Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Cauchy Estimation for Linear Scalar Systems

Moshe Idan*
Mechanical and Aerospace Engineering
University of California, Los Angeles
Los Angeles, California 90095–1597
Email: moshe.idan@ucla.edu

Jason L. Speyer
Mechanical and Aerospace Engineering
University of California, Los Angeles
Los Angeles, California 90095–1597
Email: speyer@seas.ucla.edu

Abstract*—An estimation paradigm is presented for scalar discrete linear systems entailing additive process and measurement noises that have Cauchy probability density functions (pdf). For systems with Gaussian noises, the Kalman filter has been the main estimation paradigm. However, many practical system uncertainties that have impulsive character, such as radar glint, are better described by stable non-Gaussian densities, for example, the Cauchy pdf. Although the Cauchy pdf does not have a defined mean and does have an infinite second moment, the conditional density of a Cauchy random variable, given its linear measurements with an additive Cauchy noise, has a conditional mean and a *finite* conditional variance, both being functions of the measurement. For a single measurement, simple expressions are obtained for the conditional mean and variance, by deriving closed form expressions for the infinite integrals associated with the minimum variance estimation problem. To alleviate the complexity of the multi-stage estimator, the conditional pdf is represented in a special factored form. A recursion scheme is then developed based on this factored form and closed form integrations, allowing for the propagation of the conditional mean and variance over an arbitrary number of time stages. In simulations, the performance of a newly developed scalar discrete-time Cauchy estimator is significantly superior to a Kalman filter in the presence of Cauchy noise, whereas the Cauchy estimator deteriorates only slightly compared to the Kalman filter in the presence of Gaussian noise. Remarkably, this new recursive Cauchy conditional mean estimator has parameters that are generated by *linear* difference equations with stochastic coefficients, providing computational efficiency.

I. INTRODUCTION

Many engineering applications entail random processes or noises that have significant deviations not captured by commonly assumed Gaussian probability distributions. For example, atmospheric and underwater acoustic noises, that are the governing noise types in radar and sonar applications, are non-Gaussian and have a very impulsive character [1]. The class of probabilistic models used to represent impulsive noises is called stable non-Gaussian or symmetric alpha-stable ($S\alpha$-S) distributions (see [2] for a comprehensive treatment of $S\alpha$-S densities). In this class, the Gaussians correspond to the α=2 case, whereas α=1 leads to the Cauchy probability density function (pdf).

It has been shown that a Cauchy pdf better characterizes such impulsive types of sensor noise than the Gaussian. For example, a $S\alpha$-S pdf was used to detect a radar signal in clutter, showing that the in-phase component of radar clutter time series agree extremely well with a $S\alpha$-S pdf with α=1.7 [3]. In that study both the maximum likelihood Gaussian (MLG) and Cauchy (MLC) detectors were developed. The result is that for all $\alpha\in[1, 2]$ the MLC is very close to the Cramer-Rao bound, whereas the MLG deviates significantly as α goes from 2 to 1. Similar results for a multi-user communication network are reported in [4]. There it is noted that the Gaussian and Cauchy detectors have similar complexities and there is no computational advantage in preferring one over the other. In addition it was observed that the Cauchy receiver is robust in the entire class of $S\alpha$-S noises, while the Gaussian receiver degrades very quickly even if the Gaussian assumption is only slightly violated. Moreover, this performance degradation is more serious at low probabilities of false alarm, at which any realistic detector would work.

The above observations about the nature of the impulsive noises and the robustness characteristics of the Cauchy detectors motivated the development of sequential estimators for linear systems with additive Cauchy noises presented in this study. Such heavy tailed pdf does not posses well defined and/or finite moments of any order [5], which make the associated estimation problem significantly more challenging compared to the Gaussian case.

There are many surprises when determining the properties of the conditional pdf for linear system dynamics with additive Cauchy random variables. The first surprise is that the conditional pdf of the state given the measurement is *not* Cauchy and has a finite mean and variance. Secondly, the usual difficulty in the implementation of the recursion for the conditional pdf is the evaluation of the infinite integrals required in the time propagation and normalization in the measurement update [6]. However, when formulated in a particular factored form, it is shown that these integrals can be determined in closed form by using, for example, the Cauchy residual theorem [7]. Therefore, the conditional pdf as well as the conditional mean and variance can be propagated with computational ease, which is remarkable for a pdf other than the Gaussian.

The paper is organized as follows. The problem is formulated in section II. In section III a single measurement update and time propagation are determined in a simple analytic form.

*On a sabbatical leave from the Faculty of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa, Israel, 32000. Email: moshe.idan@technion.ac.il.

Section IV presents a special factorization, used to derive a recursion for generating the conditional pdf and to compute the infinite integrals of the conditional mean and variance. For a particular dynamic system, the recursion is implemented and the results are presented in section V. The performance of this estimator in the presence of heavy tailed Cauchy measurement and process noises is examined over a large number of sample times, demonstrating its superiority over the classical Kalman filter. In section VI we conclude.

