Automatica 144 (2022) 110301

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Laplace estimation for scalar linear systems[☆]

Nhattrieu C. Duong[a,*], Jason L. Speyer[a], Moshe Idan[b]

[a] Mechanical and Aerospace Engineering, University of California, Los Angeles, Los Angeles, CA 90095, USA
[b] Aerospace Engineering, Technion-Israel Institute of Technology, Haifa 31000, Israel

ARTICLE INFO

Article history:
Received 7 February 2021
Received in revised form 20 January 2022
Accepted 3 February 2022
Available online xxxx

Keywords:
Estimation and filtering
Linear systems

ABSTRACT

Uncertainties in many physical systems have impulsive properties poorly modeled by Gaussian distributions. Refocusing previous work, an estimator is derived for a scalar discrete-time linear system with additive Laplace measurement and process noises. The *a priori* and *a posteriori* conditional probability density functions (pdf) of the state given a measurement sequence are propagated recursively and in closed form, and the *a posteriori* conditional mean and variance are derived analytically from the conditional pdf. A simulation for an estimator is presented, demonstrating marked resilience to large, un-modeled spikes in the measurements.

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 light-tailed, *i.e.* the decay rate of the tail is exponential or faster (Bryson, 1974). While its structure lends itself to compact, closed-form analytical state estimation and control, 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 (Kuruoglu, Fitzgerald, & Rayner, 1998), and stock market volatility (Linden, 2005), and algorithms derived assuming Gaussian distributions are susceptible to such outliers. While ad-hoc methods, such as pre-filters, have been developed to compensate for this limitation, we instead wish to exploit the properties of the Laplace distribution for this purpose.

With the advent of fast, inexpensive computational capabilities, simulation or Monte Carlo 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 (Musso, Bui Quang, & Le Gland, 2011) in robotics (Kozierski, Sadalla, Owczarkowski, & Drgas, 2016), navigation, and image processing, using both Gaussian and non-Gaussian noise. Laplace densities have been used in areas such as image (Rabbani, Vafadust, & Gazor, 2006) and speech (Laska, Bolic, & Goubran, 2010) processing. However, these techniques are approximate by nature and do not produce explicit closed-form expressions for the minimum variance error of the state estimate.

In contrast with Monte Carlo methods and prior to this work, there were two known analytic recursive linear estimators: the Kalman filter and the Cauchy estimator. The former, derived based on Gaussian noise, has been applied to a large class of problems. However, it is susceptible to outliers, which result in both large estimation error as well as under-estimated error variances. The Cauchy estimator was developed in Idan and Speyer (2010, 2012, 2014), which addresses these limitations. It is considered a heavy-tailed estimator, and its behavior is quite different compared to the Kalman filter. In particular, the conditional variance was shown to be a function of the measurements as well as the noise parameters, whereas the Kalman filter variance can be computed *a priori*. One can see the consequences of this when the Kalman filter processes data with Cauchy noise (Idan & Speyer, 2012).

Earlier work was done in the estimation of a Laplace random vector corrupted by Gaussian noise (Selesnick, 2008) and state estimation for linear systems driven by Laplace noise using a bank of Kalman filters (Farokhi, Milosevic, & Sandberg, 2016). However, in this work, we derive a recursive analytical state estimator in closed-form. Like the Kalman filter and Cauchy state estimator, the solution is the exact minimum-variance 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 of the state given the measurement history (referred to as cdpf for brevity). The *a posteriori* cpdf of the Kalman filter is Gaussian and is therefore both symmetric and unimodal. In contrast, the *a posteriori* cpdf of the Cauchy estimator is neither symmetric nor unimodal. Sitting between the other two, the *a posteriori* cpdf of the Laplace estimator is not symmetric but is unimodal.

[☆] This paper was presented at The 57th IEEE Conference on Decision and Control, December 17–19, 2018, Miami Beach, Florida, USA. This paper was recommended for publication in revised form by Associate Editor Chanying Li under the direction of Editor Miroslav Krstic.

* Corresponding author.
E-mail addresses: ncduong@ucla.edu (N.C. Duong), speyer@seas.ucla.edu (J.L. Speyer), moshe.idan@technion.ac.il (M. Idan).

https://doi.org/10.1016/j.automatica.2022.110301
0005-1098/© 2022 Published by Elsevier Ltd.

Like the Cauchy estimator, the Laplace estimator is resilient to outliers, but it also retains all its moments for the *a posteriori* cpdf, whereas the Cauchy estimator only preserves the first two moments. In addition, the *a priori* cpdf of the Cauchy estimator has no moments, whereas the Laplace estimator's *a priori* cpdf retains all its moments. Furthermore, note that in the development of the Cauchy estimator for vector states, the characteristic function of the cpdf is used. The characteristic function of the Cauchy pdf in the spectral variable is functionally similar to the Laplace pdf. This allows the integral formula used in Appendix of Idan and Speyer (2014) for measurement updates to be applicable to time propagation for the Laplace estimator derivation, as will be shown in Section 3. In this paper, make use of characteristic functions but only to normalize the cpfs and obtain the moments for the Laplace estimator. This method is simpler to implement compared to using the integral formula to directly obtain the moments, as was done in Duong, Speyer, Yoneyama, and Idan (2018).

In Section 2, we define the scalar discrete-time linear system and state our problem. In Section 3, we clarify and expand on the derivation of the unnormalized conditional pdfs (ucpdf) for the first steps presented previously in Duong et al. (2018) and present the general form of the *a priori* ucpdf up to $k-1$ measurements. We then prove the recursion by induction in Section 4 and detail a robust algorithm to combine terms in Section 5. In Section 6 we derive the closed-form equations for the mean and variance using characteristic functions and examine unique features of the mean and variance as a function of the measurements. In Section 7, we present and discuss a 50-step numerical simulation which shows the robustness of the scalar Laplace estimator. Finally, we offer some concluding remarks in Section 8.

2. Problem statement

A scalar discrete-time linear system with scalar state $\tilde{x}_k$, measurement $\tilde{z}_k$, independent measurement noise v_k and independent process noise w_k is given by

$$\tilde{x}_{k+1} = \Phi\tilde{x}_k + w_k, \qquad \tilde{z}_k = H\tilde{x}_k + v_k, \tag{1}$$

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

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

$$f_W(w_k) = \frac{1}{2\beta}e^{-\frac{1}{\beta}|w_k|} \triangleq \frac{1}{2\beta}\bar{f}_W(w_k) \tag{3}$$

$$f_V(v_k) = \frac{1}{2\gamma}e^{-\frac{1}{\gamma}|v_k|} \triangleq \frac{1}{2\gamma}\bar{f}_V(v_k), \tag{4}$$

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

$$\tilde{x}_k = \bar{x}_k + x_k, \qquad \tilde{z}_k = \bar{z}_k + z_k, \tag{5}$$

where the deterministic part is

$$\bar{x}_{k+1} = \Phi\bar{x}_k, \qquad \bar{z}_k = H\bar{x}_k, \tag{6}$$

and the stochastic part is

$$x_{k+1} = \Phi x_k + w_k, \qquad z_k = Hx_k + v_k, \tag{7}$$

with the stochastic initial conditions

$$f_{X_1}(x_1) = \frac{1}{2\alpha}e^{-\frac{1}{\alpha}|x_1|} \triangleq \frac{1}{2\alpha}\bar{f}_{X_1}(x_1). \tag{8}$$

We define the stochastic part of the measurement history up to step k as a random sequence $\boldsymbol{Y}_k = \{Z_1, \ldots, Z_k\}$, with the associated realization $\boldsymbol{y}_k = \{z_1, \ldots, z_k\}$. For the remainder of the estimation derivation, we consider only the stochastic part of the system.

