the standard basis of $\mathbb{R}^d$. In $\mathbb{R}^k$, for every $1 \leq i \leq k+1$ we define $H'_i = \left\{x \in \mathbb{R}^k | \langle P(v_i), x\rangle = c_i\right\}$ and let $\sigma'_i : \mathbb{R}^k \to \mathbb{R}$ be the indicator function of $\left\{x \in \mathbb{R}^k | \langle P(v_i), x\rangle > c_i\right\}$. Observe that for every $x \in \mathbb{R}^d$ and $1 \leq i \leq k+1$,

$$\sigma_i(x) = \sigma'_i(P(x)). \tag{7.47}$$

Because $k < d$, we can apply the induction hypothesis for dimension k and conclude that $\prod_{i=1}^{k+1} \sigma'_i$ is equal to a linear combination of products of k or less of $\sigma'_1, \cdots, \sigma'_{k+1}$. Because of (7.47) it follows that $\prod_{i=1}^{k+1} \sigma_i$ is equal to a linear combination of k or les of $\sigma_1, \cdots, \sigma_{k+1}$.

Consequently, $\prod_{i=1}^{d+1} \sigma_i = \prod_{i=1}^{k+1} \sigma_i \cdot \prod_{i=k+2}^{d+1}$ is equal to a linear combination of products of d or less of $\sigma_1, \cdots, \sigma_{d+1}$.

Case 2. Every set of d vectors from $v_1, \cdots, v_{d+1}$ is linearly independent. We split into two possible subcases.

Case 2a. $\cap_{i=1}^{d+1} H_i = \emptyset$. In this case $H_1, \cdots, H_{d+1}$ are the $d+1$ affine hyper-planes supporting the facets of the d-simplex Δ_d whose vertices are $v_j = \cap_{i=1, i\neq j}^{d+1} h_i$ for $1 \leq j \leq d+1$. We may assume, without loss of generality, that $0 \in \Delta_d$. Let $I = \{1, \cdots, d+1\}$. Let I_a and I_b be a partition of I into two parts such that if $i \in I_a$, then $c_i < 0$, and if $i \in I_b$, then $c_i > 0$. Then $\prod_{i\in I_a} \sigma_i \cdot \prod_{i \in I_b}(1-\sigma_i)$ is the indicator function of the interior of Δ_d. Applying (7.46) of Lemma 7.7.2 yields $\prod_{i\in I_a}(1-\sigma_i) \cdot \prod_{i\in I_b} \sigma_i = 0$, which proves the theorem for this case.

Case 2b. $\cap_{i=1}^{d+1} H_i \neq \emptyset$. In this case $\cap_{i=1}^{d+1} H_i$ is a single point, because $v_1, \cdots, v_d$ are linearly independent. Without loss of generality we assume that this single point is 0. Consequently, $c_i = 0$ for $1 \leq i \leq d+1$. Let $\alpha_1, \cdots, \alpha_{d+1}$ be real numbers, not all zero, such that $\sum_{i=1}^{d+1} \alpha_i v_i = 0$. Notice that in this case $\alpha_i \neq 0$ for all $1 \leq i \leq d+1$, because every set of d vectors from $v_1, \cdots, v_{d+1}$ is linearly independent.

Define $I_a = \{1 \leq i \leq d+1 | \alpha_i > 0\}$ and $I_b = \{1 \leq i \leq d+1 | \alpha_i < 0\}$. Notice that I_a and I_b form a partition of $1, \cdots, d+1$. We claim that

$$\prod_{i\in I_a} \sigma_i \cdot \prod_{i\in I_b} (1-\sigma_i) = 0. \tag{7.48}$$

Observe that once (7.48) is established we are done, as (7.48) implies that $\prod_{i=1}^{d+1} \sigma_i$ is a linear combination of the products of d or less of $\sigma_1, \cdots, \sigma_{d+1}$.

To prove (7.48), notice that the contrary assumption is that there exits a vector u such that for every $i \in I_a$ we have $\langle u, v_i \rangle > 0$ and for every $i \in I_b$ we have $\langle u, v_i \rangle < 0$. It follows now from the definition of I_a and I_b that for every $1 \leq i \leq d+1$ we have $\alpha_i \langle u, v_i \rangle > 0$. This is a contradiction as

$$\sum_{i=1}^{d+1} \alpha_i \langle u, v_i \rangle = \langle u, \sum_{i=1}^{d+1} \alpha_i v_i \rangle = \langle u, 0 \rangle = 0. \tag{7.49}$$

$\square$

This concludes the excerpt from Rom Pinchasi and Moshe Idan.

7.8 Modification to the integral formula of Appendix B in [IS14]

Appendix B of [IS14] gives the solution for

$$I = \int_{-\infty}^{\infty} g\left(\sum_{i=1}^{n} \varrho_i \text{sgn}\left(\xi_i - \eta\right)\right) \exp\left(-\sum_{i=1}^{n} \rho_i \left|\xi_i - \eta\right| + j\nu\eta\right) d\eta \tag{7.50}$$

where g is an explicit function of a sum of sign functions. We extend the solution to include g as a function of a sum of products of sign functions.

$$I = \int_{-\infty}^{\infty} g\left(\eta\right) \exp\left(-\sum_{l=1}^{n} \rho_l \left|\xi_l - \eta\right| + j\nu\eta\right) d\eta \tag{7.51}$$

where, for some m and unique $\sigma_j \subseteq \{1, \cdots, n\}$,

$$g(\eta) = g\left(\sum_{l=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} \text{sgn}\left(\xi_\ell - \eta\right)\right). \tag{7.52}$$

As in the original derivation, $\text{sgn}\left(\xi_\ell - \eta\right)$ is constant on the interval $[\xi_i, \xi_{i+1}]$ such that

$$\text{sgn}\left(\xi_\ell - \eta\right) \triangleq s_i^\ell = \begin{cases} \text{sgn}\left(\xi_\ell - \xi_i\right), & i \neq \ell \\ -1, & i = \ell \end{cases}. \tag{7.53}$$

The definition of (7.53) is clear from Figures 7.2 and 7.3. We can see that when $\xi_\ell = \xi_i$, $\text{sgn}\left(\xi_\ell - \eta\right) = -1$ on $[\xi_i, \xi_{i+1}]$. When $\xi_\ell \neq \xi_i$, we evaluate at ξ_i to obtain $\text{sgn}\left(\xi_\ell - \eta\right) = \text{sgn}\left(\xi_\ell - \xi_i\right)$ on $[\xi_i, \xi_{i+1}]$.

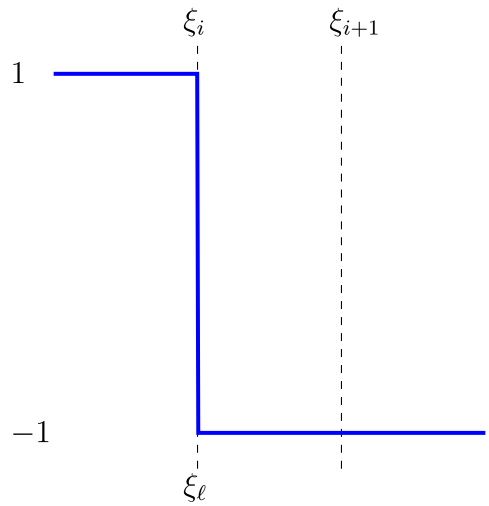


Figure 7.2: Definition of $\mathrm{sgn}\,(\xi_\ell - \eta)$ on $[\xi_i, \xi_{i+1}]$ for $\ell = i$

We perform the same procedure as in Appendix B to get

$$\begin{aligned} I &= \sum_{i=0}^{n} \left\{ \int_{\xi_i}^{\xi_{i+1}} g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} \mathrm{sgn}\,(\xi_\ell - \eta) \right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \eta) \mathrm{sgn}\,(\xi_\ell - \eta) \right) + j\nu\eta \right] d\eta \right\} \\ &= \sum_{i=0}^{n} \left\{ \int_{\xi_i}^{\xi_{i+1}} g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_i^\ell \right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \eta) s_i^\ell \right) + j\nu\eta \right] d\eta \right\}. \end{aligned} \tag{7.54}$$

This is then integrated to get

$$\begin{aligned} I = \sum_{i=0}^{n} &\left\{ \frac{ g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_i^\ell \right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_{i+1}) s_i^\ell \right) + j\nu\xi_{i+1} \right] }{ j\nu + \sum_{\ell=1}^{n} \rho_\ell s_i^\ell } \right\} \\ &- \sum_{i=0}^{n} \left\{ \frac{ g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_i^\ell \right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^\ell \right) + j\nu\xi_i \right] }{ j\nu + \sum_{\ell=1}^{n} \rho_\ell s_i^\ell } \right\} \end{aligned} \tag{7.55}$$

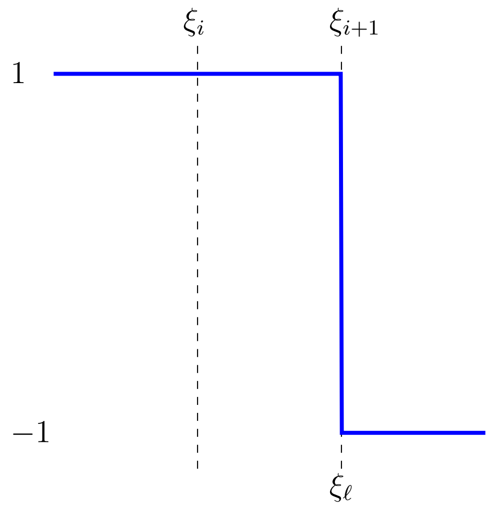


Figure 7.3: Definition of $\operatorname{sgn}(\xi_\ell - \eta)$ on $[\xi_i, \xi_{i+1}]$ for $\ell \neq i$

As in Appendix B of [IS14], the first sum in (7.55) can be manipulated as follows

$$\begin{aligned}
&\sum_{i=0}^{n}\left\{\frac{g\left(\sum_{j=1}^{m}\varrho_j\prod_{\ell\in\sigma_j}s_i^\ell\right)\exp\left[\left(-\sum_{\ell=1}^{n}\rho_\ell(\xi_\ell-\xi_{i+1})s_i^\ell\right)+j\nu\xi_{i+1}\right]}{j\nu+\sum_{\ell=1}^{n}\rho_\ell s_i^\ell}\right\}\\
&\qquad=\sum_{i=0}^{n-1}\left\{\frac{g\left(\sum_{j=1}^{m}\varrho_j\prod_{\ell\in\sigma_j}s_i^\ell\right)\exp\left[\left(-\sum_{\ell=1}^{n}\rho_\ell(\xi_\ell-\xi_{i+1})s_i^\ell\right)+j\nu\xi_{i+1}\right]}{j\nu+\sum_{\ell=1}^{n}\rho_\ell s_i^\ell}\right\}\\
&\qquad\quad+\frac{g\left(\sum_{j=1}^{m}\varrho_j\prod_{\ell\in\sigma_j}s_n^\ell\right)\exp\left[\left(-\sum_{\ell=1}^{n}\rho_\ell(\xi_\ell-\xi_{n+1})s_n^\ell\right)+j\nu\xi_{n+1}\right]}{j\nu+\sum_{\ell=1}^{n}\rho_\ell s_n^\ell}\\
&\qquad=\sum_{i=1}^{n}\left\{\frac{g\left(\sum_{j=1}^{m}\varrho_j\prod_{\ell\in\sigma_j}s_{i-1}^\ell\right)\exp\left[\left(-\sum_{\ell=1}^{n}\rho_\ell(\xi_\ell-\xi_i)s_{i-1}^\ell\right)+j\nu\xi_i\right]}{j\nu+\sum_{\ell=1}^{n}\rho_\ell s_{i-1}^\ell}\right\},
\end{aligned}\tag{7.56}$$

using the fact that $\xi_{n+1} = +\infty$ and $s_n^\ell = -1 \forall \ell \leq n$ to eliminate the extra term. Since $s_{i-1}^i = 1$ and $s_{i-1}^\ell = s_i^\ell$ for $i \neq \ell$, we can perform a substitution while being careful around