II. Problem Statement

We consider a scalar linear, discrete, time invariant dynamic system given by

$$x(k+1) = \Phi x(k) + \Gamma u(k) + w(k) \quad (1)$$

$$z(k) = Hx(k) + v(k) \quad (2)$$

where x, u and z are, respectively, the state, known input, and measurement. w and v are uncorrelated, white process and measurement noises, with Cauchy probability density functions (pdf) given by

$$f_{W_k}\big(w(k)\big) = \frac{\beta/\pi}{w^2(k) + \beta^2} \quad (3)$$

$$f_{V_k}\big(v(k)\big) = \frac{\gamma/\pi}{v^2(k) + \gamma^2} \quad (4)$$

β and γ are known Cauchy scaling parameters. The initial condition of x is assumed to be also Cauchy distributed with

$$f_{X_0}\big(x(0)\big) = \frac{\alpha/\pi}{\big(x(0) - \bar{x}_0\big)^2 + \alpha^2} \quad (5)$$

with a known positive scaling parameter α and median $\bar{x}_0$.

The goal is to design an optimal, minimum variance estimator of $x(k)$ given measurements $z(i), i = 0, 1, \ldots, k$.

III. One Step Update

To better understand the Cauchy estimator, one measurement update followed by a time propagation step is examined.

A. Measurement Update

Given the initial distribution of $x(0)$, Eq. (5), and a measurement $z(0)$, the minimum variance estimate of $x(0)$ is the conditional expectation, $E\big[x(0)|z(0)\big]$, determined from the conditional density function $f_{X_0|Z_0}\big(x(0)|z(0)\big)$, constructed as

$$f_{X_0|Z_0}\big(x(0)|z(0)\big) = \frac{f_{X_0,Z_0}\big(x(0), z(0)\big)}{f_{Z_0}\big(z(0)\big)} \quad (6)$$

The joint probability $f_{X_0,Z_0}\big(x(0), z(0)\big)$, computed using measurement Eq. (2) and noise pdf of Eq. (4), is

$$f_{X_0,Z_0}\big(x(0), z(0)\big) = f_{X_0}\big(x(0)\big)\, f_{V_0}\big(z(0) - Hx(0)\big) \quad (7)$$

The marginal probability $f_{Z_0}\big(z(0)\big)$ is obtained by

$$f_{Z_0}\big(z(0)\big) = \int_{-\infty}^{\infty} f_{X_0,Z_0}\big(x(0), z(0)\big)\, dx(0) \quad (8)$$

Solving analytically the above integral yields

$$f_{X_0|Z_0}\big(x(0)|z(0)\big) = \quad (9)$$

$$\frac{\alpha\gamma/\pi}{|H|\alpha + \gamma} \frac{\big(z(0) - H\bar{x}_0\big)^2 + \big(|H|\alpha + \gamma\big)^2}{\Big[\big(x(0) - \bar{x}_0\big)^2 + \alpha^2\Big]\Big[\big(z(0) - Hx(0)\big)^2 + \gamma^2\Big]}$$

Examining the conditional pdf of Eq. (9) reveals that although the original pdf in Eq. (5) had no well defined mean and an infinite second moment, the former has two computable finite moments. This implies that a minimum variance estimate of $x(0)$ given the measurement $z(0)$ can be determined as

$$\hat{x}(0) = E\big[x(0)|z(0)\big] = \int_{-\infty}^{\infty} x(0) f_{X_0|Z_0}\big(x(0)|z(0)\big)\, dx(0) \quad (10)$$

After a lengthy analytical manipulation, the conditional mean computation results in a strikingly simple relation

$$\hat{x}(0) = \bar{x}_0 + \frac{\alpha\, \mathrm{sign}(H)}{|H|\alpha + \gamma} \big(z(0) - H\bar{x}_0\big) \quad (11)$$

Similarly, one can determine the minimum variance of the estimation error $\tilde{x}(0) = x(0) - \hat{x}(0)$ as

$$E\left[\tilde{x}^2(0)|z(0)\right] = \frac{\alpha\gamma}{|H|} \left[\frac{\big(z(0) - H\bar{x}_0\big)^2}{\big(|H|\alpha + \gamma\big)^2} + 1\right] \quad (12)$$

Note that the estimate is linear in the measurement, resembling a measurement update equation of the Kalman filter, with a different gain multiplying the difference $z(0) - H\bar{x}_0$. Contrary to the Kalman filter, the variance in the Cauchy case depends on the measurement: it is proportional to the square of the difference between the measurement $z(0)$ and its median value $H\bar{x}_0$ that would have been attained with no measurement noise and a median value $\bar{x}_0$ for $x(0)$. In particular, a large variance would be obtained when $z(0) - H\bar{x}_0$ is large.