Throughout this paper, we will refer to the sgn (x) function, which is defined using the convention

$$\operatorname{sgn}(x) = \begin{cases} -1, & x \le 0 \\ +1, & x > 0. \end{cases} \tag{9}$$

The goal of this paper is to develop the conditional density function (cpdf) in an analytic form for the system described in (7). From the cpdf, we derive expressions for the mean and variance for an estimator.

3. Laplace conditional PDF

To motivate the *a priori* cpdf $f_{X_k|\boldsymbol{Y}_{k-1}}\left(x_k|\boldsymbol{y}_{k-1}\right)$ at step $k-1$, we begin with the initial pdf f_{X_1} and sequentially determine the conditional pdf as we make a measurement update and perform time propagation. We then deduce the general form of $f_{X_k|\boldsymbol{Y}_{k-1}}\left(x_k|\boldsymbol{y}_{k-1}\right)$.

We will use a special superscript to keep track of the index of the state and last measurement. The superscript $k|k$ refers to components of the *a posteriori* pdf of x_k conditioned on the measurement sequence $\boldsymbol{Y}_k$, or $f_{X_k|\boldsymbol{Y}_k}(x_k|\boldsymbol{y}_k)$. After a time propagation, x_k becomes x_{k+1}, and components of the *a priori* conditional pdf, $f_{X_{k+1}|\boldsymbol{Y}_k}(x_{k+1}|\boldsymbol{y}_k)$, are indicated with the superscript $k+1|k$.

3.1. Measurement update at $k = 1$

We make a measurement at $k=1$, whose noise pdf is given in (4). The pdf of X_1 conditioned on the measurement sequence $\boldsymbol{Y}_1 = \{Z_1\}$ is, by Bayes' Theorem,

$$f_{X_1|\boldsymbol{Y}_1}(x_1|\boldsymbol{y}_1) = \frac{f_{\boldsymbol{Y}_1|X_1}(\boldsymbol{y}_1|x_1)f_{X_1}(x_1)}{f_{\boldsymbol{Y}_1}(\boldsymbol{y}_1)}, \tag{10}$$

where $\boldsymbol{y}_1 = \{z_1\}$. The computations are less cumbersome if the conditional pdf is updated and propagated without the normalization factor, $f_{\boldsymbol{Y}_1}(\boldsymbol{y}_1)$. For the same reasons, we will also use the initial condition and process and measurement noises without their constant factors, indicated by a bar as defined in (8), (3) and (4), *i.e.* $\bar{f}_{X_1}$, $\bar{f}_W$ and $\bar{f}_V$. Therefore, let us define the unnormalized conditional pdf (ucpdf), *i.e.* the joint density, as

$$\begin{aligned}\bar{f}_{X_1|\boldsymbol{Y}_1}(x_1|\boldsymbol{y}_1) &= \bar{f}_{\boldsymbol{Y}_1|X_1}(\boldsymbol{y}_1|x_1)\bar{f}_{X_1}(x_1)\\ &= \bar{f}_{Z_1|X_1}(z_1|x_1)\bar{f}_{X_1}(x_1).\end{aligned} \tag{11}$$

Using the measurement equation in (7) and $\bar{f}_V$ in (4), $\bar{f}_{Z_1|X_1}(z_1|x_1) = \bar{f}_V(z_1 - Hx_1)$, or

$$\bar{f}_{Z_1|X_1}(z_1|x_1) = \exp\left[-\frac{|H|}{\gamma}\left|\frac{z_1}{H} - x_1\right|\right]. \tag{12}$$

Combining (12) and (8), we rewrite the ucpdf as $\bar{f}_{X_1|\boldsymbol{Y}_1}(x_1|\boldsymbol{y}_1) = \bar{f}_V(z_1 - Hx_1)\bar{f}_{X_1}(x_1)$, or

$$\bar{f}_{X_1|\boldsymbol{Y}_1}(x_1|\boldsymbol{y}_1) = \exp\left[-\frac{|H|}{\gamma}\left|\frac{z_1}{H} - x_1\right| - \frac{1}{\alpha}|x_1|\right]. \tag{13}$$

3.2. Propagation from $k = 1$ to $k = 2$

The time propagation from step $k=1$ to step $k=2$ involves determining the conditional density $\bar{f}_{X_2|\boldsymbol{Y}_1}(x_2|\boldsymbol{y}_1)$. To do so, first construct the joint density,

$$\begin{aligned}\bar{f}_{X_1,X_2|\boldsymbol{Y}_1}(x_2, x_1|\boldsymbol{y}_1) &= \bar{f}_{X_1|\boldsymbol{Y}_1}(x_1|\boldsymbol{y}_1)\bar{f}_{X_2|X_1,\boldsymbol{Y}_1}(x_2|x_1,\boldsymbol{y}_1)\\ &= \bar{f}_{X_1|\boldsymbol{Y}_1}(x_1|\boldsymbol{y}_1)\bar{f}_{X_2|X_1}(x_2|x_1),\end{aligned} \tag{14}$$

where the second factor simplifies because $f_{X_2|X_1}$ does not explicitly depend on $\boldsymbol{y}_1$. That is, given $\boldsymbol{y}_1$ does not change the pdf having been given x_1. Using the dynamical equation in (7) and f_W in (3), $\bar{f}_{X_2|X_1}(x_2|x_1) = \bar{f}_W(x_2 - \Phi x_1)$, or

$$\bar{f}_{X_2|X_1}(x_2|x_1) = \exp\left[\frac{|\Phi|}{\beta}\left|\frac{x_2}{\Phi} - x_1\right|\right], \tag{15}$$

and we can rewrite (14) as

$$\begin{aligned}\bar{f}_{X_1,X_2|\mathbf{Y}_1}(x_1, x_2|\mathbf{y}_1) &= \bar{f}_{X_1|\mathbf{Y}_1}(x_1|\mathbf{y}_1)\bar{f}_{X_2|X_1}(x_2|x_1)\\ &= \bar{f}_{X_1|\mathbf{Y}_1}(x_1|\mathbf{y}_1)\bar{f}_W(x_2 - \Phi x_1)\\ &= \exp\left[-\frac{1}{\alpha}|x_1| - \frac{|H|}{\gamma}\left|\frac{z_1}{H} - x_1\right| - \frac{|\Phi|}{\beta}\left|\frac{x_2}{\Phi} - x_1\right|\right].\end{aligned} \tag{16}$$

We then use the Chapman–Kolmogorov equation by integrating (16) over x_1,

$$\bar{f}_{X_2|\mathbf{Y}_1}(x_2|\mathbf{y}_1) = \int_{-\infty}^{\infty} \bar{f}_{X_1,X_2|\mathbf{Y}_1}(x_1, x_2|\mathbf{y}_1)\, dx_1. \tag{17}$$