$i = l$. Luckily, the exponential goes to zero when $i = \ell$. Then, the term becomes

$$\sum_{i=1}^{n} \left\{ \frac{g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_{i-1}^{\ell}\right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_{i-1}^{\ell}\right) + j\nu\xi_i\right]}{j\nu + \sum_{\ell=1}^{n} \rho_\ell s_{i-1}^{\ell}} \right\}$$

$$= \sum_{i=1}^{n} \left\{ \frac{g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_{i-1}^{\ell}\right) \exp\left[\left(-\sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^{\ell}\right) j\nu\xi_i\right]}{j\nu + \rho_i + \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^{\ell}} \right\} \tag{7.57}$$

Similarly, the second sum in (7.55) can be manipulated as follows

$$\sum_{i=0}^{n} \left\{ \frac{g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_i^{\ell}\right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^{\ell}\right) + j\nu\xi_i\right]}{j\nu + \sum_{\ell=1}^{n} \rho_\ell s_i^{\ell}} \right\}$$

$$= \sum_{i=1}^{n} \left\{ \frac{g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_i^{\ell}\right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^{\ell}\right) + j\nu\xi_i\right]}{j\nu + \sum_{\ell=1}^{n} \rho_\ell s_i^{\ell}} \right\}$$

$$+ \frac{g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_0^{\ell}\right) \exp\left[\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_0) s_0^{\ell}\right) + j\nu\xi_0\right]}{j\nu + \sum_{\ell=1}^{n} \rho_\ell s_0^{\ell}} \tag{7.58}$$

$$= \sum_{i=1}^{n} \left\{ \frac{g\left(\sum_{j=1}^{m} \varrho_j \prod_{\ell \in \sigma_j} s_i^{\ell}\right) \exp\left[\left(-\sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^{\ell}\right) + j\nu\xi_i\right]}{j\nu - \rho_i + \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^{\ell}} \right\},$$

where the extra term is eliminated since $\xi_0 = -\infty$ and $s_0^{\ell} = 1 \forall \ell \geq 0$.

Combining the two manipulated terms yields

$$\begin{aligned} I &= \sum_{i=1}^{n} \left[\frac{g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell \in \sigma_j} s_{i-1}^{\ell}\right)}{j\nu + \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^\ell} - \frac{g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell \in \sigma_j} s_{i}^{\ell}\right)}{j\nu - \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^\ell} \right] \exp\left(-\sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^\ell + j\nu\xi_i \right) \\ &= \sum_{i=1}^{n} G_i \exp\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^\ell + j\nu\xi_i \right) \end{aligned} \tag{7.59}$$

where

$$G_i = \frac{g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell \in \sigma_j} s_{i-1}^{\ell}\right)}{j\nu + \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^\ell} - \frac{g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell \in \sigma_j} s_{i}^{\ell}\right)}{j\nu - \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^\ell} \tag{7.60}$$

Alternatively, we can express G_i as

$$\begin{aligned} G_i &= \frac{g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell \in \sigma_j} s_{i}^{\ell}\right)\Bigg|_{s_i^i=1}}{j\nu + \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^\ell} - \frac{g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell \in \sigma_j} s_{i}^{\ell}\right)}{j\nu - \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell s_i^\ell} \\ &= \left[\frac{g(\xi_i)}{j\nu + \sum\limits_{\ell=1}^{n} \rho_\ell s_i^\ell} \right]_{s_i^i=1} - \frac{g(\xi_i)}{j\nu + \sum\limits_{\ell=1}^{n} \rho_\ell s_i^\ell} \end{aligned} \tag{7.61}$$

We do not simplify this further, because in practice, it is easier to leave s_i^ℓ as defined in (7.53) rather than explicitly using sign functions. This is due to the complication that not all terms in $g(\eta)$ contains s_i^i, whose sign flips depending on which fraction of G_i it is in. Note that when $\xi_\ell = \xi_i$, the term in the exponent is zero, so

$$\exp\left(-\sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^\ell + j\nu\xi_i \right) = \exp\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^\ell + j\nu\xi_i \right). \tag{7.62}$$

Instead of the first form, used in Appendix B, I use the second form to match G_i.

Scalar Example: The generalized integral formula (7.59) was demonstrated using three sign functions. For $x \in \mathbb{R}$, let

$$f(x) = g(x) \exp\left(-\sum_{i=1}^{3} \rho_i |\xi_i - x|\right) \tag{7.63}$$

where

$$g = s_1 + s_2 + s_3 + s_1 s_2 + s_1 s_3 + s_2 s_3 + s_1 s_2 s_3 \tag{7.64}$$

and $s_i = \text{sgn}(\xi_i - x)$. For

$$\begin{aligned} \xi_1 &= -1 \quad \rho_1 = 0.5 \\ \xi_2 &= 2 \qquad \rho_2 = 0.2 \\ \xi_3 &= 3 \qquad \rho_3 = 0.1 \end{aligned} \tag{7.65}$$

we numerically integrated $f(x)$ from -10 to 10 using a Riemann sum with $dx = 0.0001$, which resulted in

$$I_{\text{numerical}} = 2.0987. \tag{7.66}$$

Using (7.59), we got

$$I_{\text{generalized}} = 2.0986. \tag{7.67}$$

Note that $g(x)$ could have been simplified to a simple sum, but we left it in the longer form to show the flexibility of (7.59).

2D Example: For $\boldsymbol{x} \in \mathbb{R}^2$, let

$$\begin{aligned} s_1 &= \text{sgn}(1 + x_1 + x_2) = -\text{sgn}(\xi_1 - x_2) \\ s_2 &= \text{sgn}(1 - x_1 + x_2) = -\text{sgn}(\xi_2 - x_2) \end{aligned} \tag{7.68}$$

where $\xi_1 = -1 - x_1$ and $\xi_2 = -1 + x_1$. Then, consider the integral where we integrate with respect to x_2

$$f(x_1) = \int_{-\infty}^{\infty} s_1 s_2 \exp(-2|\xi_1 - x_2| - 3|\xi_2 - x_2|)\, dx_2 \tag{7.69}$$

Using the generalized integral formula in (7.59),

$$\begin{aligned} f(x_1) &= G_1(x_1) \exp(-3|\xi_2 - \xi_1|) + G_2(x_1) \exp(-2|\xi_1 - \xi_2|) \\ &= G_1(x_1) \exp(-3|2x_1|) + G_2(x_1) \exp(-2|2x_1|) \end{aligned} \tag{7.70}$$

where

$$
\begin{aligned}
G_1(x_1) &= \frac{\operatorname{sgn}(\xi_2-\xi_1)}{2+3\operatorname{sgn}(\xi_2-\xi_1)} - \frac{-\operatorname{sgn}(\xi_2-\xi_1)}{-2+3\operatorname{sgn}(\xi_2-\xi_1)} \\
&= \frac{\operatorname{sgn}(2x_1)}{2+3\operatorname{sgn}(2x_1)} + \frac{\operatorname{sgn}(2x_1)}{-2+3\operatorname{sgn}(2x_1)} \\
&= \frac{\operatorname{sgn}(x_1)}{2+3\operatorname{sgn}(x_1)} + \frac{\operatorname{sgn}(x_1)}{-2+3\operatorname{sgn}(x_1)} \\
G_2(x_2) &= \frac{\operatorname{sgn}(\xi_1-\xi_2)}{2+2\operatorname{sgn}(\xi_1-\xi_2)} - \frac{-\operatorname{sgn}(\xi_1-\xi_2)}{-2+2\operatorname{sgn}(\xi_1-\xi_2)} \\
&= \frac{\operatorname{sgn}(-2x_1)}{3+2\operatorname{sgn}(-2x_1)} - \frac{-\operatorname{sgn}(-2x_1)}{-3+2\operatorname{sgn}(-2x_1)} \\
&= \frac{-\operatorname{sgn}(x_1)}{3-2\operatorname{sgn}(x_1)} - \frac{\operatorname{sgn}(x_1)}{-3-2\operatorname{sgn}(x_1)}
\end{aligned}
\tag{7.71}
$$

The integral (7.69) was solved numerically using a Riemann sum and plotted against the closed-form solution (7.70) in Figure 7.4.

7.8.1 Re-writing output of generalized integral formula in standard form

The solution to the generalized integral formula in (7.59) is compact, but in order to facilitate incorporating it into an algorithm, it needs to be rewritten in standard form. That is, the arguments of s_i^l must be explicitly written in the form $\psi + \theta^T \boldsymbol{x}$.

We start with one term of the output of (7.59)

$$
I_i = G_i \exp\left(-\sum_{\ell=1}^{n} \rho_\ell (\xi_\ell - \xi_i) s_i^\ell + j\nu\xi_i \right) \tag{7.72}
$$

where

$$
G_i = \frac{ g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell\in\sigma_j} s_i^\ell \right) \Bigg|_{s_i^i=1} }{ j\nu + \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell\neq i}}^{n} \rho_\ell s_i^\ell } - \frac{ g\left(\sum\limits_{j=1}^{m} \varrho_j \prod\limits_{\ell\in\sigma_j} s_i^\ell \right) }{ j\nu - \rho_i + \sum\limits_{\substack{\ell=1 \\ \ell\neq i}}^{n} \rho_\ell s_i^\ell } = \left[\frac{g}{j\nu + \sum\limits_{\ell=1}^{n} \rho_\ell s_i^\ell} \right]_{s_i^i=1} - \frac{g}{j\nu + \sum\limits_{\ell=1}^{n} \rho_\ell s_i^\ell}
\tag{7.73}
$$