Furthermore, $f_{X_0|Z_0}\big(x(0)|z(0)\big)$ of Eq. (9) has a fourth order denominator with two pairs of complex conjugate roots at $\bar{x}_0 \pm j\alpha$ and $z(0)/H \pm j\gamma/|H|$, while the initial pdf $f_{X_0}\big(x(0)\big)$ of Eq. (5) had only one pair at $\bar{x}_0 \pm j\alpha$. This implies that a measurement update increases the order of the pdf denominator by two, adding to it a conjugate pair of roots with measurement dependent values $z(0)/H \pm j\gamma/|H|$. This fact is exploited in the subsequent derivations.

A measurement updated pdf is shown in Fig. 1 for parameter values stated on the plot. It is interesting to note that unlike in the Gaussian, the conditional pdf $f_{X_0|Z_0}\big(x(0)|z(0)\big)$ is non-symmetric and non-unitary, the extent of which is determined by the size of the difference $z(0) - H\bar{x}_0$.

B. Time Propagation

In the time propagation step we construct the conditional pdf of the state at time k=1 given the measurement $z(0)$, i.e.,

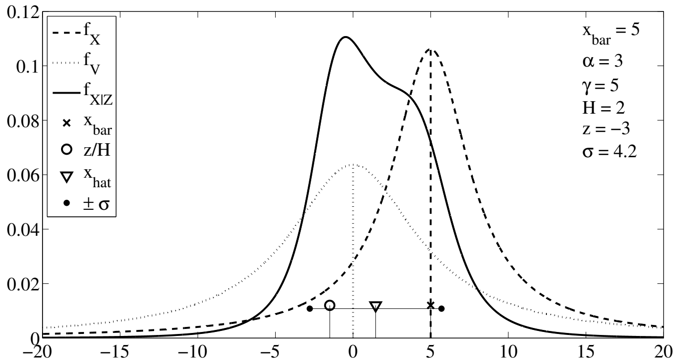


Fig. 1. One step measurement update.

$f_{X_1|Z_0}\big(x(1)|z(0)\big)$. Using the Chapman-Kolmogorov equation, this conditional pdf can be computed by

$$f_{X_1|Z_0}\big(x(1)|z(0)\big) = \int_{-\infty}^{\infty} f_{X_1|X_0}\big(x(1)|x(0)\big)\, f_{X_0|Z_0}\big(x(0)|z(0)\big)\, dx(0) \quad (13)$$

In this expression, $f_{X_0|Z_0}\big(x(0)|z(0)\big)$ is given in Eq. (9). $f_{X_1|X_0}\big(x(1)|x(0)\big)$ is computed using the dynamics Eq. (1) and the process noise pdf of Eq. (3):

$$f_{X_1|X_0}\big(x(1)|x(0)\big) = f_{W_0}\big(x(1) - \Phi x(0) - \Gamma u(0)\big) \quad (14)$$

Substituting Eqs. (3), (9) and (14) into Eq. (13) and integrating yields the conditional pdf

$$f_{X_1|Z_0}\big(x(1)|z(0)\big) = \frac{A\big(x(1) - B\big)^2 + C^2}{\Big[\big(x(1) - \bar{x}_{11}\big)^2 + \alpha_1^2\Big]\Big[\big(x(1) - \bar{x}_{12}\big)^2 + \alpha_2^2\Big]} \quad (15)$$

where

$$\bar{x}_{11} = \Phi\bar{x}_0 + \Gamma u(0) \qquad \alpha_1 = |\Phi|\alpha + \beta \quad (16a)$$

$$\bar{x}_{12} = \frac{\Phi}{H}z(0) + \Gamma u(0) \qquad \alpha_2 = \left|\frac{\Phi}{H}\right|\gamma + \beta \quad (16b)$$

and A, B and C are complicated functions of the various system parameters and the signals $z(0)$ and $u(0)$. Since the exact expressions do not provide valuable information or insight, they are not detailed in this exposition.

Note that this time propagated pdf has a fourth order denominator, similar to the pdf we obtained after the measurement update in Eq. (9). Two major difference can be observed though. First, the values of the denominator roots have been changed to $\bar{x}_{1i} \pm j\alpha_i,\ i=1,2$. Second, the numerator of the pdf in Eq. (15) is second order in $x(1)$, indicating that this pdf, similar to a Cauchy pdf, has an undefined mean and an infinite second moment. This implies that contrary to the Gaussian case, one cannot compute a prior to measurement $z(1)$ estimate of the state $x(1)$ given $z(0)$ only.

Carrying out an additional measurement update with $z(1)$ when starting with the propagated fourth order pdf of Eq. (15) proved extremely complex and unpractical. Therefore, we have adopted an alternative approach, that builds on the facts learned from the one step measurement update and time propagation presented above.