The solution to the integral in (17) was shown in Appendix B of Idan and Speyer (2014) and is restated in (A.1) of the Appendix for convenience. Using that result and defining $g(x) = 1$, (17) is evaluated as

$$\begin{aligned}\bar{f}_{X_2|\mathbf{Y}_1}(x_2, \mathbf{y}_1) &= \bar{g}_1^{2|1}\cdot\exp\left[-\frac{1}{\alpha}\left|\frac{z_1}{H}\right| - \frac{1}{\beta}\left|\frac{\Phi z_1}{H} - x_2\right|\right]\\ &+ \bar{g}_2^{2|1}\cdot\exp\left[-\frac{1}{\gamma}|z_1| - \frac{1}{\beta}|-x_2|\right]\\ &+ \bar{g}_3^{2|1}\cdot\exp\left[-\frac{|H|}{\gamma|\Phi|}\left|\frac{\Phi z_1}{H} - x_2\right| - \frac{1}{\alpha|\Phi|}|-x_2|\right],\end{aligned} \tag{18}$$

where, for $\delta_i(\rho_l) = \sum_{l=1, l\neq i}^{3} \rho_l \text{sgn}(\xi_l - \xi_i)$,

$$\bar{g}_i^{2|1} = \frac{1}{\rho_i + \delta_i(\rho_l)} - \frac{1}{-\rho_i + \delta_i(\rho_l)}, \tag{19}$$

and $\rho_1 = \frac{|H|}{\gamma}, \rho_2 = \frac{1}{\alpha}, \rho_3 = \frac{|\Phi|}{\beta}, \xi_1 = \frac{z_1}{H}, \xi_2 = 0, \xi_3 = \frac{x_2}{\Phi}$. Since the exponentials without x_2 are constant, we can collapse them into the coefficients to get

$$\begin{aligned}\bar{f}_{X_2|\mathbf{Y}_1}(x_2, \mathbf{y}_1) &= \bar{g}_1^{2|1}\cdot\exp\left[-\frac{1}{\beta}\left|\frac{\Phi z_1}{H} - x_2\right|\right]\\ &+ \bar{g}_2^{2|1}\cdot\exp\left[-\frac{1}{\beta}|x_2|\right]\\ &+ \bar{g}_3^{2|1}\cdot\exp\left[-\frac{|H|}{\gamma|\Phi|}\left|\frac{\Phi z_1}{H} - x_2\right| - \frac{1}{\alpha|\Phi|}|-x_2|\right].\end{aligned} \tag{20}$$

The coefficient terms $\bar{g}_i^{2|1}$ are constant except for one or two step changes. From this observation, we state the following theorem.

Theorem 1. *For*

$$\begin{aligned}\mathcal{A} &= \{A_0, A_1, \ldots, A_m\}\\ &\triangleq \{(-\infty, \xi_1), [\xi_1, \xi_2), \ldots, [\xi_m, +\infty)\},\end{aligned} \tag{21}$$

with $\xi_1 < \cdots < \xi_m \in \mathbb{R}$, any $g : \mathbb{R} \to \mathbb{R}$ constant on $A_i \in \mathcal{A}$, can be expressed as

$$g(x) = \rho_0 + \sum_{i=1}^{m} \rho_i \text{sgn}(\xi_i - x), \tag{22}$$

where, $\rho_0 = \frac{g(x_0) + g(x_m)}{2}$ and $\rho_i = \frac{g(x_{i-1}) - g(x_i)}{2}$, for $x_i \in A_i$.

This theorem is a special case of a more general theorem whose proof is presented in Duong, Idan, Pinchasi, and Speyer (2021). Therefore, the complicated expressions for $\bar{g}_i^{2|1}$ can be simplified to

$$\bar{g}_i^{2|1} = \bar{\rho}_{i0}^{2|1} + \sum_{l=1}^{M_i^{2|1}} \bar{\rho}_{il}^{2|1}\text{sgn}\left(\xi_{il}^{2|1} - x_2\right), \tag{23}$$

where $M_i^{2|1}$ is the ith element of $M^{2|1} = \begin{bmatrix}1 & 1 & 2\end{bmatrix}^T$. This conversion from a nested fractional form in (19) to a sum in (23) is key to developing the recursive structure of the Laplace estimator. The number of sign functions in each term corresponds to the number of addends with x_2 in the associated exponential. We will refer to these addends as "elements". Therefore, $M_i^{2|1}$ refers to the number of unique sign functions of term i.

In addition, we can form an indicator matrix which maps between a certain ordering of terms with an ordering of elements from newest to oldest, or

$$B_2 = \begin{bmatrix}1 & 0\\ 0 & 1\\ 1 & 1\end{bmatrix}. \tag{24}$$

We can see how rows and columns of B_2 correspond to the terms and elements in the exponentials of (20). This indicator matrix is used in the estimation algorithm to keep track of the elements in each term.

3.3. General form of ucpdf

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

$$\bar{f}_{X_k|Y_{k-1}}(x_k|\mathbf{y}_{k-1}) = \sum_{i=1}^{N^{k|k-1}} \bar{g}_i^{k|k-1}\epsilon_i^{k|k-1}, \tag{25}$$

where

$$\begin{aligned}\bar{g}_i^{k|k-1} &= \bar{\rho}_{i0}^{k|k-1} + \sum_{l=1}^{M_i^{k|k-1}} \bar{\rho}_{il}^{k|k-1}\text{sgn}\left(\xi_{il}^{k|k-1} - x_k\right)\\ \epsilon_i^{k|k-1} &= \exp\left[-\sum_{l=1}^{M_i^{k|k-1}} \eta_{il}^{k|k-1}\left|\xi_{il}^{k|k-1} - x_k\right|\right].\end{aligned} \tag{26}$$

Note that the arguments of the sign functions match those of the associated absolute value functions in (26).

The number of terms in (25) is

$$N^{k|k-1} = \sum_{i=1}^{k} i = \frac{k(k+1)}{2}. \tag{27}$$

The number of elements (or sign functions) in each term is given by

$$M^{k|k-1} = \begin{bmatrix}\mathbf{1}^k & \times 1\\ \mathbf{1}^{k-1} & \times 2\\ & \vdots\\ \mathbf{1}^1 & \times k\end{bmatrix}, \tag{28}$$

where $\mathbf{1}^k$ is a column vector of length k composed of 1's and each row corresponds to a term. 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. We can see that $M^{2|1}$ after Eq. (23) has the form of (28) for $k = 2$.

As was done in (24), we can order the elements and terms so that they are mapped using an indicator matrix,

$$B_k = \begin{bmatrix} I_k \\ \left[\mathbf{1}^{k-1\times 1}, \quad I_{k-1}\right] \\ \vdots \\ \left[\mathbf{1}^{2\times k-1}, \quad I_2\right] \\ \mathbf{1}^{1\times k} \end{bmatrix}, \tag{29}$$

where I_m is the identity matrix of dimension $m \times m$, and the rows and columns represent the terms and elements, respectively. We can see that (24) is equal to B_k for $k = 2$ and that the sum of the columns of B_k equals $M^{k|k-1}$. B_k provides an accounting for which terms contains which elements, and it is used to manage the growing number of terms when implementing the estimation algorithm.

An important property of the conditional pdf is that it is log-concave and unimodal. Consider the joint density function of the initial condition and one measurement. To compute the conditional pdf under time propagation, we use the Chapman–Kolmogorov equation, which involves evaluating a convolution integral involving the product of the density functions of the initial condition, measurement noise and process noise. Since each of these density functions is log-concave, and log-concavity is preserved under multiplication and convolution, it follows that both the *a priori* and *a posteriori* conditional pdfs at time step 1, and by extension step k, are also log-concave. This property has important implications when considering algorithms to extremize cost functions based on the *a posteriori* conditional pdfs, such as computing the maximum *a posteriori* estimates. However, in this paper, we will focus on the conditional mean estimate.

4. General form ucpdf proved by induction

We generalize the propagation and update structures presented in Section 3 and show that the general form (25) is preserved under measurement update and time propagation by induction. The following steps should also serve as a guide to develop an implementation of the recursive Laplace estimation algorithm.