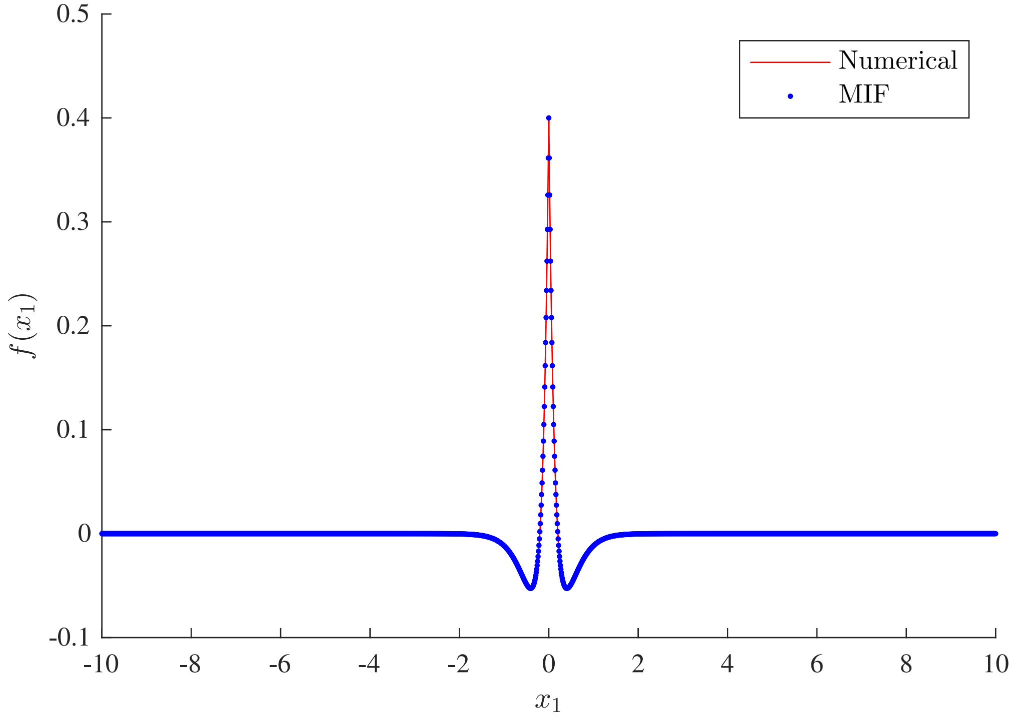


Figure 7.4: Numerical and generalized (MIF) closed-form solutions to integral formula in two dimensions

and g is defined in (7.52). Nominally, s_i^l is defined by (7.53), or

$$\begin{aligned} s_i^l &= \operatorname{sgn}(\xi_l - \xi_i) \\ &= \operatorname{sgn}\left(\psi_l + \theta_l^T \boldsymbol{x} - \psi_i - \theta_i^T \boldsymbol{x}\right) \\ &= \operatorname{sgn}\left(\psi_l - \psi_i + (\theta_l - \theta_i)^T \boldsymbol{x}\right) \\ &\triangleq \operatorname{sgn}\left(\psi_{il} + \theta_{il}^T \boldsymbol{x}\right) \end{aligned} \tag{7.74}$$

7.9 Collapsed hyperplanes

When the θ component of hyperplane $\xi = \psi + \theta \boldsymbol{x}$ vanishes, the hyperplane collapses into a constant and must be factored out from the standard function expression.

7.9.1 Example: Collapsed hyperplane

It's easiest to start with an example. Given a pdf term in standard form such that

$$\begin{aligned} f_i(\boldsymbol{x}) &= g(\boldsymbol{x}) \exp\left(-\sum_{j=1}^{3} \eta_j \left|\psi_j + \theta_j^T \boldsymbol{x}\right|\right) \\ g(\boldsymbol{x}) &= \rho_0 + \sum_{j=1}^{6} \rho_j \prod_{l \in \sigma_j} \operatorname{sgn}\left(\psi_l + \theta_l^T \boldsymbol{x}\right) \end{aligned} \tag{7.75}$$

where

$$\begin{aligned} \psi &= \begin{bmatrix} \psi_1 & \psi_2 & \psi_3 \end{bmatrix} \\ \theta &= \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 \end{bmatrix} \\ \eta &= \begin{bmatrix} \eta_1 & \eta_2 & \eta_3 \end{bmatrix} \\ \rho &= \begin{bmatrix} \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 & \rho_6 \end{bmatrix} \\ \sigma &= \{\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\} \end{aligned} \tag{7.76}$$

Let $\theta_2 = 0$. The exponential function becomes

$$\exp\left(-\eta_2 \left|\psi_2\right|\right) \cdot \exp\left(-\eta_1 \left|\psi_1 + \theta_1 \boldsymbol{x}\right| - \eta_3 \left|\psi_3 + \theta_3 \boldsymbol{x}\right|\right) \tag{7.77}$$

The constant factor $\zeta = \exp\left(-\eta_2 \left|\psi_2\right|\right)$ will be distributed into the coefficient function. In the coefficient function, term 2 also becomes constant, while terms 4 and 6 simply factor out constant coefficients as

$$\begin{aligned} \rho_0 &\to \rho_0 + \rho_2 \operatorname{sgn}\left(\psi_2\right) \\ \rho_2 &\to 0 \\ \rho_4 &\to \rho_4 \operatorname{sgn}\left(\psi_2\right) \\ \rho_6 &\to \rho_6 \operatorname{sgn}\left(\psi_2\right) \end{aligned} \tag{7.78}$$

and

$$\sigma \to \{\{1\}, \emptyset, \{3\}, \{1\}, \{1,3\} \{3\}\} \tag{7.79}$$

Terms 1 and 3 now have the same elements as terms 3 and 6, respectively, so they can be combined as

$$\begin{aligned}\bar{\rho}_0 &= \zeta\left(\rho_0 + \rho_2 \operatorname{sgn}(\psi_2)\right)\\ \bar{\rho}_1 &= \zeta\left(\rho_1 + \rho_4 \operatorname{sgn}(\psi_2)\right)\\ \bar{\rho}_3 &= \zeta\left(\rho_3 + \rho_6 \operatorname{sgn}(\psi_2)\right)\\ \bar{\rho}_5 &= \zeta \rho_5\end{aligned} \tag{7.80}$$

and

$$\sigma \to \{\{1\},\{3\},\{1,3\}\} \tag{7.81}$$

Finally, we can re-index all the terms to get

$$\begin{aligned}\hat{f}_i(\boldsymbol{x}) &= \hat{g}(\boldsymbol{x}) \exp\left(-\sum_{j=1}^{2} \hat{\eta}_j \left|\hat{\psi}_j + \hat{\theta}_j \boldsymbol{x}\right|\right)\\ \hat{g}(\boldsymbol{x}) &= \hat{\rho}_0 + \sum_{j=1}^{2} \hat{\rho}_j \prod_{l\in\sigma_j} \operatorname{sgn}\left(\hat{\psi}_l + \hat{\theta}_l \boldsymbol{x}\right)\end{aligned} \tag{7.82}$$

where

$$\begin{aligned}\hat{\psi} &= \begin{bmatrix} \hat{\psi}_1 & \hat{\psi}_2 \end{bmatrix} = \begin{bmatrix} \psi_1 & \hat{\psi}_3 \end{bmatrix}\\ \hat{\theta} &= \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_3 \end{bmatrix}\\ \hat{\eta} &= \begin{bmatrix} \hat{\eta}_1 & \hat{\eta}_2 \end{bmatrix} = \begin{bmatrix} \hat{\eta}_1 & \hat{\eta}_3 \end{bmatrix}\\ \hat{\rho}_0 &= \zeta\left(\rho_0 + \rho_2 \operatorname{sgn}(\psi_2)\right)\\ \hat{\rho} &= \begin{bmatrix} \hat{\rho}_1 & \hat{\rho}_2 & \hat{\rho}_3 \end{bmatrix} = \zeta \begin{bmatrix} \rho_1 + \rho_4 \operatorname{sgn}(\psi_2) & \rho_3 + \rho_6 \operatorname{sgn}(\psi_2) & \rho_5 \end{bmatrix}\\ \hat{\sigma} &= \{\{1\},\{2\},\{1,2\}\}\end{aligned} \tag{7.83}$$

7.10 Merging hyperplanes

When two hyperplanes coalesce into one, it is necessary to either merge them into one or to otherwise keep track of them. This is due to the fact that certain naive implementations of the integral formulas may result in incorrect sign of certain terms.

First, we define two hyperplanes ξ_i and ξ_j as equal when, for

$$\xi_i = \psi_i + \theta_i \boldsymbol{x}, \quad \xi_j = \psi_j + \theta_j \boldsymbol{x}, \tag{7.84}$$

the equation

$$A \triangleq \begin{bmatrix} \theta_i \\ \theta_j \end{bmatrix} \boldsymbol{x} = 0 \tag{7.85}$$

has no non-trivial solutions and

$$\frac{\psi_i}{\|\theta_i\|_2} - \frac{\psi_j}{\|\theta_j\|_2} = 0 \tag{7.86}$$

Equivalently, A has a non-trivial nullspace, or the directions defined by θ_i and θ_j are parallel. When $\psi_i = \psi_j = 0$, θ_i and θ_j may be anti-parallel and still satisfy the above conditions, so the hyperplanes must still be merged, but the signs are opposite.

In practice, one is limited to machine precision and must decide whether that is actually too small. Perhaps a higher threshold is needed to coalesce hyperplanes which are "close enough", though care must be taken to avoid overdoing it and corrupting the expressions for the pdf or characteristic functions.

7.10.1 Example: Merging hyperplanes when computing characteristic function

It's easiest to start with an example. Given a pdf term in standard form such that

$$\begin{aligned} f_i(\boldsymbol{x}) &= g(\boldsymbol{x}) \exp\left(-\sum_{j=1}^{3} \eta_j \left|\psi_j + \theta_j^T \boldsymbol{x}\right|\right) \\ g(\boldsymbol{x}) &= \rho_0 + \sum_{j=1}^{6} \rho_j \prod_{l \in \sigma_j} \operatorname{sgn}\left(\psi_l + \theta_l^T \boldsymbol{x}\right) \end{aligned} \tag{7.87}$$

where

$$\begin{aligned}
\psi &= \left[\begin{array}{ccc} \psi_1 & \psi_2 & \psi_3 \end{array} \right] \\
\theta &= \left[\begin{array}{ccc} \theta_1 & \theta_2 & \theta_3 \end{array} \right] \\
\eta &= \left[\begin{array}{ccc} \eta_1 & \eta_2 & \eta_3 \end{array} \right] \\
\rho &= \left[\begin{array}{cccccc} \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 & \rho_6 \end{array} \right] \\
\sigma &= \{\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}\}
\end{aligned} \tag{7.88}$$

Let hyperplanes 1 and 2 be coincident and are equal to within a scale factor ζ so that

$$\psi_1 + \theta_1^T \boldsymbol{x} = \zeta\left(\psi_2 + \theta_2^T \boldsymbol{x}\right) \implies \frac{\psi_1}{\zeta} + \frac{\theta_1^T}{\zeta}\boldsymbol{x} = \psi_2 + \theta_2^T \boldsymbol{x} \tag{7.89}$$

$$\tag{7.90}$$

Replacing hyperplane 2 with hyperplane 1 results in the following changes

$$\begin{aligned}
\eta_2\left|\psi_2 + \theta_2^T \boldsymbol{x}\right| &= \eta_2\left|\frac{\psi_1}{\zeta} + \frac{\theta_1^T}{\zeta}\boldsymbol{x}\right| \\
&= \frac{\eta_2}{|\zeta|}\left|\psi_1 + \theta_1^T \boldsymbol{x}\right| \\
\rho_2 \mathrm{sgn}\left(\psi_2 + \theta_2^T \boldsymbol{x}\right) &= \rho_2 \mathrm{sgn}\left(\frac{\psi_1}{\zeta} + \frac{\theta_1^T}{\zeta}\boldsymbol{x}\right) \\
&= \rho_2 \mathrm{sgn}\left(\zeta\right) \mathrm{sgn}\left(\psi_1 + \theta_1^T \boldsymbol{x}\right) \\
\rho_4 \mathrm{sgn}\left(\psi_1 + \theta_1^T \boldsymbol{x}\right) \mathrm{sgn}\left(\psi_2 + \theta_2^T \boldsymbol{x}\right) &= \rho_4 \mathrm{sgn}\left(\psi_1 + \theta_1^T \boldsymbol{x}\right) \mathrm{sgn}\left(\frac{\psi_1}{\zeta} + \frac{\theta_1^T}{\zeta}\boldsymbol{x}\right) \\
&= \rho_4 \mathrm{sgn}\left(\zeta\right) \mathrm{sgn}\left(\psi_1 + \theta_1^T \boldsymbol{x}\right)^2 \\
&= \rho_4 \mathrm{sgn}\left(\zeta\right) \\
\rho_6 \mathrm{sgn}\left(\psi_2 + \theta_2^T \boldsymbol{x}\right) \mathrm{sgn}\left(\psi_3 + \theta_3^T \boldsymbol{x}\right) &= \rho_6 \mathrm{sgn}\left(\frac{\psi_1}{\zeta} + \frac{\theta_1^T}{\zeta}\boldsymbol{x}\right) \mathrm{sgn}\left(\psi_3 + \theta_3^T \boldsymbol{x}\right) \\
&= \rho_6 \mathrm{sgn}\left(\zeta\right) \mathrm{sgn}\left(\psi_1 + \theta_1^T \boldsymbol{x}\right) \mathrm{sgn}\left(\psi_3 + \theta_3^T \boldsymbol{x}\right)
\end{aligned} \tag{7.91}$$