IV. The Cauchy Estimator

A. pdf Factorization

Motivated by the structure of the rational pdfs encountered during the one step measurement update and time propagation process, for the derivation of a sequential estimator, at step k, just before the update using the measurement $z(k)$ is performed, we represent the conditional pdf of $x(k)$ given past data $y(k-1)$ in a factored form given by

$$f_{X_k|Y_{k-1}}\big(x(k)|y(k-1)\big) = \sum_{i=1}^{k+1} \frac{a_i(k|k-1)x(k) + b_i(k|k-1)}{\big(x(k) - \sigma_i(k|k-1)\big)^2 + \omega_i^2(k|k-1)} \quad (17)$$

where $y(k-1)$ denotes the measurement history from time step 0 to $k-1$, i.e., $y(k-1) = [z(0) \ \cdots \ z(k-1)]$. At $k=0$, the above pdf represents the a-priori distribution of $x(0)$ given in Eq. (5). Hence at $k=0$ there is only one term in the sum, with $a_1(0|-1)=0$, $b_1(0|-1)=\alpha/\pi$, $\sigma_1(0|-1)=\bar{x}_0$, and $\omega_1(0|-1)=\alpha$. In the consequent steps, the parameters $a_i(k|k-1)$, $b_i(k|k-1)$, $\sigma_i(k|k-1)$, and $\omega_i(k|k-1)$ will be functions of the measurements $y(k-1)$ as is detailed below.

This pdf changes during measurement update and time propagation steps, as presented in section III. Each measurement update increases the pdf denominator order by two, adding a pair of complex measurement dependent roots to it. This is represented by an additional term in the pdf sum of Eq. (17), while the rest of the pdf poles remains unchanged. All the numerator parameters $a_i(k|k-1)$ and $b_i(k|k-1)$ will change during the measurement update. The time propagation does not change the number of terms in the pdf sum, but affect all its parameters.

B. Measurement Update

Following the steps in section III-A, we compute

$$f_{X_k|Y_k}\big(x(k)|y(k)\big) = \frac{f_{X_k,Z_k|Y_{k-1}}\big(x(k), z(k)|y(k-1)\big)}{f_{Z_k|Y_{k-1}}\big(z(k)|y(k-1)\big)} \quad (18)$$

The density function in the above numerator is computed as

$$\begin{aligned} &f_{X_k,Z_k|Y_{k-1}}\big(x(k), z(k)|y(k-1)\big) \\ &= f_{X_k|Y_{k-1}}\big(x(k)|y(k-1)\big)\, f_{V_k}\big(z(k) - Hx(k)\big) \\ &= \sum_{i=1}^{k+1} \frac{a_i(k|k-1)x(k) + b_i(k|k-1)}{\big(x(k) - \sigma_i(k|k-1)\big)^2 + \omega_i^2(k|k-1)} \times \\ &\qquad \frac{\gamma/\pi}{\big(z(k) - Hx(k)\big)^2 + \gamma^2} \end{aligned} \quad (19)$$

Using partial fraction expansions on the terms in the above sum, we obtain

$$f_{X_k,Z_k|Y_{k-1}}\big(x(k),z(k)|y(k-1)\big) = \sum_{i=1}^{k+2} \frac{\bar{a}_i(k|k)x(k)+\bar{b}_i(k|k)}{\big(x(k)-\sigma_i(k|k)\big)^2+\omega_i^2(k|k)} \quad (20)$$

$\sigma_i(k|k){=}\sigma_i(k|k{-}1)$, $\omega_i(k|k){=}\omega_i(k|k{-}1)$, $i{=}1,\ldots,k{+}1$ represent the unchanged roots of the pdf denominator, while $\sigma_{k+2}(k|k){=}z(k)/H$ and $\omega_{k+2}(k|k){=}\gamma/|H|$ represent its new complex measurement dependent roots. $\bar{a}_i(k|k)$ and $\bar{b}_i(k|k)$ are computed from the partial fraction expansions as

$$\begin{Bmatrix}\bar{a}_i(k|k)\\ \bar{b}_i(k|k)\end{Bmatrix} = F_i(k)\begin{Bmatrix}a_i(k|k{-}1)\\ b_i(k|k{-}1)\end{Bmatrix}, \quad i=1,\ldots,k+1 \quad (21)$$

where

$$F_i(k) = \frac{1}{\Delta_i(k)}\times \begin{bmatrix} \delta_i(k)-\frac{\sigma(k|k)}{\omega(k|k)}\theta_i(k) & -\frac{1}{\omega(k|k)}\theta_i(k) \\ \frac{\sigma^2(k|k)+\omega^2(k|k)}{\omega(k|k)}\theta_i(k) & \delta_i(k)+\frac{\sigma(k|k)}{\omega(k|k)}\theta_i(k) \end{bmatrix} \quad (22)$$

$$\delta_i(k) = \big(\sigma_{k+2}(k|k)-\sigma_i(k|k)\big)^2+\omega_{k+2}^2(k|k)-\omega_i^2(k|k) \quad (23a)$$

$$\theta_i(k) = 2\omega_i(k|k)\big(\sigma_i(k|k)-\sigma_{k+2}(k|k)\big) \quad (23b)$$

$$\Delta_i(k) = \frac{\pi\gamma}{\omega_{k+2}^2(k|k)}\big(\delta_i^2(k)+\theta_i^2(k)\big) \quad (23c)$$