4.1. Update from $k|k-1$ to $k|k$

Assuming the ucpdf of the form in (25), we perform a measurement update to step k by

$$\begin{aligned} \bar{f}_{X_k|\boldsymbol{Y}_k}(x_k|\boldsymbol{y}_k) &= \bar{f}_{X_k|\boldsymbol{Y}_{k-1}}(x_k|\boldsymbol{y}_{k-1})\bar{f}_{Z_k|X_k}(z_k|x_k) \\ &= \bar{f}_{X_k|\boldsymbol{Y}_{k-1}}(x_k|\boldsymbol{y}_{k-1})\bar{f}_V\left(\frac{z_k}{H} - x_k\right). \end{aligned} \tag{30}$$

4.2. Propagate from $k|k$ to $k+1|k$

Next, we perform the time propagation to step $k+1$ by using the Chapman–Kolmogorov equation, forming the joint density $\bar{f}_{X_{k+1},X_k|Y_k}$ and evaluating the convolution integral

$$\begin{aligned} \bar{f}_{X_{k+1}|Y_k}(x_{k+1}|\boldsymbol{y}_k) &= \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k}\bar{f}_{X_{k+1}|X_k}\, dx_k \\ &= \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k}\bar{f}_W\left(\frac{x_{k+1}}{\Phi} - x_k\right) dx_k. \end{aligned} \tag{31}$$

Each term i in $\bar{f}_{X_{k+1}|Y_k}$ must be evaluated using the integral formula (A.1) of Appendix. We choose temporary parameters ρ, η and ξ (obviating the need for superscripts), to directly match their counterparts in the integral formula. In addition, we use the parameters from $\bar{f}_{X_k|Y_{k-1}}$ for convenience, because it is simple to append the parameters associated with $\bar{f}_V$ from (30) and $\bar{f}_W$ from (31) to the temporary parameters to maintain a convenient indexing scheme. Using the integral formula, we generate the *a priori* ucpdf at $k+1|k$ using the *a priori* ucpdf at $k|k-1$ in (25). Based on the patterns for $2|1$, we determine that the parameters at $k+1|k$ are

$$\begin{array}{lll} \rho_{i1} = \bar{\rho}_{i1}^{k|k-1} & \xi_{i1} = \xi_{i1}^{k|k-1} & \eta_{i1} = \eta_{i1}^{k|k-1} \\ \rho_{i2} = \bar{\rho}_{i2}^{k|k-1} & \xi_{i2} = \xi_{i2}^{k|k-1} & \eta_{i2} = \eta_{i2}^{k|k-1} \\ & \vdots & \\ \rho_{i,\tilde{M}_i} = \bar{\rho}_{i,\tilde{M}_i}^{k|k-1} & \xi_{i,\tilde{M}_i} = \xi_{i,\tilde{M}_i}^{k|k-1} & \eta_{i,\tilde{M}_i} = \eta_{i,\tilde{M}_i}^{k|k-1} \\ \rho_{i,\tilde{M}_i+1} = 0 & \xi_{i,\tilde{M}_i+1} = \frac{z_k}{H} & \eta_{i,\tilde{M}_i+1} = \frac{|H|}{\gamma} \\ \rho_{i,\tilde{M}_i+2} = 0 & \xi_{i,\tilde{M}_i+2} = \frac{x_{k+1}}{\Phi} & \eta_{i,\tilde{M}_i+2} = \frac{|\Phi|}{\beta} \end{array} \tag{32}$$

where the temporary variable $\tilde{M}_i = M_i^{k|k-1}$ is used here only for compactness. The solution to the integral is determined using the integral formula (A.1) of Appendix.

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

After performing the integral, the arguments of the sign functions are of the form $\xi_l - \xi_i$, and only some of these involve $\xi_{i,\tilde{M}_i+1}$ in (32) containing x_{k+1}. Every term not involving x_{k+1} is constant, so the corresponding exponentials are factored out as a scalar multiple. Some algebra is required to isolate x_{k+1} so that it has the form shown in (26). For example, an element of the exponential function can be re-factored as

$$\eta_j \left| \frac{x_{k+1}}{\Phi} - \frac{z_k}{H} \right| = \frac{\eta_j}{|\Phi|} \left| \frac{\Phi z_k}{H} - x_{k+1} \right|. \tag{33}$$

4.4. Simplify coefficient function and combine terms

The complicated coefficient (g) functions are simplified using the method described in Theorem 1. This allows for the terms with identical exponential parts to be combined by adding the coefficients of the common sign functions. Finally, the expression is rewritten in the general form given in (25), with the indices $k-1$ and k incremented to k and $k+1$, respectively, as

$$\bar{f}_{X_{k+1}|Y_k}(x_{k+1}|\boldsymbol{y}_k) = \sum_{i=1}^{N^{k+1|k}} \bar{g}_i^{k+1|k} \epsilon_i^{k+1|k}, \tag{34}$$

where

$$\begin{aligned} \bar{g}_i^{k+1|k} &= \bar{\rho}_{i0}^{k+1|k} + \sum_{l=1}^{M_i^{k+1|k}} \bar{\rho}_{il}^{k+1|k} \operatorname{sgn}\left(\xi_{il}^{k+1|k} - x_{k+1}\right) \\ \epsilon_i^{k+1|k} &= \exp\left[- \sum_{l=1}^{M_i^{k+1|k}} \eta_{il}^{k+1|k} \left| \xi_{il}^{k+1|k} - x_{k+1} \right| \right]. \end{aligned} \tag{35}$$

5. Term combination

After the isolation step in Section 4.3, the ucpdf has many terms which have the same exponential functions and may be combined by summing their coefficient functions. The number of terms, $N^{k|k-1}$ in (27), elements per term, $M^{k|k-1}$ in (28), and mapping between each term and each element, B_k in (29), are only valid if all terms that can combine are combined. We can determine which terms of the *a priori* ucpdf combine at any step $k|k-1$ by using an indicator matrix that can be constructed recursively.

First, note that every element, ξ, contains a particular measurement z_k or the initial condition $\bar{x}_1$ and assume that they are ordered from newest to oldest. Then, for $k \geq 2$, we can construct an indicator matrix E_k similarly to B_k, composed of sub-matrices E_{ki}, for $i = 2, \ldots, k$. Let

$$A_{ki} = \begin{bmatrix} I_i \\ \mathbf{1}^{1\times i} \end{bmatrix} \triangleq \begin{bmatrix} \bar{A}_{ki} & a_{ki} \end{bmatrix}, \tag{36}$$

where a_{ki} is the last column of A_{ki}, I_i is the identity matrix of size $i \times i$ and $\mathbf{1}$ is a vector with elements all equal to 1. Then,

$$E_{ki} = \begin{bmatrix} \bar{A}_{ki} & a_{ki} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{A}_{ki} & \mathbf{0} & \cdots & a_{ki} \end{bmatrix}. \tag{37}$$

The matrix E_{ki} in (37) as shown is true for $k \geq i > 4$ for the sake of demonstrating the form. However, the pattern should be apparent for smaller k and i. Finally, we define the full indicator matrix

$$E_k = \begin{bmatrix} E_{k2} \\ E_{k3} \\ \vdots \\ E_{kk} \end{bmatrix}. \tag{38}$$

Terms corresponding to rows of E_k with the same elements will combine, which is done by adding their coefficient functions. Since the element is ordered, this simply involves summing the coefficients, ρ. We have found that, barring a numerical coincidence, only terms with one element will sum. Once all of the terms are combined, E_k becomes B_k.