Note that $\text{sgn}(\zeta)$ almost always equals 1, since the only time two hyperplanes coincide when $\theta_1^T\theta_2 < 0$ is when $\psi_1 = \psi_2 = 0$. The parameters then become

$$\begin{aligned}
\bar{\psi} &= \begin{bmatrix} \psi_1 & \psi_1 & \psi_3 \end{bmatrix} \\
\bar{\theta} &= \begin{bmatrix} \theta_1 & \theta_1 & \theta_3 \end{bmatrix} \\
\bar{\eta} &= \begin{bmatrix} \eta_1 & \frac{\eta_2}{|\zeta|} & \eta_3 \end{bmatrix} \\
\bar{\rho}_0 &= \rho_0 + \rho_4 \text{sgn}(\zeta) \\
\bar{\rho} &= \begin{bmatrix} \rho_1 & \rho_2 \text{sgn}(\zeta) & \rho_3 & \rho_5 & \rho_6 \text{sgn}(\zeta) \end{bmatrix} \\
\bar{\sigma} &= \{\{1\},\{1\},\{3\},\{1,3\},\{1,3\}\}
\end{aligned} \tag{7.92}$$

Combining like terms and shifting the old index 3 in σ to the vacated index 2 yields

$$\begin{aligned}
\hat{\psi} &= \begin{bmatrix} \psi_1 & \psi_3 \end{bmatrix} \\
\hat{\theta} &= \begin{bmatrix} \theta_1 & \theta_3 \end{bmatrix} \\
\hat{\eta} &= \begin{bmatrix} \eta_1 + \frac{\eta_2}{|\zeta|} & \eta_3 \end{bmatrix} \\
\hat{\rho}_0 &= \rho_0 + \rho_4 \text{sgn}(\zeta) \\
\hat{\rho} &= \begin{bmatrix} \rho_1 + \rho_2 \text{sgn}(\zeta) & \rho_3 & \rho_5 + \rho_6 \text{sgn}(\zeta) \end{bmatrix} \\
\hat{\sigma} &= \{\{1\},\{2\},\{1,2\}\}
\end{aligned} \tag{7.93}$$

To reiterate, the indices in $\hat{\sigma}$ reflect the positions of elements of $\hat{\psi}$ and $\hat{\theta}$.

7.11 Algorithm for computing characteristic function in $\mathbb{R}^2$

Since it is not practical to show the characteristic function(CF) for the general case in $\mathbb{R}^n$, we will instead provide an algorithm for computing the CF in $\mathbb{R}^2$. For each term in $\phi_{X|Y}$, we have

$$
\begin{aligned}
\phi^i_{X|Y}(\nu_1,\nu_2) &= \int_{\mathbb{R}} e^{j\nu_1 x_1} \int_{\mathbb{R}} f^i_{X|Y}(\boldsymbol{x}) e^{j\nu_2 x_2}\, dx_2\, dx_1 \\
&= \int_{\mathbb{R}} e^{j\nu_1 x_1} \int_{\mathbb{R}} g_i(\boldsymbol{x}) \exp\left(-\sum_{l=1}^{m_i} \eta_l \left|\psi_l + \theta_l \boldsymbol{x}\right| + j\nu_2 x_2 \right) dx_2\, dx_1 \\
&\triangleq \int_{\mathbb{R}} e^{j\nu_1 x_1} I^i_2(\nu_2, x_1)\, dx_1
\end{aligned} \tag{7.94}
$$

7.11.1 Isolate x_2

We first evaluate the inner integral $I^i_2(\nu, x_1)$. Similarly to the propagation step, we multiply $f^i_{X|Y}$ by an exponential, $\exp(j\nu_2 x_2)$, and manipulate the argument $\psi_l + \theta_l^T \boldsymbol{x}$ to look like $\xi_l - x_2$.

For $\theta_{l2} \neq 0$,

$$
\begin{aligned}
\tilde{\rho}_q &= \rho_q \prod_{l \in \sigma_q} -\mathrm{sgn}\left(\theta_{l2}\right) \\
\tilde{\xi}_l &= -\frac{\psi_l + \theta_{l1} x_1}{\theta_{l2}} \\
\tilde{\eta}_l &= \eta_l \left|\theta_{l2}\right|
\end{aligned} \tag{7.95}
$$

Note that x_1 is actually implicit, so when we write this algorithm in code, we have to implicitly carry it forward.

For $\theta_{l2} = 0$, the particular element is constant with respect to x_2, so it is not considered a hyperplane for this integral. Therefore, if $\sigma_q \setminus l \neq \emptyset$,

$$
\begin{aligned}
\tilde{\sigma}_q &= \sigma_q \setminus l \\
\tilde{\rho}_q &= \rho_q \mathrm{sgn}\left(\psi_l - \theta_{l1} x_1\right).
\end{aligned} \tag{7.96}
$$

This moves the constant exponential $\exp\left(-\eta_l \left|\psi_l - \theta_{l1} x_1\right|\right)$ out of the integral to join $\exp(j\nu_1 x_1)$. In $g_i(\boldsymbol{x})$, $\mathrm{sgn}\left(\psi_l - \theta_{l1} x_1\right)$ moves out to join ρ_q as a constant coefficient.

If $\sigma_{iq} \setminus l = \emptyset$, the entire q term in $g_i(\boldsymbol{x})$ is constant with respect to x_2, and it is then grouped with the constant coefficient ρ_{i0} as

$$\tilde{\rho}_{i0} = \rho_{i0+}\rho_{il}\operatorname{sgn}(\psi_{il} - \theta_{i,l1}x_1) \tag{7.97}$$

These new parameters now define term i in a new form

$$\tilde{f}^i_{X|Y} = \tilde{g}_i(\tilde{\xi}_l - x_2)\exp\left(-\sum_{l=1}^{\tilde{m}_i}\tilde{\eta}_l\left|\tilde{\xi}_l - x_2\right| + j\nu_2 x_2\right) \tag{7.98}$$

7.11.2 Integration with respect to x_2

After the conditioning in the first stage, we can use the generalized integral formula (7.59) to integrate with respect to x_2 to get

$$I^i_2(\nu_2, x_1) = \sum_{j=1}^{\tilde{m}_i} G_{ij}(\nu_2, x_1)\exp\left(-\sum_{\substack{l=1\\ l\neq j}}^{\tilde{m}_{ij}+\bar{m}_{ij}}\tilde{\eta}_{il}\left|\tilde{\xi}_{il} - \tilde{\xi}_{ij}\right| + j\nu_2\tilde{\xi}_{ij}\right) \tag{7.99}$$

where

$$\begin{aligned} G_{ij}(\nu_2, x_1) = & \frac{\tilde{g}^{\dagger}_i\left(\tilde{\xi}_{ij}\right)}{j\nu_2 + \eta_{ij} + \sum\limits_{\substack{l=1\\ l\neq j}}^{\tilde{m}_{ij}}\eta_{il}\operatorname{sgn}\left(\tilde{\xi}_{il} - \tilde{\xi}_{ij}\right)} \\ & - \frac{\tilde{g}_i\left(\tilde{\xi}_{ij}\right)}{j\nu_2 - \eta_{ij} + \sum\limits_{\substack{l=1\\ l\neq j}}^{\tilde{m}_{ij}}\eta_{il}\operatorname{sgn}\left(\tilde{\xi}_{il} - \tilde{\xi}_{ij}\right)} \end{aligned} \tag{7.100}$$

$g^{\dagger}\left(\tilde{\xi}_{ij}\right)$ refers to $\tilde{g}_i(x_1)$ for which $\operatorname{sgn}(0)$ is evaluated as 1 instead of the usual -1, $\bar{m}_{ij}$ accounts for any constant terms that were factored out of the first integration, and

$$\begin{aligned} \xi_{il} - \xi_{ij} &= (\psi_{il} - \psi_{ij}) - (\theta_{il} - \theta_{ij})\,x_1 \\ &\triangleq \psi_{ijl} + \theta_{ijl}x_1 \end{aligned} \tag{7.101}$$

Unfortunately, we will not be able to flatten the resulting $G_{ij}(\nu_2, x_1)$, since it has non-constant ν_2 which are not in sign functions.

7.11.3 Re-index terms

Since each term i produces $\tilde{m}_i$ terms, $\phi_{X|Y}$ is a sum of sum of terms, whose indexing can be cumbersome to manage. Therefore, we can re-index all of the sub-terms from the first integration so we again have a sum of the form

$$\begin{aligned}\phi_{X|Y}(\nu_1,\nu_2) &= \int_{\mathbb{R}} j\nu x_1 \left[\sum_{i=1}^{N} G_i(\nu_2,x_1)\exp\left(-\sum_{j=1}^{m_i}\tilde{\eta}_{il}\left|\psi_{il}+\theta_{il}x_1\right| + j\nu_2\xi_i\right)\right] dx_1 \\ &= \sum_{i=1}^{N}\left[\int_{\mathbb{R}} G_i(\nu_2,x_1)\exp\left(-\sum_{j=1}^{m_i}\eta_{il}\left|\psi_{il}+\theta_{il}x_1\right| + j\nu_1 x_1 + j\nu_2\xi_i\right) dx_1\right]\end{aligned} \tag{7.102}$$

where the previous ij indices now become the new i indices, and N now accounts for all the sub-terms. Note that we also remove the tilde.

7.11.4 Isolate x_1

We can integrate each term separately once more to obtain a term of $\phi_{X|Y}$. As in the first integral, we must isolate x_1 in order to use the integral formula. For $\theta_{il} \neq 0$,

$$\begin{aligned}\hat{\xi}_{il} &= \frac{\psi_{il}}{-\theta_{il}} \\ \hat{\eta}_{il} &= \eta_{il}\left|\theta_{il}\right| \\ \hat{\rho}_{il} &= \rho_{il}\operatorname{sgn}\left(-\theta_{il}\right)\end{aligned} \tag{7.103}$$

When $\theta_{il} = 0$, the corresponding constant exponential $\exp\left(-\eta_{il}\left|\psi_{il}\right|\right)$ is factored out of the integration and the corresponding constant $\operatorname{sgn}\left(\psi_{il}\right)$ is accounted for in G_i similarly to the first integration. Note that this will affect both the numerator and denominator of G_i, though the denominator is straight-forward because it will not have products of signs to deal with.