Note that the matrices $F_i(k)$, $i{=}1,\ldots,k{+}1$ are not functions of the numerator parameters $a_i(k|k{+}1)$ and $b_i(k|k{+}1)$, generating a linear, time-dependent, stochastic update equation (21). The numerator parameters of the new term in the pdf sum are

$$\bar{a}_{k+2}(k|k) = \sum_{i=1}^{k+1}\bar{a}_{i,k+2}(k|k) \quad (24)$$

$$\bar{b}_{k+2}(k|k) = \sum_{i=1}^{k+1}\bar{b}_{i,k+2}(k|k) \quad (25)$$

where for $i{=}1,\ldots,k{+}1$

$$\begin{Bmatrix}\bar{a}_{i,k+2}(k|k)\\ \bar{b}_{i,k+2}(k|k)\end{Bmatrix} = F_{i,k+2}(k)\begin{Bmatrix}\bar{a}_i(k|k)\\ \bar{b}_i(k|k)\end{Bmatrix} \quad (26a)$$

and

$$F_{i,k+2}(k) = -\begin{bmatrix} 1 & 0\\ 2\big(\sigma_i(k|k)-\sigma_{k+2}(k|k)\big) & 1\end{bmatrix}. \quad (26b)$$

Next, $f_{Z_k|Y_{k-1}}\big(z(k)|y(k{-}1)\big)$ is determined by integrating the pdf in Eq. (20) with respect to $x(k)$. Due to the modular structure of this pdf, the integral is evaluated analytically as

$$f_{Z_k|Y_{k-1}}\big(z(k)|y(k-1)\big) = \pi\sum_{i=1}^{k+2}\frac{\bar{a}_i(k|k)\sigma_i(k|k)+\bar{b}_i(k|k)}{\omega_i(k|k)} \quad (27)$$

Finally, the conditional pdf in Eq. (18) is obtained by dividing the result in Eq. (20) by the one in Eq. (27), yielding

$$f_{X_k|Y_k}\big(x(k)|y(k)\big) = \sum_{i=1}^{k+2}\frac{a_i(k|k)x(k)+b_i(k|k)}{\big(x(k)-\sigma_i(k|k)\big)^2+\omega_i^2(k|k)} \quad (28)$$

where $a_i(k|k)$ and $b_i(k|k)$ are obtained by dividing $\bar{a}_i(k|k)$ and $\bar{b}_i(k|k)$ by the results in Eq. (27), i.e., for $i{=}1,\ldots,k{+}2$

$$a_i(k|k) = \bar{a}_i(k|k)/f_{Z_k|Y_{k-1}}\big(z(k)|y(k-1)\big) \quad (29a)$$

$$b_i(k|k) = \bar{b}_i(k|k)/f_{Z_k|Y_{k-1}}\big(z(k)|y(k-1)\big). \quad (29b)$$

Eq. (28) indicates that the pdf structure defined in Eq. (17) is maintained after a measurement update. Moreover, one can see the explicit dependence of updated pdf parameters on the measurement $z(k)$. Firstly, $\sigma_{k+2}(k|k){=}z(k)/H$. In addition, all the numerator parameters $a_i(k|k)$ and $b_i(k|k)$, $i{=}1,\ldots,k+2$, depend on $z(k)$ though the dependence of the matrices $F_i(k)$ and $F_{i,k+2}(k)$ on $\sigma_{k+2}(k|k)$.

The modular form of Eq. (28) allows also an efficient computation of its first two moments, the former yielding the state estimate $\hat{x}(k|k)$. Those are given by

$$\hat{x}(k|k) = \pi\sum_{i=1}^{k+2}\frac{a_i\left(\sigma_i^2-\omega_i^2\right)+b_i\sigma_i}{\omega_i} \quad (30)$$

$$E\left(x^2|y(k)\right) = \pi\sum_{i=1}^{k+2}\frac{\left(a_i\sigma_i+b_i\right)\left(\sigma_i^2-\omega_i^2\right)-2a_i\sigma_i\omega_i^2}{\omega_i} \quad (31)$$

The error variance is obtained by

$$E\left(\tilde{x}^2|y(k)\right) = E\left(x^2|y(k)\right)-\hat{x}^2(k|k) \quad (32)$$

Note that in Eqs. (30) and (31) the time tag $(k|k)$ of the pdf parameters is dropped for simplicity.

C. Time Propagation

Consider the time propagation from step k to step $k{+}1$. Given the conditional pdf in Eq. (28) our goal is to construct $f_{X_{k+1}|Y_k}\big(x(k{+}1)|y(k)\big)$. Similarly to the computation in section III-B, using the Chapman-Kolmogorov equation we have

$$f_{X_{k+1}|Y_k}\big(x(k{+}1)|y(k)\big) = \int_{-\infty}^{\infty} f_{X_{k+1}|X_k}\big(x(k{+}1)|x(k)\big)\, f_{X_k|Y_k}\big(x(k)|y(k)\big)\,dx(k) \quad (33)$$

where

$$f_{X_{k+1}|X_k}\big(x(k{+}1)|x(k)\big) = f_{W_k}\big(x(k{+}1)-\Phi x(k)-\Gamma u(k)\big) = \frac{\beta/\pi}{\big(x(k{+}1)-\Phi x(k)-\Gamma u(k)\big)^2+\beta^2} \quad (34)$$