6. Conditional mean and variance

Now that we have the *a priori* ucpdf, we can obtain the *a posteriori* ucpdf, $\bar{f}_{X_k|\mathbf{Y}_k}(x_k|\mathbf{y}_k)$, simply by multiplying $\bar{f}_V\left(\frac{z_k}{H} - x_k\right)$. We can then use it to compute the minimum variance estimate, or conditional mean, $\hat{x}_k$ and estimation error variance $\hat{\sigma}^2$, which are given as

$$\begin{aligned} \hat{x}_k &= E\left[X_k|Y_k\right] \\ \hat{\sigma}_k^2 &= E\left[(X_k - \hat{x}_k)^2|Y_k\right] = E\left[X_k^2|Y_k\right] - \hat{x}_k^2, \end{aligned} \tag{39}$$

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.

First, we need to normalize the *a posteriori* ucpdf $\bar{f}_{X_k|Y_k}$, which is given by (25) and (30), or more explicitly,

$$\begin{aligned} \bar{f}_{X_k|Y_k}(x_k|\mathbf{y}_k) = &\left[\sum_{i=1}^{N^{k|k-1}+1} g_i^{k|k-1}\epsilon_i^{k|k-1}\right] \\ &\times \frac{1}{2\gamma}\exp\left[-\frac{|H|}{\gamma}\left|\frac{z_k}{H} - x_k\right|\right], \end{aligned} \tag{40}$$

where $g_i^{k|k-1}$ and $\epsilon_i^{k|k-1}$ are given in (26). Since $\frac{1}{2\gamma}$ also appears in the normalization factor, we can pre-divide it simply by neglecting it at this stage. After re-indexing the terms, we get

$$\bar{f}_{X_k|Y_k}(x_k|\mathbf{y}_k) = \sum_{i=1}^{N^{k|k}} \bar{g}_i^{k|k}\epsilon_i^{k|k}, \tag{41}$$

where

$$\begin{aligned} \bar{g}_i^{k|k} &= \bar{\rho}_{i0}^{k|k} + \sum_{l=1}^{M_i^{k|k}} \bar{\rho}_{il}^{k|k}\mathrm{sgn}\left(\xi_{il}^{k|k} - x_k\right) \\ \epsilon_i^{k|k} &= \exp\left[-\sum_{l=1}^{M_i^{k|k-1}} \eta_{il}^{k|k-1}\left|\xi_{il}^{k|k-1} - x_k\right| - \frac{|H|}{\gamma}\left|\frac{z_k}{H} - x_k\right|\right] \\ &\triangleq \exp\left[-\sum_{l=1}^{M_i^{k|k}} \eta_{il}^{k|k}\left|\xi_{il}^{k|k} - x_k\right|\right]. \end{aligned} \tag{42}$$

The normalization of $\bar{f}_{X_k|Y_k}(x_k|\mathbf{y}_k)$, $f_{\mathbf{Y}_k}$, is given by

$$f_{\mathbf{Y}_k} = \int_{-\infty}^{\infty} \bar{f}_{X_k|Y_k}(x_k|\mathbf{y}_k)\,dx_k, \tag{43}$$

which can be evaluated using the integral formula (A.1) in Appendix. The first moment is defined as

$$E[X_k|\mathbf{Y}_k] = \int_{-\infty}^{\infty} x_k f_{X_k|\mathbf{Y}_k}\,dx_k = \frac{\int_{-\infty}^{\infty} x_k \bar{f}_{X_k|\mathbf{Y}_k}\,dx_k}{f_{\mathbf{Y}_k}}, \tag{44}$$

where $\bar{f}_{X_k|\mathbf{Y}_k}$ is given in (41) and $f_{\mathbf{Y}_k}$ is given in (43). Unlike the normalization factor, determining the first and second moments is more involved. The numerator of (44) can be integrated by parts. However, the intermediate steps can quickly become cumbersome. While we have performed the computation, it is difficult to present succinctly. Fortunately, we can avoid this complexity by making use of characteristic functions, which only require one integration using the integral formula (A.1) in the Appendix, followed by much simpler differentiation.

6.1. Characteristic function

For $\nu, x \in \mathbb{R}$, the characteristic function of $\bar{f}_{X_k|\mathbf{Y}_k}$ is

$$\bar{\phi}_{X_k|\mathbf{Y}_k}(\nu) = \int_{-\infty}^{\infty} e^{j\nu x_k}\bar{f}_{X_k|\mathbf{Y}_k}(x_k|\mathbf{y}_k)\,dx_k. \tag{45}$$

The derivation of the characteristic function can be found in Appendix A.2 of the Appendix, where it simply is an application of the same integral formula (A.1) that was used in propagation. Using (A.8) for each term of $\bar{f}^i_{X_k|\mathbf{Y}_k} \triangleq g_i\epsilon_i$, the corresponding term in the characteristic function is

$$\bar{\phi}^i_{X_k|\mathbf{Y}_k} = G_i(\nu)\mathcal{E}_i(\nu) \tag{46}$$

where, for $\delta_i(\cdot) = \sum_{\substack{l=0 \\ l\neq j}}^{M_i^{2|1}+1} (\cdot)\mathrm{sgn}\left(\xi_{il} - \xi_{ij}\right)$,

$$\begin{aligned} G_i(\nu) &= \frac{\rho_0 + \rho_i + \delta_i(\rho_l)}{j\nu + \eta_i + \delta_i(\eta_l)} - \frac{\rho_0 - \rho_i + \delta_i(\rho_l)}{j\nu - \eta_i + \delta_i(\eta_l)} \\ &\triangleq \frac{a_{i1}}{j\nu + a_{i2}} - \frac{b_{i1}}{j\nu + b_{i2}} \\ \mathcal{E}_i(\nu) &= \exp\left[-\sum_{\substack{l=1 \\ l\neq i}}^{M_i} \eta_l\left|\xi_l - \xi_i\right| + j\nu\xi_i\right] \\ &\triangleq \exp\left(c_{i1} + c_{i2}j\nu\right). \end{aligned} \tag{47}$$

6.2. Normalization

From (A.10) of the Appendix, the normalization factor $f_{\mathbf{Y}_k}$ is simply $f_{\mathbf{Y}_k} = \bar{\phi}_{X_k|\mathbf{Y}_k}(0)$. Therefore, using (46), we get

$$f_{\mathbf{Y}_k} = \sum_{i=1}^{m}\left(\frac{a_{i1}}{a_{i2}} - \frac{b_{i1}}{b_{i2}}\right)\exp\left(c_{i1}\right). \tag{48}$$

6.3. First moment and second moments at step k

The conditional mean and variance are given by $\hat{x}_k = E\left[X_k|\mathbf{Y}_k\right]$ and $\mathrm{Var}(\hat{x}_k) = E\left[X_k^2|\mathbf{Y}_k\right] - \hat{x}_k^2$, where

$$E\left[X_k|\mathbf{Y}_k\right] = \left[\frac{1}{j}\frac{d\phi_{X_k|\mathbf{Y}_k}(\nu)}{d\nu}\right]_{\nu=0}$$
$$E\left[X_k^2|\mathbf{Y}_k\right] = \left[-\frac{d^2\phi_{X_k|\mathbf{Y}_k}(\nu)}{d\nu^2}\right]_{\nu=0}. \tag{49}$$