Equation (7.102) then looks like

$$\phi_{X|Y}(\nu_1,\nu_2) = \sum_{i=1}^{N}\left[e^{j\nu_2\psi_i}\int_{\mathbb{R}} G_i(\nu_2,x_1)\exp\left(-\sum_{j=1}^{m_i}\hat{\eta}_{il}\left|\hat{\xi}_{il}-x_1\right| + j\nu_1 x_1 + j\nu_2\theta_i x_1\right) dx_1\right] \tag{7.104}$$

Note that $\exp(j\nu_2\psi_i)$ and $\exp(\nu_2\theta_i x_1)$ both came from $\tilde{\xi}_{ij}$ in (7.99).

7.11.5 Integration with respect to x_1

Finally, we can use the integral formula to integrate each term of $\phi_{X|Y}$ to get

$$\phi^i_{X|Y} = \sum_{j=1}^{m_i} \hat{G}_{ij}(\nu_1, \nu_2) \exp\left(-\sum_{\substack{l=1 \\ l\neq j}}^{m_{ij}} \hat{\eta}_{il} \left| \hat{\xi}_{il} - \hat{\xi}_{ij} \right| + j\nu_1 \hat{\xi}_{ij} + j\nu_2 \left(\psi_i + \theta_i \hat{\xi}_{ij} \right) \right) \tag{7.105}$$

where

$$\begin{aligned} \hat{G}_{ij}(\nu_1, \nu_2) = & \frac{G_i^*\left(\nu_2, \hat{\xi}_{ij}\right)}{j\nu_1 + j\nu_2\theta_i + \hat{\eta}_{ij} + \sum\limits_{\substack{l=1 \\ l\neq j}} \hat{\eta}_{il} \mathrm{sgn}\left(\hat{\xi}_{il} - \hat{\xi}_{ij}\right)} \\ & - \frac{G_i\left(\nu_2, \hat{\xi}_{ij}\right)}{j\nu_1 + j\nu_2\theta_i - \hat{\eta}_{ij} + \sum\limits_{\substack{l=1 \\ l\neq j}} \hat{\eta}_{il} \mathrm{sgn}\left(\hat{\xi}_{il} - \hat{\xi}_{ij}\right)} \end{aligned} \tag{7.106}$$

and $G^*(\nu_2, \hat{\xi}_{ij})$ denotes $G_i(\nu_2, x_1)$ evaluated at $\hat{\xi}_{ij}$ where $\mathrm{sgn}(0)$ is evaluated to 1 instead of the normal -1. $G_i(\nu_2, x_1)$ is the re-indexed $G_{ij}(\nu_2, x_1)$ defined in (7.100).

7.11.6 Simplify characteristic function

After the second integration (the outer integration), the characteristic function has the form

$$\phi_{X|Y} = \sum_{i=1}^{N} \hat{G}_i(\nu_1, \nu_2) \exp\left(c_{i1} \cdot j\nu_1 + c_{i2} \cdot j\nu_2 + c_{i3}\right) \tag{7.107}$$

where

$$\hat{G}_i(\nu_1, \nu_2) = \frac{\frac{a_{i1}}{a_{i2}+j\nu_2} - \frac{a_{i3}}{a_{i4}+j\nu_2}}{a_{i5} + j\nu_1 + a_{i6} \cdot j\nu_2} - \frac{\frac{b_{i1}}{b_{i2}+j\nu_2} - \frac{b_{i3}}{b_{i4}+j\nu_2}}{b_{i5} + j\nu_1 + b_{i6} \cdot j\nu_2} \tag{7.108}$$

and $a_{i1}, \cdots, a_{i6}$, $b_{i1}, \cdots, b_{i6}$ and $c_{i1}, \cdots, c_{i3}$ are constants.

7.12 Example: Normalization of $f_{X_1|Y_1}$ in $\mathbb{R}^2$

Given independent initial densities $f_{X_{11}}$ and $f_{X_{12}}$ with means $\bar{x}_1, \bar{x}_2$ and spread parameter α, the initial pdf is

$$\begin{aligned} f_{X_1}(\boldsymbol{x}_1) &= \frac{1}{2\alpha}\exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right)\frac{1}{2\alpha}\exp\left(-\frac{1}{\alpha}\left|\bar{x}_2 - x_2\right|\right) \\ &= \frac{1}{4\alpha^2}\exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - E_1\boldsymbol{x}_1\right| - \frac{1}{\alpha}\left|\bar{x}_2 - E_2\boldsymbol{x}_1\right|\right) \end{aligned} \tag{7.109}$$

where for $E_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $E_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$,

Given measurement $Y_1 = \{z_1\}$ and using (3.17), where $H = \begin{bmatrix} h_1 & h_2 \end{bmatrix}$, $h_1 \neq 0, h_2 \neq 0$, the unnormalized conditional pdf, $f_{X_1|Y_1}$, is given by

$$f_{X_1|Y_1}(\boldsymbol{x}_1|z_1) = \frac{1}{2\gamma}\exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - E_1\boldsymbol{x}_1\right| - \frac{1}{\alpha}\left|\bar{x}_2 - E_2\boldsymbol{x}_1\right| - \frac{1}{\gamma}\left|z_1 - H\boldsymbol{x}_1\right|\right) \tag{7.110}$$

We can ignore the $\frac{1}{4\alpha^2}$ and $\frac{1}{2\gamma}$ because it will get taken care of when we normalize. The normalization factor of (7.110) is given by

$$\begin{aligned} f_{Y_1} &= \iint_{\mathbb{R}^2} f_{X_1|Y_1}(\boldsymbol{x}_1)\, d\boldsymbol{x} \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - E_1\boldsymbol{x}_1\right| - \frac{1}{\alpha}\left|\bar{x}_2 - E_2\boldsymbol{x}_1\right| - \frac{1}{\gamma}\left|z_1 - H\boldsymbol{x}_1\right|\right) dx_2\, dx_1 \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - x_1\right| - \frac{1}{\alpha}\left|\bar{x}_2 - x_2\right| - \frac{1}{\gamma}\left|z_1 - h_1x_1 - h_2x_2\right|\right) dx_2\, dx_1 \\ &= \int_{\mathbb{R}} \exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right)\int_{\mathbb{R}} \exp\left(-\frac{1}{\alpha}\left|\bar{x}_2 - x_2\right| - \frac{1}{\gamma}\left|z_1 - h_1x_1 - h_2x_2\right|\right) dx_2\, dx_1 \\ &\triangleq \int_{\mathbb{R}} I_2(x_1)\exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) dx_1 \end{aligned} \tag{7.111}$$

Assuming $h_2 \neq 0$, we isolate x_2 of the inner integral to get

$$\begin{aligned} I_2(x_1) &= \int_{\mathbb{R}} \exp\left(-\frac{1}{\alpha}\left|\bar{x}_2 - x_2\right| - \frac{1}{\gamma}\left|z_1 - h_1x_1 - h_2x_2\right|\right) dx_2 \\ &= \int_{\mathbb{R}} \exp\left(-\frac{1}{\alpha}\left|\bar{x}_2 - x_2\right| - \frac{|h_2|}{\gamma}\left|\frac{z_1}{h_2} - \frac{h_1}{h_2}x_1 - x_2\right|\right) dx_2 \\ &\triangleq \int_{\mathbb{R}} \exp\left(-\eta_1\left|\xi_1 - x_2\right| - \eta_2\left|\xi_2 - x_2\right|\right) dx_2 \end{aligned} \tag{7.112}$$

where

$$\begin{aligned}\eta_1 &= \frac{1}{\alpha}\\ \xi_1 &= \bar{x}_2 + 0 \cdot x_1 \triangleq \psi_1 + \theta_1 x_1\\ \eta_2 &= \frac{|h_2|}{\gamma}\\ \xi_2 &= \frac{z_1}{h_2} + \frac{-h_1}{h_2} x_1 \triangleq \psi_2 + \theta_2 x_1\end{aligned} \tag{7.113}$$

Using the generalized integral formula (7.59), we obtain

$$\begin{aligned}I_2(x_1) &= g_1(x_1) \exp\left(-\eta_2 \left|\xi_2 - \xi_1\right|\right) + g_2(x_1) \exp\left(-\eta_1 \left|\xi_1 - \xi_2\right|\right)\\ &\triangleq I_{21}(x_1) + I_{22}(x_1)\end{aligned} \tag{7.114}$$

where

$$\begin{aligned}g_1(x_1) &= \frac{1}{j\nu_2 + \eta_1 + \eta_2 \operatorname{sgn}\left(\xi_2 - \xi_1\right)} - \frac{1}{j\nu_2 - \eta_1 + \eta_2 \operatorname{sgn}\left(\xi_2 - \xi_1\right)}\\ g_2(x_1) &= \frac{1}{j\nu_2 + \eta_2 + \eta_1 \operatorname{sgn}\left(\xi_1 - \xi_2\right)} - \frac{1}{j\nu_2 - \eta_2 + \eta_1 \operatorname{sgn}\left(\xi_1 - \xi_2\right)}\end{aligned} \tag{7.115}$$

7.12.1 Integrate term 1 with respect to x_1

Let's expand $I_{21}(x_1)$ and perform the second integration separately from $I_{22}(x_1)$.

$$\begin{aligned}I_{11} &= \int_{\mathbb{R}} I_{21}(x_1) \exp\left(-\frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) dx_1\\ &= \int_{\mathbb{R}} g_1(x_1) \exp\left(-\eta_2 \left|\xi_2 - \xi_1\right| - \frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) dx_1\\ &= \int_{\mathbb{R}} g_1(x_1) \exp\left(-\eta_2 \left|(\psi_2 - \psi_1) + (\theta_2 - \theta_1)\, x_1\right| - \frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) dx_1\\ &= \int_{\mathbb{R}} g_1(x_1) \exp\left(-\eta_2 \left|(\psi_2 - \psi_1) + \theta_2 x_1\right| - \frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) dx_1\end{aligned} \tag{7.116}$$

where

$$\begin{aligned} g_1(x_1) &= \frac{1}{\eta_1 + \eta_2 \text{sgn}\left(\xi_2 - \xi_1\right)} - \frac{1}{-\eta_1 + \eta_2 \text{sgn}\left(\xi_2 - \xi_1\right)} \\ &= \frac{1}{\eta_1 + \eta_2 \text{sgn}\left(\left(\psi_2 - \psi_1\right) + \left(\theta_2 - \theta_1\right) x_1\right)} - \frac{1}{-\eta_1 + \eta_2 \text{sgn}\left(\left(\psi_2 - \psi_1\right) + \left(\theta_2 - \theta_1\right) x_1\right)} \\ &= \frac{1}{\eta_1 + \eta_2 \text{sgn}\left(\left(\psi_2 - \psi_1\right) + \theta_2 x_1\right)} - \frac{1}{-\eta_1 + \eta_2 \text{sgn}\left(\left(\psi_2 - \psi_1\right) + \theta_2 x_1\right)} \end{aligned} \tag{7.117}$$