Using Eqs. (28) and (34), the integral in Eq. (33) can be computed for each term in the sum analytically, leading to

$$f_{X_{k+1}|Y_k}\big(x(k{+}1)|y(k)\big) = \sum_{i=1}^{k+2}\frac{a_i(k{+}1|k)x(k{+}1)+\,b_i(k{+}1|k)}{\big(x(k{+}1)-\sigma_i(k{+}1|k)\big)^2+\,\omega_i^2(k{+}1|k)} \quad (35)$$

where for $i{=}1,\ldots,k{+}2$

$$\sigma_i(k{+}1|k) = \Phi\sigma_i(k|k) + \Gamma u(k) \tag{36}$$

$$\omega_i(k{+}1|k) = |\Phi|\omega_i(k|k) + \beta \tag{37}$$

$$\begin{Bmatrix} a_i(k{+}1|k) \\ b_i(k{+}1|k) \end{Bmatrix} = G_i(k) \begin{Bmatrix} a_i(k|k) \\ b_i(k|k) \end{Bmatrix} \tag{38}$$

with

$$G_i(k) = \begin{bmatrix} \text{sign}(\Phi) & 0 \\ \dfrac{\sigma_i(k|k)\beta}{\omega_i(k|k)} - \text{sign}(\Phi)\Gamma u(k) & \dfrac{\omega_i(k{+}1|k)}{\omega_i(k|k)} \end{bmatrix} \tag{39}$$

Again we observe that the factored pdf structure of Eq. (17) is maintained also after the time propagation in Eq. (35).

D. Discussion

The measurement update and time propagation procedures presented above can be repeated for any number of steps, while maintaining the factored form of the pdfs. The measurement updated extends the pdf, increasing the order of its denominator and thus adding an additional term in the factored sum. The time propagation modifies the parameters in the sum, without increasing its size. This type of a recursion will lead to a constantly increasing number of terms in the pdf sum. To overcome this continuous growth, we analyze whether the number of elements in this sum can be truncated, thus approximating the pdf to within a reasonable accuracy.

E. Finite Dimensional Approximation

We are interested in analyzing the dynamic behavior of the parameters $a_i(k|k)$ and $b_i(k|k)$ as a function of time. Their decay would imply that the pdf sum could be truncated, leading to a finite dimensional representation and a finite dimensional estimator. Unfortunately, the division presented in Eqs. (29) complicates the analysis to a point that no definite conclusion can be drawn regarding the dynamics of those parameters.

However, using the same derivation and factorization presented above, an alternative estimator can be constructed by propagating the joint (and not the conditional) pdf of the state $x(k)$ and the measurement history $y(k)$. Specifically, similar to Eq. (17), prior to the measurement update the joint pdf is represented by

$$f_{X_k Y_{k-1}}\big(x(k)y(k{-}1)\big) = \sum_{i=1}^{k+1} \frac{\mathcal{a}_i(k|k{-}1)x(k) + \mathcal{b}_i(k|k{-}1)}{\big(x(k) - \sigma_i(k|k{-}1)\big)^2 + \omega_i^2(k|k{-}1)} \tag{40}$$

This pdf is initialized the same way as the one in Eq. (17). A measurement update leads to

$$f_{X_k Y_k}\big(x(k)y(k)\big) = \sum_{i=1}^{k+2} \frac{\mathcal{a}_i(k|k)x(k) + \mathcal{b}_i(k|k)}{\big(x(k) - \sigma_i(k|k)\big)^2 + \omega_i^2(k|k)} \tag{41}$$

obtained using the same computations performed to attain Eq. (20). Thus, $\sigma_i(k|k){=}\sigma_i(k|k{-}1)$, $\omega_i(k|k){=}\omega_i(k|k{-}1)$ for $i{=}1,\ldots,k{+}1$ represent the unchanged roots of the pdf denominator, while $\sigma_{k+2}(k|k){=}z(k)/H$ and $\omega_{k+2}(k|k){=}\gamma/|H|$ are the new roots. The numerator parameters are propagated and generated exactly as in the conditional pdf case, i.e., through Eqs. (21)-(26), while replacing (a_i, b_i) with $(\mathcal{a}_i, \mathcal{b}_i)$.

The time propagation of this joint pdf is identical to the one presented in section IV-C. In particular, Eqs. (36), (37) and (39) are unchanged, and in Eq. (38) (a_i, b_i) are replaced by $(\mathcal{a}_i, \mathcal{b}_i)$.