Note that the moments are derived from the *normalized* characteristic function $\phi_{X_k|\mathbf{Y}_k}$, which is simply $\bar{\phi}_{X|Y}$ divided by $f_{\mathbf{Y}_k}$ in (48). Using the formulas in Appendix A.3 of the Appendix, the first and second moments of term i are given by

$$E_i[X_k|\mathbf{Y}_k] = \frac{1}{f_{\mathbf{Y}_k}}\sum_{i=1}^{m}\frac{1}{j}\left[\frac{\partial G_i}{\partial \nu}\mathcal{E}_i + G_i\frac{\partial \mathcal{E}_i}{\partial \nu}\right] \triangleq \frac{\bar{E}_i[X_k|\mathbf{Y}_k]}{f_{\mathbf{Y}_k}}$$
$$E_i[X_k^2|\mathbf{Y}_k] = -\frac{1}{f_{\mathbf{Y}_k}}\sum_{i=1}^{m}\left[\frac{\partial^2 G_i}{\partial \nu^2}\mathcal{E}_i + 2\frac{\partial G_i}{\partial \nu}\cdot\frac{\partial \mathcal{E}_i}{\partial \nu} + G_i\frac{\partial^2 \mathcal{E}_i}{\partial \nu^2}\right] \triangleq \frac{\bar{E}_i[X_k^2|\mathbf{Y}_k]}{f_{\mathbf{Y}_k}}, \tag{50}$$

where, $\bar{E}_i[X_k|\mathbf{Y}_k]$ and $\bar{E}_i[X_k^2|\mathbf{Y}_k]$ are defined as the unnormalized part of term i of the first and second moments, respectively. For brevity, the partial derivatives are shown in Appendix A.3 of Appendix.

6.4. Example: a posteriori conditional mean and variance of $f_{X_1|\mathbf{Y}_1}$

We examine the mean and variance of the Laplace estimator after one update step to see distinguishing features of the Laplace estimates compared with those of a Kalman filter and Cauchy estimator. Consider a generic example of the mean for the ucpdf after the first update, $\bar{f}_{X_1|\mathbf{Y}_1}$. For simplicity, let $H = 1$. From (13), $\bar{f}_{X_1|\mathbf{Y}_1} = \frac{1}{4\alpha\gamma}\exp\left(-\frac{1}{\gamma}|z_1 - x_1| - \frac{1}{\alpha}|x_1|\right)$. Then, from (46), the unnormalized characteristic function, $\bar{\phi}_{X_1|\mathbf{Y}_1}(\nu)$, is determined. Letting ν go to zero gives the normalization factor,

$$f_{\mathbf{Y}_1} = \frac{\alpha\exp\left(-\frac{1}{\alpha}|z_1|\right) - \gamma\exp\left(-\frac{1}{\gamma}|z_1|\right)}{2\left(\alpha^2-\gamma^2\right)}, \tag{51}$$

and from (50) evaluated at ν equal to zero, the normalized first moment is obtained as

$$\hat{x}_1 \triangleq \frac{\bar{E}[X_1|\mathbf{Y}_1]}{f_{\mathbf{Y}_1}} = \frac{\alpha z_1 - \frac{2\alpha^2\gamma^2\mathrm{sgn}(z_1)}{\alpha^2-\gamma^2}}{\alpha - \gamma\exp\left(-\frac{\alpha-\gamma}{\alpha\gamma}|z_1|\right)} + \frac{2\alpha^2\gamma^2\mathrm{sgn}(z_1)}{\alpha^2-\gamma^2}\cdot\frac{1}{\alpha\exp\left(\frac{\alpha-\gamma}{\alpha\gamma}|z_1|\right)-\gamma}. \tag{52}$$

Fig. 1 shows $\hat{x}_1$ 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, its effects should be attenuated. When they are equal, it splits the difference.

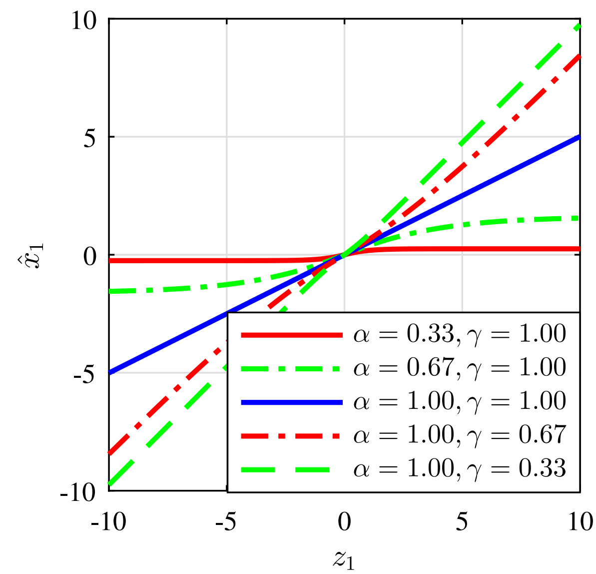

Fig. 1. Conditional mean $\hat{x}_1$ versus measurement z_1.

6.4.1. Example: a posteriori variance at $k = 1$

Consider the variance of $f_{X_1|\mathbf{Y}_1}$. Using (51) and (50), the second moment is then given by

$$E[X_1^2|\mathbf{Y}_1] = \frac{z_1^2\alpha - \frac{4\alpha^2\gamma^2|z_1|}{\alpha^2-\gamma^2} + \frac{2\alpha^3\gamma^2\left(\alpha^2+3\gamma^2\right)}{\left(\alpha^2-\gamma^2\right)^2}}{\alpha - \gamma\exp\left(-\frac{\alpha-\gamma}{\alpha\gamma}|z_1|\right)} - \frac{2\alpha^2\gamma^3\left(\gamma^2+3\alpha^2\right)}{\left(\alpha^2-\gamma^2\right)^2}\cdot\frac{1}{\alpha\exp\left(\frac{\alpha-\gamma}{\alpha\gamma}|z_1|\right)-\gamma}. \tag{53}$$

We examine some of the properties of $\mathrm{Var}(\hat{x}_1) = E[X_1^2|\mathbf{Y}_1] - \hat{x}_1^2$ at its extremes. Fig. 2 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 $\alpha = \gamma$.

It is also instructive to qualitatively compare the Laplace error variance with those of the Cauchy estimator and Kalman filter. Figs. 3 and 4 show the error variance of $\hat{x}_1$ as a function of measurement z_1 for these three estimators. We can see that the qualitative behavior of the Laplace variance approaches that of the Kalman filter as the difference between α and γ becomes 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 $\frac{1}{4}$ Idan & Speyer, 2014).

7. Numerical example for $k = 1, \ldots, 50$

We recursively and analytically computed the conditional pdf as well as the minimum variance state estimate and state estimation error variance for 50 steps using the parameters $\Phi = 0.9, H = 1, \alpha = 1/5, \beta = 1/4, \gamma = 1/3$. 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 un-modeled anomalies. For comparison, we used a Kalman filter to process the same measurement history. By a least-squares fit of a Gaussian pdf to a Laplace pdf, the standard deviation was set approximately equal to the corresponding Laplace spread parameter.

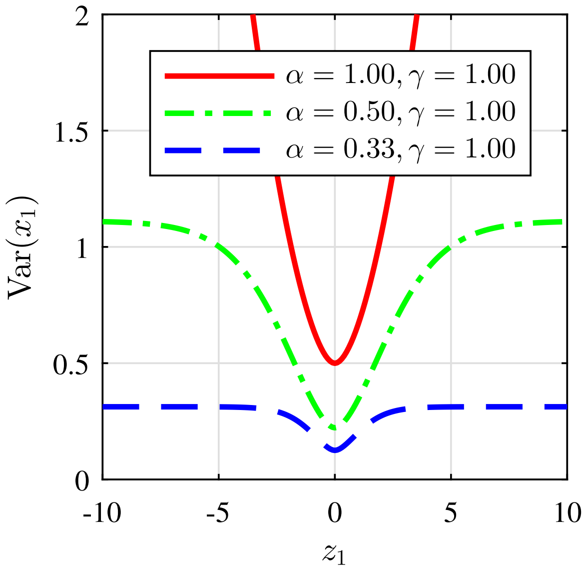

Fig. 2. Variance versus measurement z_1.