Next, we isolate x_1 to get

$$\begin{aligned} I_{11} &= \int_{\mathbb{R}} g_1(x_1) \exp\left(-\eta_2 \left|(\psi_2 - \psi_1) + \theta_2 x_1\right| - \frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) \, dx_1 \\ &= e^{j\nu_2\psi_1} \int_{\mathbb{R}} g_1(\nu_2, x_1) \exp\left(-\eta_2 \left|(\psi_2 - \psi_1) + \theta_2 x_1\right| - \frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) \, dx_1 \\ &= e^{j\nu_2\psi_1} \int_{\mathbb{R}} g_1(\nu_2, x_1) \exp\left(-\eta_2 \left|\theta_2\right| \left|\frac{\psi_2 - \psi_1}{-\theta_2} - x_1\right| - \frac{1}{\alpha}\left|\bar{x}_1 - x_1\right|\right) \, dx_1 \\ &\triangleq e^{j\nu_2\psi_1} \int_{\mathbb{R}} g_1(\nu_2, x_1) \exp\left(-\hat{\eta}_{11} \left|\hat{\xi}_{11} - x_1\right| - \hat{\eta}_{12}\left|\hat{\xi}_{12} - x_1\right|\right) \, dx_1 \end{aligned} \tag{7.118}$$

where

$$\begin{aligned} g_1(x_1) &= \frac{1}{\eta_1 + \eta_2 \text{sgn}\left(\left(\psi_2 - \psi_1\right) + \theta_2 x_1\right)} - \frac{1}{-\eta_1 + \eta_2 \text{sgn}\left(\left(\psi_2 - \psi_1\right) + \theta_2 x_1\right)} \\ &= \frac{1}{\eta_1 + \eta_2 \text{sgn}\left(-\theta_2\right) \text{sgn}\left(\frac{\psi_2 - \psi_1}{-\theta_2} - x_1\right)} - \frac{1}{-\eta_1 + \eta_2 \text{sgn}\left(-\theta_2\right) \text{sgn}\left(\frac{\psi_2 - \psi_1}{-\theta_2} - x_1\right)} \\ &\triangleq \frac{1}{\eta_1 + \hat{\rho}_{11} \text{sgn}\left(\hat{\xi}_{11} - x_1\right)} - \frac{1}{-\eta_1 + \hat{\rho}_{11} \text{sgn}\left(\hat{\xi}_{11} - x_1\right)} \end{aligned} \tag{7.119}$$

and

$$\begin{aligned} \hat{\eta}_{11} &= \eta_2 \left|\theta_2\right| \\ \hat{\eta}_{12} &= \frac{1}{\alpha} \\ \hat{\rho}_{11} &= \eta_2 \text{sgn}\left(-\theta_2\right) \\ \hat{\rho}_{12} &= 0 \\ \hat{\xi}_{11} &= -\frac{\psi_2 - \psi_1}{\theta_2} \\ \hat{\xi}_{12} &= \bar{x}_1 \end{aligned} \tag{7.120}$$

We can then use the generalized integral formula (actually, the original (7.3) works just fine) to get

$$\begin{aligned} I_{11} &= G_{11} \exp\left(-\hat{\eta}_{12} \left|\hat{\xi}_{12} - \hat{\xi}_{11}\right|\right) + G_{12} \exp\left(-\hat{\eta}_{11} \left|\hat{\xi}_{11} - \hat{\xi}_{12}\right|\right) \\ &= G_{11} \exp\left(c_{11} \cdot j\nu_1 + c_{12} \cdot j\nu_2 + c_{13}\right) \\ &\quad + G_{12} \exp\left(c_{21} \cdot j\nu_1 + c_{22} \cdot j\nu_2 + c_{23}\right) \end{aligned} \tag{7.121}$$

where

$$\begin{aligned} G_{11} &= \frac{g_1(\hat{\xi}_{11})}{j\nu_1 + \hat{\eta}_{11} + \hat{\eta}_{12}\mathrm{sgn}\left(\hat{\xi}_{12} - \hat{\xi}_{11}\right)} - \frac{g_1(\nu_2, \hat{\xi}_{11})}{j\nu_1 - \hat{\eta}_{11} + \hat{\eta}_{12}\mathrm{sgn}\left(\hat{\xi}_{12} - \hat{\xi}_{11}\right)} \\ &= \frac{\frac{1}{j\nu_2+\eta_1+\hat{\rho}_{11}} - \frac{1}{j\nu_2-\eta_1+\hat{\rho}_{11}}}{j\nu_1 + \hat{\eta}_{11} + \hat{\eta}_{12}\mathrm{sgn}\left(\hat{\xi}_{12} - \hat{\xi}_{11}\right)} - \frac{\frac{1}{j\nu_2+\eta_1-\hat{\rho}_{11}} - \frac{1}{j\nu_2-\eta_1-\hat{\rho}_{11}}}{j\nu_1 - \hat{\eta}_{11} + \hat{\eta}_{12}\mathrm{sgn}\left(\hat{\xi}_{12} - \hat{\xi}_{11}\right)} \\ &\triangleq \frac{\frac{a_{11}}{a_{12}+j\nu_2} - \frac{a_{13}}{a_{14}+j\nu_2}}{a_{15} + j\nu_1 + a_{16} \cdot j\nu_2} - \frac{\frac{b_{11}}{b_{12}+j\nu_2} - \frac{b_{13}}{b_{14}+j\nu_2}}{b_{15} + j\nu_1 + b_{16} \cdot j\nu_2} \\ G_{12} &= \frac{g_1(\hat{\xi}_{12})}{\hat{\eta}_{12} + \hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right)} - \frac{g_1(\nu_2, \hat{\xi}_{12})}{-\hat{\eta}_{12} + \hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right)} \\ &= \frac{\frac{1}{\eta_1+\hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11}-\hat{\xi}_{12}\right)} - \frac{1}{-\eta_1+\hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11}-\hat{\xi}_{12}\right)}}{\hat{\eta}_{12} + \hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right)} - \frac{\frac{1}{\eta_1+\hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11}-\hat{\xi}_{12}\right)} - \frac{1}{-\eta_1+\hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11}-\hat{\xi}_{12}\right)}}{-\hat{\eta}_{12} + \hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right)} \\ &\triangleq \frac{\frac{a_{21}}{a_{22}+j\nu_2} - \frac{a_{23}}{a_{24}+j\nu_2}}{a_{25} + j\nu_1 + a_{26} \cdot j\nu_2} - \frac{\frac{b_{21}}{b_{22}+j\nu_2} - \frac{b_{23}}{b_{24}+j\nu_2}}{b_{25} + j\nu_1 + b_{26} \cdot j\nu_2} \end{aligned} \tag{7.122}$$

and

$$\begin{aligned} &a_{11} = 1 && b_{11} = 1 && c_{11} = \hat{\xi}_{11} \\ &a_{12} = \eta_1 + \hat{\rho}_{11} && b_{12} = \eta_1 - \hat{\rho}_{11} && c_{12} = \psi_1 \\ &a_{13} = 1 && b_{13} = 1 && c_{13} = -\hat{\eta}_{12}\left|\xi_{12} - \xi_{11}\right| \\ &a_{14} = -\eta_1 + \hat{\rho}_{11} && b_{14} = -\eta_1 - \hat{\rho}_{11} && \\ &a_{15} = \hat{\eta}_{11} + \hat{\eta}_{12}\mathrm{sgn}\left(\hat{\xi}_{12} - \hat{\xi}_{11}\right) && b_{15} = -\hat{\eta}_{11} + \hat{\eta}_{12}\mathrm{sgn}\left(\hat{\xi}_{12} - \hat{\xi}_{11}\right) && \\ &a_{16} = 0 && b_{16} = 0 && \end{aligned} \tag{7.123}$$

and

$$
\begin{aligned}
&a_{21} = 1 && b_{21} = 1 && c_{21} = \hat{\xi}_{12} \\
&a_{22} = \eta_1 + \hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right) && b_{22} = \eta_1 + \hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right) && c_{22} = \psi_1 \\
&a_{23} = 1 && b_{23} = 1 && c_{23} = -\hat{\eta}_{11}\left|\xi_{11} - \xi_{12}\right| \\
&a_{25} = -\eta_1 + \hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right) && b_{25} = -\eta_1 + \hat{\rho}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right) && \\
&a_{27} = \hat{\eta}_{12} + \hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right) && b_{27} = -\hat{\eta}_{12} + \hat{\eta}_{11}\mathrm{sgn}\left(\hat{\xi}_{11} - \hat{\xi}_{12}\right) && \\
&a_{29} = 0 && b_{29} = 0 &&
\end{aligned}
\tag{7.124}
$$

7.12.2 Integrate term 2 with respect to x_1

We repeat the integration for $I_{22}(x_1)$, though by careful inspection, we can say that for

$$
\begin{aligned}
\hat{\eta}_{21} &= \eta_1 \left|\theta_2\right| \\
\hat{\eta}_{22} &= \frac{1}{\alpha} \\
\hat{\rho}_{21} &= \eta_1 \mathrm{sgn}\left(\theta_2\right) \\
\hat{\rho}_{22} &= 0 \\
\hat{\xi}_{21} &= \frac{\psi_1 - \psi_2}{\theta_2} \\
\hat{\xi}_{22} &= \bar{x}_1
\end{aligned}
\tag{7.125}
$$

We can then use the modified integral formula (actually, the original works just fine) to get

$$
\begin{aligned}
I_{12} &= G_{21}\exp\left(-\hat{\eta}_{22}\left|\hat{\xi}_{22} - \hat{\xi}_{21}\right|\right) + G_{22}\exp\left(-\hat{\eta}_{21}\left|\hat{\xi}_{21} - \hat{\xi}_{22}\right|\right) \\
&= G_{21}\exp\left(c_{11}\cdot j\nu_1 + c_{12}\cdot j\nu_2 + c_{13}\right) \\
&\quad + G_{22}\exp\left(c_{21}\cdot j\nu_1 + c_{22}\cdot j\nu_2 + c_{23}\right)
\end{aligned}
\tag{7.126}
$$

where

$$
\begin{aligned}
G_{21} &\triangleq \frac{\frac{a_{11}}{a_{12}+j\nu_2} - \frac{a_{13}}{a_{14}+j\nu_2}}{a_{15} + j\nu_1 + a_{16}\cdot j\nu_2} - \frac{\frac{b_{11}}{b_{12}+j\nu_2} - \frac{b_{13}}{b_{14}+j\nu_2}}{b_{15} + j\nu_1 + b_{16}\cdot j\nu_2} \\
G_{22} &\triangleq \frac{\frac{a_{21}}{a_{22}+j\nu_2} - \frac{a_{23}}{a_{24}+j\nu_2}}{a_{25} + j\nu_1 + a_{26}\cdot j\nu_2} - \frac{\frac{b_{21}}{b_{22}+j\nu_2} - \frac{b_{23}}{b_{24}+j\nu_2}}{b_{25} + j\nu_1 + b_{26}\cdot j\nu_2}
\end{aligned}
\tag{7.127}
$$