Propagating the joint pdf changes the computation of the state estimate and the second moment of the state to

$$\begin{aligned}\hat{x}(k|k) &= \left[\int_{-\infty}^{\infty} x(k) f_{X_k Y_k}\big(x(k)y(k)\big)\, dx(k)\right] / f_{Y_k}\big(y(k)\big) \\ &= \frac{\pi}{f_{Y_k}\big(y(k)\big)} \sum_{i=1}^{k+2} \frac{\mathcal{a}_i\left(\sigma_i^2 - \omega_i^2\right) + \mathcal{b}_i \sigma_i}{\omega_i}\end{aligned} \tag{42}$$

$$\begin{aligned}E\left(x^2|y(k)\right) &= \frac{\int_{-\infty}^{\infty} x^2(k) f_{X_k Y_k}\big(x(k)y(k)\big)\, dx(k)}{f_{Y_k}\big(y(k)\big)} \\ &= \frac{\pi}{f_{Y_k}\big(y(k)\big)} \sum_{i=1}^{k+2} \frac{\left(\mathcal{a}_i\sigma_i + \mathcal{b}_i\right)\left(\sigma_i^2 - \omega_i^2\right) - 2\mathcal{a}_i\sigma_i\omega_i^2}{\omega_i}\end{aligned} \tag{43}$$

where

$$\begin{aligned}f_{Y_k}\big(y(k)\big) &= \int_{-\infty}^{\infty} f_{X_k Y_k}\big(x(k)y(k)\big)\, dx(k) \\ &= \pi \sum_{i=1}^{k+2} \frac{\mathcal{a}_i(k|k)\sigma_i(k|k) + \mathcal{b}_i(k|k)}{\omega_i(k|k)}\end{aligned} \tag{44}$$

In Eqs. (42) and (43) the $(k|k)$ dependance of the pdf parameters is suppressed for presentation brevity.

Note that the time updated parameters $a_i(k|k)$ and $b_i(k|k)$ of the conditional pdf discussed in section IV-B can be computed from those of the joint pdf, $\mathcal{a}_i(k|k)$ and $\mathcal{b}_i(k|k)$, by dividing the latter by $f_{Y_k}\big(y(k)\big)$ of Eq. (44). Obviously, there is no change in the parameters $\sigma_i(k|k)$ and $\omega_i(k|k)$ between the conditional and joint pdfs. Consequently, both formulations lead to the same estimates and error variances.

The advantage of the derivation based on the joint pdf is in the fact that both the measurement update and time propagation equations for the pdf parameters are linear. Combining the measurement update and time propagation, the measurement updated parameters at steps k and $k{+}1$ are related linearly as

$$\sigma_i(k{+}1|k{+}1) = \Phi\sigma_i(k|k) + \Gamma u(k) \tag{45}$$

$$\omega_i(k{+}1|k{+}1) = |\Phi|\omega_i(k|k) + \beta \tag{46}$$

$$\begin{Bmatrix} \mathcal{a}_i(k{+}1|k{+}1) \\ \mathcal{b}_i(k{+}1|k{+}1) \end{Bmatrix} = F_i(k{+}1)G_i(k) \begin{Bmatrix} \mathcal{a}_i(k|k) \\ \mathcal{b}_i(k|k) \end{Bmatrix} \tag{47}$$

where matrices $F_i(k)$ and $G_i(k)$ are defined in Eqs. (22) and (39), respectively.

Eqs. (45) and (46) are simple, decoupled linear equations that can often be solved analytically for a given sequence $u(k)$. Eq. (47) is linear in the parameters $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$, but is coupled to $\sigma_i(k|k)$ and $\omega_i(k|k)$, that determine the matrix $F_i(k{+}1)G_i(k)$. Thus Eq. (47) is a linear, time-dependent, stochastic difference equation which we wish to analyze. In particular we wish to determine whether $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ vanish when $k{\to}\infty$.

The characteristic polynomial of $F_i(k)$ in Eq. (22) is

$$\left|\lambda I_2 - F_i(k)\right| = \lambda^2 - 2\frac{\delta_i(k)}{\Delta_i(k)}\lambda + \frac{\delta_i^2(k)+\theta_i^2(k)}{\Delta_i^2(k)} \quad (48)$$

where I_2 is a two-by-two identity matrix. The eigenvalues of $F_i(k)$ are

$$\lambda_{1,2} = \frac{\delta_i(k) \pm j\theta_i(k)}{\Delta_i(k)} \quad (49)$$

These eigenvalues are a complex conjugate pair when $\theta_i(k) \neq 0$. They are real and equal when $\theta_i(k){=}0$, i.e., when $z(k) = H\sigma_i(k|k)$, a probability zero event. In order to determine whether the eigenvalues of $F_i(k)$ are inside the unit circle, it is sufficient to test the size of its determinant

$$\begin{aligned} \det F_i(k) &= \frac{\delta_i^2(k)+\theta_i^2(k)}{\Delta_i^2(k)} \\ &= \frac{1}{\pi^2\gamma^2}\cdot\frac{1}{c^4+2c^2(d^2+1)+(d^2-1)^2} \end{aligned} \quad (50)$$

where

$$c = \frac{\sigma_i(k|k) - z(k)/H}{\gamma/|H|}, \quad d = \frac{\omega_i(k|k)}{\gamma/|H|} \quad (51)$$

The matrix $G_i(k)$ in Eq. (39) is lower triangular. Consequently, the eigenvalues of the matrix $F_i(k{+}1)G_i(k)$ in Eq. (47) cannot be easily determined, thus complicating the analysis of its dynamic characteristics. Nonetheless, it is possible to analyze its asymptotic behavior.