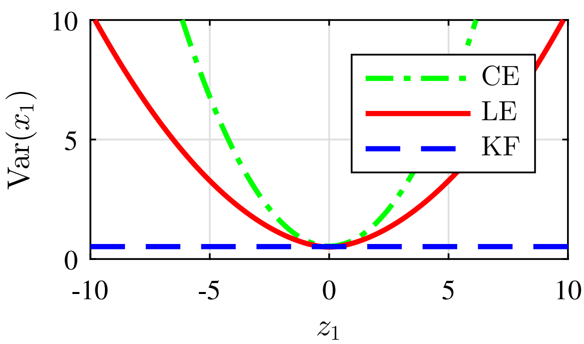

Fig. 3. Variance versus measurement z_1 for $\alpha = 1, \gamma = 1$.

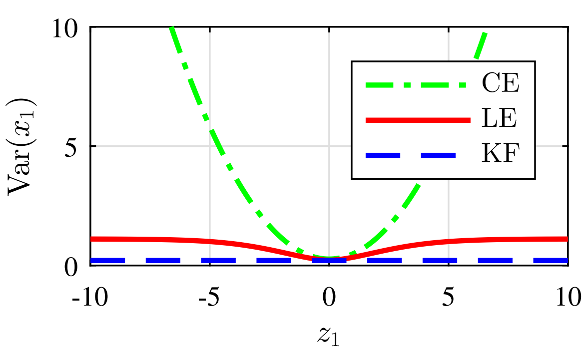

Fig. 4. Variance versus measurement z_1 for $\alpha = 0.5, \gamma = 1$.

Fig. 5 shows the measurement noise as well as the estimation errors and estimated standard deviation bounds for the Kalman filter and Laplace estimator. We can see that the Kalman estimate reacts strongly to the spikes, while its variance remains constant, thus severely underestimating the uncertainty. In contrast, the Laplace estimate appears to attenuate the measurement spikes and the estimated standard deviation increases to account for the increased uncertainty. This is similar to the behavior of the Cauchy estimator in Idan and Speyer (2012), which was shown to attenuate noise spikes in Cauchy-distributed measurement noise and increase its estimated error variance accordingly.

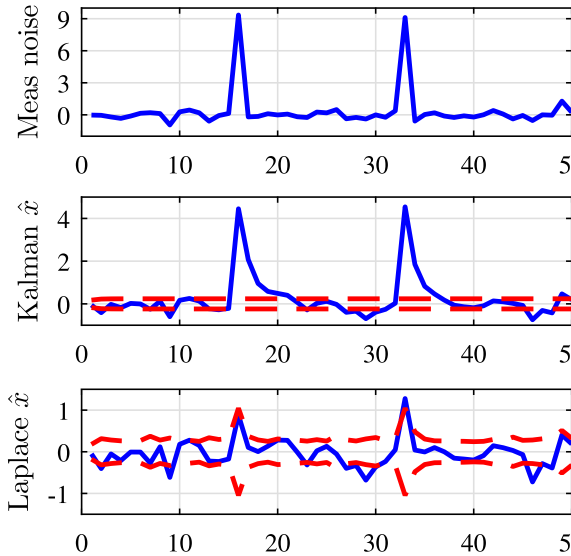

Fig. 5. Measurement noise and estimation error for the Kalman filter and Laplace estimator.

8. Conclusions

We have developed the minimum-variance estimator for a scalar discrete-time linear system with additive independent Laplace measurement and process noises. We used the characteristic function in lieu of evaluating the moments directly. Using these results, we implemented the estimator in MATLAB, demonstrating both novel, intrinsic and desirable non-linear properties that differ considerably from the Kalman filter. The results of this paper highlight promising avenues for future research into robust estimation and L_1 optimal control schemes.

Acknowledgments

This material is based on work supported by the National Science Foundation, USA under Grant Numbers NSF/ENG/ECCS-BSF 1607502 and 1934467 as well as the United States–Israel Binational Science Foundation (BSF), Israel joint NSF-BSF ECCS program under Grant No. 2015702.

Appendix

A.1. Key integral formula

A key integral formula was derived by Idan and Speyer in Appendix B of Idan and Speyer (2014) and is restated here for convenience.

$$I = \int_{-\infty}^{\infty} g(x) \exp\left[-\sum_{i=1}^{M} \eta_i \left|\xi_i - x\right| + jzx\right] dx$$
$$= \sum_{i=1}^{M} g_i \exp\left[-\sum_{\substack{l=1 \\ l \neq i}}^{M} \eta_l \operatorname{sgn}(\xi_l - \xi_i) + jz\xi_i\right], \tag{A.1}$$

where, for $\delta_i(\eta_l) = \sum_{\substack{l=1 \\ l \neq i}}^{M} \eta_l \operatorname{sgn}(\xi_l - \xi_i)$,

$$g_i = \frac{\rho_i + \delta_i(\rho_l)}{jz + \eta_i + \delta_i(\eta_l)} - \frac{-\rho_i + \delta_i(\rho_l)}{jz - \eta_i + \delta_i(\eta_l)}. \tag{A.2}$$

A.2. Scalar characteristic function

Given the form of the general scalar ucpdf in (25), the characteristic function is the expectation of $e^{j\nu x}$, or

$$\bar{\phi}_{X|\mathbf{Y}} = \int_{\mathbb{R}} e^{j\nu x}\bar{f}_{X|\mathbf{Y}}(x|\mathbf{y})\,dx \tag{A.3}$$

where

$$\begin{aligned} g_i(x) &= \rho_0 + \sum_{l=1}^{P} \rho_l \operatorname{sgn}(\xi_l - x) \\ \epsilon_i(x) &= \exp\left[-\sum_{l=1}^{M} \eta_l\,|\xi_l - x|\right]. \end{aligned} \tag{A.4}$$

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

$$\bar{\phi}^i_{X|\mathbf{Y}} = \int_{\mathbb{R}} g_i(x)\exp\left[-\sum_{l=1}^{M}\eta_l\,|\xi_l - x| + j\nu x\right] dx. \tag{A.5}$$

The solution to this integral is given by (A.1), or

$$\bar{\phi}^i_{X|\mathbf{Y}} = \sum_{i=1}^{n} G_i(\nu)\exp\left[-\sum_{\substack{l=1\\ l\neq i}}^{n}\eta_l \operatorname{sgn}(\xi_l - \xi_i) + j\nu\xi_i\right], \tag{A.6}$$

where, for $\delta_i(*) = \sum_{\substack{l=1\\ l\neq i}}^{n}(*)\operatorname{sgn}(\xi_l - \xi_i)$,

$$G_i = \frac{\rho_i + \delta_i(\rho_l)}{j\nu + \eta_i + \delta_i(\eta_l)} - \frac{-\rho_i + \delta_i(\rho_l)}{j\nu - \eta_i + \delta_i(\eta_l)}. \tag{A.7}$$

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

$$\begin{aligned} G_i(\nu) &= \frac{a_{i1}}{j\nu + a_{i2}} - \frac{b_{i1}}{j\nu + b_{i2}} \\ \epsilon_i(\nu) &= \exp(c_{i1} + c_{i2}\cdot j\nu) \end{aligned} \tag{A.8}$$