and

$$
\begin{aligned}
&a_{11} = 1 && b_{11} = 1 && c_{11} = \hat{\xi}_{21} \\
&a_{12} = \eta_2 + \hat{\rho}_{21} && b_{12} = \eta_2 - \hat{\rho}_{21} && c_{12} = \psi_1 \\
&a_{13} = 1 && b_{13} = 1 && c_{13} = -\hat{\eta}_{22}\left|\hat{\xi}_{22} - \hat{\xi}_{21}\right| \\
&a_{14} = -\eta_2 + \hat{\rho}_{21} && b_{14} = -\eta_2 - \hat{\rho}_{21} && \\
&a_{15} = \hat{\eta}_{21} + \hat{\eta}_{22}\mathrm{sgn}\left(\hat{\xi}_{22} - \hat{\xi}_{21}\right) && b_{15} = -\hat{\eta}_{21} + \hat{\eta}_{22}\mathrm{sgn}\left(\hat{\xi}_{22} - \hat{\xi}_{21}\right) && \\
&a_{16} = 0 && b_{16} = 0 &&
\end{aligned}
\tag{7.128}
$$

and

$$
\begin{aligned}
&a_{21} = 1 && b_{21} = 1 && c_{21} = \hat{\xi}_{22} \\
&a_{22} = \eta_2 + \hat{\rho}_{21}\mathrm{sgn}\left(\hat{\xi}_{21} - \hat{\xi}_{22}\right) && b_{22} = \eta_2 + \hat{\rho}_{21}\mathrm{sgn}\left(\hat{\xi}_{21} - \hat{\xi}_{22}\right) && c_{22} = \psi_1 \\
&a_{23} = 1 && b_{23} = 1 && c_{23} = -\hat{\eta}_{21}\left|\hat{\xi}_{21} - \hat{\xi}_{22}\right| \\
&a_{24} = -\eta_2 + \hat{\rho}_{21}\mathrm{sgn}\left(\hat{\xi}_{21} - \hat{\xi}_{22}\right) && b_{24} = -\eta_2 + \hat{\rho}_{21}\mathrm{sgn}\left(\hat{\xi}_{21} - \hat{\xi}_{22}\right) && \\
&a_{25} = \hat{\eta}_{22} + \hat{\eta}_{21}\mathrm{sgn}\left(\hat{\xi}_{21} - \hat{\xi}_{22}\right) && b_{25} = -\hat{\eta}_{22} + \hat{\eta}_{21}\mathrm{sgn}\left(\hat{\xi}_{21} - \hat{\xi}_{22}\right) && \\
&a_{26} = 0 && b_{26} = 0 &&
\end{aligned}
\tag{7.129}
$$

Once the CF is in the standard form, the equations for finding the moments are identical to those in the previous example for f_{X_1}!

7.13 First two moments in $\mathbb{R}^2$ using characteristic functions

The first two moments in $\mathbb{R}^2$ are

$$
\begin{aligned}
E\left[X\right] &= \begin{bmatrix} \frac{1}{j}\frac{\partial\phi}{\partial\nu_1} \\ \frac{1}{j}\frac{\partial\phi}{\partial\nu_2} \end{bmatrix}_{\boldsymbol{\nu}=\mathbf{0}} \\
E\left[XX^T\right] &= \begin{bmatrix} -\frac{\partial^2\phi}{\partial\nu_1^2} & -\frac{\partial^2\phi}{\partial\nu_1\partial\nu_2} \\ -\frac{\partial^2\phi}{\partial\nu_1\partial\nu_2} & -\frac{\partial^2\phi}{\partial\nu_2^2} \end{bmatrix}_{\boldsymbol{\nu}=\mathbf{0}}
\end{aligned}
\tag{7.130}
$$

where

$$\begin{aligned}
\frac{\partial \phi}{\partial \nu_1} &= \frac{\partial G}{\partial \nu_1} \cdot \epsilon + G \cdot \frac{\partial \epsilon}{\partial \nu_1} \\
\frac{\partial \phi}{\partial \nu_2} &= \frac{\partial G}{\partial \nu_2} \cdot \epsilon + G \cdot \frac{\partial \epsilon}{\partial \nu_2} \\
\frac{\partial^2 \phi}{\partial \nu_1^2} &= \frac{\partial^2 G}{\partial \nu_1^2} \cdot \epsilon + 2\frac{\partial G}{\partial \nu_1} \cdot \frac{\partial \epsilon}{\partial \nu_1} + G \cdot \frac{\partial^2 \epsilon}{\partial \nu_1^2} \\
\frac{\partial^2 \phi}{\partial \nu_1 \nu_2} &= \frac{\partial^2 G}{\partial \nu_1 \nu_2} \cdot \epsilon + \frac{\partial G}{\partial \nu_1} \cdot \frac{\partial \epsilon}{\partial \nu_2} + \frac{\partial G}{\partial \nu_2} \cdot \frac{\partial \epsilon}{\partial \nu_1} + G \cdot \frac{\partial^2 \epsilon}{\partial \nu_1 \nu_2} \\
\frac{\partial^2 \phi}{\partial \nu_2^2} &= \frac{\partial^2 G}{\partial \nu_2^2} \cdot \epsilon + 2\frac{\partial G}{\partial \nu_2} \cdot \frac{\partial \epsilon}{\partial \nu_2} + G \cdot \frac{\partial^2 \epsilon}{\partial \nu_2^2}
\end{aligned} \tag{7.131}$$

The first partial derivatives with respect to ν_1 and ν_2 are

$$\begin{aligned}
\frac{\partial G}{\partial \nu_1} &= -\frac{N_a \frac{\partial D_a}{\partial \nu_1}}{D_a^2} + \frac{N_b \frac{\partial D_b}{\partial \nu_1}}{D_b^2} \\
\frac{\partial G}{\partial \nu_2} &= \frac{-N_a \frac{\partial D_a}{\partial \nu_2} + D_a \frac{\partial N_a}{\partial \nu_2}}{D_a^2} - \frac{-N_b \frac{\partial D_b}{\partial \nu_2} + D_b \frac{\partial N_b}{\partial \nu_2}}{D_b^2} \\
\frac{\partial \epsilon}{\partial \nu_1} &= jc_1 \epsilon \\
\frac{\partial \epsilon}{\partial \nu_2} &= jc_2 \epsilon
\end{aligned} \tag{7.132}$$

where

$$\begin{aligned}
\frac{\partial N_a}{\partial \nu_1} &= 0 \\
\frac{\partial D_a}{\partial \nu_1} &= ja_8 \\
\frac{\partial N_a}{\partial \nu_2} &= \frac{-ja_1 a_3}{(a_2 + j\nu_2 a_3)^2} - \frac{-ja_4 a_6}{(a_5 + j\nu_2 a_6)^2} \\
\frac{\partial D_a}{\partial \nu_2} &= ja_9
\end{aligned} \tag{7.133}$$

and the corresponding partials for the N_b, D_b terms are the same except using b coefficients.

The second partial derivatives are as follows

$$\begin{aligned}
\frac{\partial^2 G}{\partial \nu_1^2} &= \frac{-\frac{\partial}{\partial \nu_1}\left(N_a \frac{\partial D_a}{\partial \nu_1}\right) D_a^2 + N_a \frac{\partial D_a}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_1} D_a^2}{D_a^4} - \frac{-\frac{\partial}{\partial \nu_1}\left(N_b \frac{\partial D_b}{\partial \nu_1}\right) D_b^2 + N_b \frac{\partial D_b}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_1} D_b^2}{D_b^4} \\
&= \frac{N_a \frac{\partial D_a}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_1} D_a^2}{D_a^4} - \frac{N_b \frac{\partial D_b}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_1} D_b^2}{D_b^4}
\end{aligned} \tag{7.134}$$

where we use $\dfrac{\partial N_a}{\partial \nu_1} = 0$ and $\dfrac{\partial^2 D_a}{\partial \nu_1^2} = 0$. Next, we have the cross derivatives

$$\begin{aligned}
\frac{\partial^2 G}{\partial \nu_1 \partial \nu_2} &= \frac{-\frac{\partial}{\partial \nu_2}\left(N_a \frac{\partial D_a}{\partial \nu_1}\right) \cdot D_a^2 + N_a \frac{\partial D_a}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_2} D_a^2}{D_a^4} - \frac{-\frac{\partial}{\partial \nu_2}\left(N_b \frac{\partial D_b}{\partial \nu_1}\right) \cdot D_b^2 + N_b \frac{\partial D_b}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_2} D_b^2}{D_b^4} \\
&= \frac{-\frac{\partial N_a}{\partial \nu_2} \cdot \frac{\partial D_a}{\partial \nu_1} \cdot D_a^2 + N_a \frac{\partial D_a}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_2} D_a^2}{D_a^4} - \frac{-\frac{\partial N_b}{\partial \nu_2} \cdot \frac{\partial D_b}{\partial \nu_1} \cdot D_b^2 = N_b \frac{\partial D_b}{\partial \nu_1} \cdot \frac{\partial}{\partial \nu_2} D_b^2}{D_b^4}
\end{aligned} \tag{7.135}$$

where we use $\dfrac{\partial^2 D_a}{\partial \nu_1 \partial \nu_2} = 0$. Finally, we have

$$\begin{aligned}
\frac{\partial^2 G}{\partial \nu_2^2} &= \frac{\frac{\partial}{\partial \nu_2}\left(-N_a \frac{\partial D_a}{\partial \nu_2} + D_a \frac{\partial N_a}{\partial \nu_2}\right) \cdot D_a^2 - \left(-N_a \frac{\partial D_a}{\partial \nu_2} + D_a \frac{\partial N_a}{\partial \nu_2}\right) \cdot \frac{\partial}{\partial \nu_2} D_a^2}{D_a^4} \\
&\quad - \frac{\frac{\partial}{\partial \nu_2}\left(-N_b \frac{\partial D_b}{\partial \nu_2} + D_b \frac{\partial N_b}{\partial \nu_2}\right) \cdot D_b^2 - \left(-N_b \frac{\partial D_b}{\partial \nu_2} + D_b \frac{\partial N_b}{\partial \nu_2}\right) \cdot \frac{\partial}{\partial \nu_2} D_b^2}{D_b^4} \\
&= \frac{\left(-\frac{\partial N_a}{\partial \nu_2} \cdot \frac{\partial D_a}{\partial \nu_2} - N_a \cdot \frac{\partial^2 D_a}{\partial \nu_2^2} + \frac{\partial D_a}{\partial \nu_2} \cdot \frac{\partial N_a}{\partial \nu_2} + D_a \frac{\partial^2 N_a}{\partial \nu_2^2}\right) \cdot D_a^2 - \left(-N_a \frac{\partial D_a}{\partial \nu_2} + D_a \frac{\partial N_a}{\partial \nu_2}\right) \cdot \frac{\partial}{\partial \nu_2} D_a^2}{D_a^4} \\
&\quad - \frac{\left(-\frac{\partial N_b}{\partial \nu_2} \cdot \frac{\partial D_b}{\partial \nu_2} - N_b \frac{\partial^2 D_b}{\partial \nu_2^2} + \frac{\partial D_b}{\partial \nu_2} \cdot \frac{\partial N_b}{\partial \nu_2} + D_b \frac{\partial^2 N_b}{\partial \nu_2^2}\right) \cdot D_b^2 - \left(-N_b \frac{\partial D_b}{\partial \nu_2} + D_b \frac{\partial^2 N_b}{\partial \nu_2^2}\right) \cdot \frac{\partial}{\partial \nu_2} D_b^2}{D_b^4} \\
&= \frac{D_a^3 \frac{\partial^2 N_a}{\partial \nu_2^2} - \left(-N_a \frac{\partial D_a}{\partial \nu_2} + D_a \frac{\partial N_a}{\partial \nu_2}\right) \cdot \frac{\partial}{\partial \nu_2} D_a^2}{D_a^4} \frac{D_b^3 \frac{\partial^2 N_b}{\partial \nu_2^2} - \left(-N_b \frac{\partial D_b}{\partial \nu_2} + D_b \frac{\partial^2 N_b}{\partial \nu_2^2}\right) \cdot \frac{\partial}{\partial \nu_2} D_b^2}{D_b^4}
\end{aligned} \tag{7.136}$$

where we use $\dfrac{\partial^2 D_a}{\partial \nu_2^2} = \dfrac{\partial^2 D_b}{\partial \nu_2^2} = 0$.