A.3. Scalar moments from characteristic function

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

$$\begin{aligned} f_{\mathbf{Y}} &= \bar{\phi}_{X|\mathbf{Y}}(0) \\ E[X|\mathbf{Y}] &= \left[\frac{1}{j}\frac{\partial\phi_{X|\mathbf{Y}}(\nu)}{\partial\nu}\right]_{\nu=0} \\ E[X^2|\mathbf{Y}] &= \left[-\frac{\partial^2\phi_{X|\mathbf{Y}}(\nu)}{\partial\nu^2}\right]_{\nu=0}. \end{aligned} \tag{A.9}$$

Using the generic form for each term, the ith term of the normalization of $\bar{f}_{X|\mathbf{Y}}$ is

$$f_{\mathbf{Y},i} = \left[\phi^i_{X|\mathbf{Y}}(0)\right] = \left(\frac{a_{i1}}{a_{i2}} - \frac{b_{i1}}{b_{i2}}\right)\exp(c_{i1}). \tag{A.10}$$

The ith term of the first moment is $E_i[X|\mathbf{Y}] = \left[\frac{1}{j}\frac{\partial\phi_{X|\mathbf{Y},i}(\nu)}{\partial\nu}\right]_{\nu=0}$, or

$$E_i[X|\mathbf{Y}] = \frac{1}{j}\left[\frac{\partial G_i}{\partial\nu}\epsilon_i + G_i\frac{\partial\epsilon_i}{\partial\nu}\right]_{\nu=0}, \tag{A.11}$$

where

$$\begin{aligned} \left[\frac{\partial G_i}{\partial\nu}\right]_{\nu=0} &= -\frac{a_{i1}\cdot j}{(a_{i2})^2} + \frac{b_{i1}\cdot j}{(b_{i2})^2}, \\ \left[\frac{\partial\epsilon_i}{\partial\nu}\right]_{\nu=0} &= jc_{i2}\exp(c_{i1}). \end{aligned} \tag{A.12}$$

The ith term of the second moment is

$$\begin{aligned} E_i[X^2|\mathbf{Y}] &= \left[-\frac{\partial^2\phi_{X|\mathbf{Y}}(\nu)}{\partial\nu^2}\right]_{\nu=0} \\ &= -\left[\frac{\partial^2 G_i}{\partial\nu^2}\epsilon_i + 2\frac{\partial G_i}{\partial\nu}\cdot\frac{\partial\epsilon_i}{\partial\nu} + G_i\frac{\partial^2\epsilon_i}{\partial\nu^2}\right]_{\nu=0} \end{aligned} \tag{A.13}$$

where

$$\begin{aligned} \left[\frac{\partial^2 G_i}{\partial\nu^2}\right]_{\nu=0} &= -\frac{2a_{i1}}{(a_{i2})^3} + \frac{2b_{i1}}{(b_{i2})^3}, \\ \left[\frac{\partial^2\epsilon_i}{\partial\nu^2}\right]_{\nu=0} &= -c_{i2}^2\exp(c_{i1}). \end{aligned} \tag{A.14}$$

References

Bryson, M. C. (1974). Heavy-tailed distributions: Properties and tests. *Technometrics*, *16*(1), 61–68.

Duong, N., Idan, M., Pinchasi, R., & Speyer, J. (2021). A note on hyper-plane arrangements in $\mathbb{R}^d$. *Discrete Mathematics Letters*, *7*(1), 79–85.

Duong, N., Speyer, J., Yoneyama, J., & Idan, M. (2018). Laplace Estimator for linear scalar systems. In *57th IEEE conference on decision and control*. Miami, FL, USA.

Farokhi, F., Milosevic, J., & Sandberg, H. (2016). Optimal state estimation with measurements corrupted by Laplace noise. In *2016 IEEE 55th conference on decision and control (CDC)* (pp. 302–307).

Idan, M., & Speyer, J. L. (2010). Cauchy estimation for linear scalar systems. *IEEE Transactions on Automatic Control*, *55*(6), 1329–1342.

Idan, M., & Speyer, J. L. (2012). State estimation for linear scalar dynamic systems with additive Cauchy noises: Characteristic function approach. *SIAM Journal on Control and Optimization*, *50*(4), 1971–1994.

Idan, M., & Speyer, J. L. (2014). Multivariate Cauchy estimator with scalar measurement and process noises. *SIAM Journal on Control and Optimization*, *52*(2), 1108–1141.

Kozierski, P., Sadalla, T., Owczarkowski, A., & Drgas, S. (2016). Particle filter in multidimensional systems. In *2016 21st international conference on methods and models in automation and robotics (MMAR)* (pp. 806–810).

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

Laska, B. N. M., Bolic, M., & Goubran, R. A. (2010). Particle filter enhancement of speech spectral amplitudes. *IEEE Transactions on Audio, Speech, and Language Processing*, *18*(8), 2155–2167.

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

Musso, C., Bui Quang, P., & Le Gland, F. (2011). Introducing the Laplace approximation in particle filtering. In *14th international conference on information fusion* (pp. 1–8).

Rabbani, H., Vafadust, M., & Gazor, S. (2006). 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).

Selesnick, I. W. (2008). The estimation of Laplace random vectors in additive white Gaussian noise. *IEEE Transactions on Signal Processing*, *56*(8), 3482–3496.

Nhattrieu C. Duong received the B.S. degree in Mechanical Engineering from the California Institute of Technology in Pasadena, CA in 2007. He received the M.S. and Ph.D. degrees in Aerospace Engineering from the University of California, Los Angeles in Los Angeles, CA in 2008 and 2020, respectively. He is currently a Navigation Engineer at Blue Origin, LLC.

Jason L. Speyer (M'71–SM'82–F'85–LF'05) received the B.S. in aeronautics and astronautics from MIT, Cambridge, in 1960 and the Ph.D. in applied mathematics from Harvard University, Cambridge, MA, in 1968. He was awarded an Honorary Doctorate from the Technion in 2013. He is the Ronald and Valerie Sugar Distinguished Professor in the Mechanical and Aerospace Engineering Department and the Electrical Engineering Department, UCLA. He coauthored, with W. H. Chung, Stochastic Processes, Estimation, and Control (SIAM, 2008), and coauthored, with D. H. Jacobson, Primer on Optimal Control Theory (SIAM, 2010). He served as Associate Editor for Technical Notes and Correspondence (1975–1976) and Stochastic Control (1978–1979), IEEE Transactions on Automatic Control, for AIAA Journal of Guidance and Control (1977–1978), and for Journal of Optimization Theory and Applications (1981–present). He is a fellow of the IEEE, an Honorary fellow of the AIAA, and was awarded the AIAA Mechanics and Control of Flight Award, AIAA Dryden Lectureship in Research, Air Force Exceptional Civilian Decoration (1991 and 2001), IEEE Third Millennium Medal, AIAA Guidance, Navigation, and Control Award, Richard E. Bellman Control Heritage Award, and membership in the National Academy of Engineering.

Moshe Idan received his B.Sc. and M.Sc. degrees from the Faculty of Aerospace Engineering at the Technion in 1983 and 1986, respectively. He earned his Ph.D. from the Department of Aerospace Engineering at Stanford University in 1990. Since 1991, Dr. Idan has been with the Faculty of Aerospace Engineering at the Technion. In 2000–2001 he was a visiting research scholar at the School of Aerospace Engineering at Georgia Tech, and in 2007–2009 he spent a two year sabbatical at the Department of Mechanical and Aerospace Engineering in UCLA. His current research interests include robust, nonlinear and adaptive flight control system design techniques and applications, control and estimation of systems with non-Gaussian noises.