The second partials for the exponential term are

$$\begin{aligned}
\frac{\partial^2 \epsilon}{\partial \nu_1^2} &= -c_1^2 \epsilon \\
\frac{\partial^2 \epsilon}{\partial \nu_1 \partial \nu_2} &= -c_1 c_2 \epsilon \\
\frac{\partial \epsilon}{\partial \nu_2} &= -c_2^2 \epsilon
\end{aligned} \tag{7.137}$$

where

$$
\begin{aligned}
\frac{\partial^2 D_a}{\partial \nu_1 \partial \nu_2} &= \frac{\partial}{\partial \nu_2}\frac{\partial D_a}{\partial \nu_1} = \frac{\partial}{\partial \nu_2} j a_9 = 0 \\
\frac{\partial^2 D_b}{\partial \nu_1 \partial \nu_2} &= \frac{\partial}{\partial \nu_2}\frac{\partial D_b}{\partial \nu_1} = \frac{\partial}{\partial \nu_2} j b_9 = 0 \\
\frac{\partial^2 D_a}{\partial \nu_1^2} &= \frac{\partial^2 D_b}{\partial \nu_1^2} = 0 \\
\frac{\partial^2 D_a}{\partial \nu_2^2} &= \frac{\partial^2 D_b}{\partial \nu_2^2} = 0 \\
\frac{\partial}{\partial \nu_1}\left(N_a \frac{\partial D_a}{\partial \nu_1}\right) &= \frac{\partial}{\partial \nu_1}\left(N_b \frac{\partial D_b}{\partial \nu_1}\right) = 0 \\
\frac{\partial}{\partial \nu_1} D_a^2 &= 2 D_a \frac{\partial D_a}{\partial \nu_1} = j 2 a_8 D_a \\
\frac{\partial}{\partial \nu_1} D_b^2 &= 2 D_b \frac{\partial D_b}{\partial \nu_1} = j 2 b_8 D_b \\
\frac{\partial^2 N_a}{\partial \nu_2^2} &= \frac{\partial}{\partial \nu_2}\frac{\partial N_a}{\partial \nu_2} = -\frac{2 a_1 a_3^2}{(a_2 + j \nu_2 a_3)^3} + \frac{2 a_4 a_6^2}{(a_5 + j \nu_2 a_6)^3} \\
\frac{\partial^2 N_b}{\partial \nu_2^2} &= \frac{\partial}{\partial \nu_2}\frac{\partial N_b}{\partial \nu_2} = -\frac{2 b_1 b_3^2}{(b_2 + j \nu_2 b_3)^3} + \frac{2 b_4 b_6^2}{(b_5 + j \nu_2 b_6)^3} \\
\frac{\partial}{\partial \nu_2}\left(N_a \frac{\partial D_a}{\partial \nu_1}\right) &= \frac{\partial N_a}{\partial \nu_2} \cdot \frac{\partial D_a}{\partial \nu_1} + N_a \cdot \frac{\partial^2 D_a}{\partial \nu_1 \partial \nu_2} = \frac{\partial N_a}{\partial \nu_2} \cdot \frac{\partial D_a}{\partial \nu_1} \\
\frac{\partial}{\partial \nu_2}\left(N_b \frac{\partial D_b}{\partial \nu_1}\right) &= \frac{\partial N_b}{\partial \nu_2} \cdot \frac{\partial D_b}{\partial \nu_1} + N_b \cdot \frac{\partial^2 D_b}{\partial \nu_1 \partial \nu_2} = \frac{\partial N_b}{\partial \nu_2} \cdot \frac{\partial D_b}{\partial \nu_1} \\
\frac{\partial}{\partial \nu_2}\left(D_a \frac{\partial N_a}{\partial \nu_2}\right) &= \frac{\partial D_a}{\partial \nu_2} \cdot \frac{\partial N_a}{\partial \nu_2} + D_a \cdot \frac{\partial^2 N_a}{\partial \nu_2^2} \\
\frac{\partial}{\partial \nu_2}\left(D_b \frac{\partial N_b}{\partial \nu_2}\right) &= \frac{\partial D_b}{\partial \nu_2} \cdot \frac{\partial N_b}{\partial \nu_2} + D_b \cdot \frac{\partial^2 N_b}{\partial \nu_2^2} \\
\frac{\partial}{\partial \nu_2} D_a^2 &= 2 D_a \frac{\partial D_a}{\partial \nu_2} \\
\frac{\partial}{\partial \nu_2} D_b^2 &= 2 D_b \frac{\partial D_b}{\partial \nu_2}
\end{aligned}
\tag{7.138}
$$

Note that the first and second moments use the *normalized* characteristic function $\phi^{k|k}$ by using f_{Y_k} to normalize $\bar{\phi}^{k|k}$. In practice, we can apply the normalization factor either before or after computing the moments, since it's just a scale factor over a sum.

REFERENCES

[ABB11] Aleksandr Y. Aravkin, Bradley M. Bell, James V. Burke, and Gianluigi Pillonetto. "An l1-Laplace Robust Kalman Smoother." *IEEE Transactions on Automatic Control*, **56**(12):2898–2911, 2011.

[ABL17] Aleksandr Y. Aravkin, James V. Burke, Lennart Ljung, Aurelie Lozano, and Gianluigi Pillonetto. "Generalized Kalman smoothing: Modeling and algorithms." *Automatica (Journal of IFAC)*, **86**(C):63–86, 2017.

[AF93] David Avis and Komei Fukuda. "Reverse Search for Enumeration." *Discrete Applied Mathematics*, **65**:21–46, 1993.

[Bry74] Maurice C. Bryson. "Heavy-Tailed Distributions: Properties and Tests." *Technometrics*, **16**(1):61–68, 1974.

[DSY18] N. Duong, J. Speyer, J. Yoneyama, and M. Idan. "Laplace Estimator for Linear Scalar Systems." In *57th IEEE Conference on Decision and Control*, Miami, FL, USA, December 2018.

[FMS16] F. Farokhi, J. Milosevic, and H. Sandberg. "Optimal state estimation with measurements corrupted by Laplace noise." In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 302–307, 2016.

[FSI15] J. Fernández, J. L. Speyer, and M. Idan. "Stochastic Estimation for Two-State Linear Dynamic Systems with Additive Cauchy Noises." *IEEE Transactions on Automatic Control*, **60**(12), 2015.

[IS10] M. Idan and J. L. Speyer. "Cauchy Estimation for Linear Scalar Systems." *IEEE Transactions on Automatic Control*, **55**(6):1329–1342, 2010.

[IS12] M. Idan and J. L. Speyer. "State Estimation for Linear Scalar Dynamic Systems with Additive Cauchy Noises: Characteristic Function Approach." *SIAM Journal on Control and Optimization*, **50**(4):1971–1994, 2012.

[IS14] M. Idan and J. L. Speyer. "Multivariate Cauchy Estimator with Scalar Measurent and Process Noises." *SIAM Journal on Control and Optimization*, **52**(2):1108–1141, 2014.

[KFR98] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner. "Near optimal detection of signals in impulsive noise modeled with a symmetric /spl alpha/-stable distribution." *IEEE Communications Letters*, **2**(10):282–284, 1998.

[KSO16] P. Kozierski, T. Sadalla, A. Owczarkowski, and S. Drgas. "Particle filter in multidimensional systems." In *2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR)*, pp. 806–810, 2016.

[LBG10] B. N. M. Laska, M. Bolic, and R. A. Goubran. "Particle Filter Enhancement of Speech Spectral Amplitudes." *IEEE Transactions on Audio, Speech, and Language Processing*, **18**(8):2155–2167, 2010.

[LBS04] U. Ledzewicz, T. Brown, and H. Schättler. "Comparison of optimal controls for a model in cancer chemotherapy with L1- and L2-type objectives." *Optimization Methods and Software*, **19**(3-4):339–350, 2004.

[Lin05] Mikael Linden. "Estimating the distribution of volatility of realized stock returns and exchange rate changes." *Physica A: Statistical Mechanics and its Applications*, **352**(2):573 – 583, 2005.

[MBL11] C. Musso, P. Bui Quang, and F. Le Gland. "Introducing the Laplace approximation in particle filtering." In *14th International Conference on Information Fusion*, pp. 1–8, 2011.

[MKP14] P.P. Markopoulos, G.N. Karystinos, and D.A. Pados. "Optimal Algorithms for L1-subspace Signal Processing." *IEEE Transactions on Signal Processing*, **62**(19):5046–5058, 2014.

[PIM14] M.R. de Pihno, Kornienko I., and H. Maurer. "Optimal Control of a SEIR Model with Mixed Constraints and L1 Cost." In *11th Portuguese Conference on Automatic Control*, Porto, Portugal, July 2014.

[RC18] Miroslav Rada and Michal Cerný. "A New Algorithm for Enumeration of Cells of Hyperplane Arrangements and a Comparison with Avis and Fukuda's Reverse Search." *SIAM J. Discret. Math.*, **32**(1):455–473, 2018.

[Ros06] I.H. Ross. "Space trajectory optimization and L1-optimal control problems." *Elsevier Astrodynamics Series*, **1**(1):155–188, 2006.

[RVG06] H. Rabbani, M. Vafadust, and S. Gazor. "Image Denoising Based on a Mixture of Laplace Distributions with Local Parameters in Complex Wavelet Domain." In *2006 International Conference on Image Processing*, pp. 2597–2600, 2006.

[SC08] J. L. Speyer and W. H. Chung. *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

[Sel08] I. W. Selesnick. "The Estimation of Laplace Random Vectors in Additive White Gaussian Noise." *IEEE Transactions on Signal Processing*, **56**(8):3482–3496, 2008.

[SIF14] Jason L. Speyer, Moshe Idan, and Javier H. Fernández. "A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise." *Automatica*, **50**:114–127, 2014.

[Sta89] R. Stanley. "Log-Concave and Unimodal Sequences in Algebra, Combinatorics, and Geometry." *Annals of the New York Academy of Sciences*, **576**(1):500–535, 1989.

[SYD18] J. Speyer, J. Yoneyama, N. Duong, and M. Idan. "Laplace One-Step Controller for Linear Scalar Systems." In *European Control Conference*, Limassol, Cyprus, June 12-15 2018.

[VM06] G. Vossen and H. Maurer. "On L1-minimization in optimal control and applications to robotics." *Optimal Control Applications and Methods*, **6**(27):301–312, 2006.