As explained at the beginning of this section, the element i in the pdf sum of Eq. (41) is introduced during the measurement update of time step $i{-}1$. Therefore, its parameters at step $k{>}i{-}1$ where propagated $m{=}k{-}i{+}1$ times through Eqs. (45)-(47). We will analyze those equation now when $k{\gg}i$ or equivalently $m{\gg}1$.

The analysis is different for stable ($|\Phi|{<}1$), marginally stable ($|\Phi|{=}1$), and unstable ($|\Phi|{>}1$) systems. It also depends on the sign of Φ. Thus, for simplicity, we will limit our analysis to the most common case of positive Φ. Moreover, we will assume a constant input $u(k){=}u_c$ for all $k{\geq}0$.

1) Case 1: $\Phi{<}1$:

Solving Eqs. (45) and (46) analytically, while letting $m{\to}\infty$, and substituting the results into Eq. (39) yields

$$\sigma_i(k|k) \to \frac{\Gamma u_c}{1-\Phi}, \quad \omega_i(k|k) \to \frac{\beta}{1-\Phi}, \quad G_i(k) \to I_2 \quad (52)$$

This implies that the dynamic characteristics of Eq. (47) is determined entirely by the eigenvalues of $F_i(k)$, or, as stated earlier, by its determinant, given in Eq. (50), with parameters

$$c \to \left(\frac{\Gamma u_c}{1-\Phi} - \frac{z(k)}{H}\right)\frac{|H|}{\gamma}, \quad d \to \frac{|H|\beta}{(1-\Phi)\gamma} \quad (53)$$

The parameters $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ will decay to zero iff the eigenvalues of $F_i(k{+}1)G_i(k)$ are inside the unit circle, or equivalently in this case, if $\det F_i(k){\ll}1$. The determinant, given in Eq. (50), is random due to the measurement noise in $z(k)$ and thus in c. However, $\det F_i(k)$ is largest when $c{=}0$, or when $z(k) = H\Gamma u_c/(1-\Phi)$. Although this is a zero probability event, a conservative condition for a decay to zero of $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ is that $\det F_i(k){\ll}1$ when $c{=}0$. Substituting Eq. (53) with $c{=}0$ into (50), this sufficient condition can be stated as

$$\left[\frac{\beta^2}{\gamma^2}\frac{H^2}{(1-\Phi)^2} - 1\right]^2 > \frac{1}{\pi^2\gamma^2} \quad (54)$$

It is striking that this deterministic, easily verifiable condition is sufficient to guarantee the stability of the stochastic, two-dimensional Eq. (47).

2) Case 2: $\Phi{=}1$:

In this case, the analytical solutions of Eq. (45) and (46) for $m{\to}\infty$, and the resulting $G_i(k)$ are

$$\sigma_i(k|k) \to m\Gamma u_c, \quad \omega_i(k|k) \to m\beta, \quad G_i(k) \to I_2 \quad (55)$$

Here too, $\det F_i(k){\ll}1$ will guarantee the decay of $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$. As in the previous case, $c{=}0$, or $z(k) = mH\Gamma u_c$ will lead to a largest $\det F_i(k)$, that is given by

$$\det F_i(k) = \frac{1}{\pi^2\gamma^2}\cdot\frac{1}{(d^2-1)^2} \to \frac{\gamma^2}{\pi^2 H^4\beta^4}\cdot\frac{1}{m^4} \to 0 \quad (56)$$

Hence a decay of $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ to zero is ensured in this case without posing additional convergence conditions.

3) Case 3: $\Phi{>}1$:

For unstable systems, the parameters $\sigma_i(k|k)$ and $\omega_i(k|k)$ diverge when $m{\to}\infty$ with dominant terms

$$\sigma_i(k|k) \to \left(\frac{z(k)}{H} - \frac{\Gamma u_c}{1-\Phi}\right)\Phi^m \quad (57a)$$

$$\omega_i(k) \to \left(\frac{\gamma}{|H|} - \frac{\beta}{1-\Phi}\right)\Phi^m \quad (57b)$$

i.e., they diverge as Φ^m. The matrix $G_i(k)$ converges to

$$G_i(k) \to \begin{bmatrix} 1 & 0 \\ \dfrac{\beta z(k) - \gamma\Gamma u_c \mathrm{sign}(H)}{\gamma \mathrm{sign}(H) + H\beta/(\Phi-1)} & \Phi \end{bmatrix} \quad (58)$$

and $\det F_i(k{+}1) \to 0$. Examining the elements of the matrix $F_i(k+1)$ reveals that the entire matrix tends to zero as $1/\Phi^m$ and faster. Hence we conclude that also in this case the parameters $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ are guarantied to decay to zero as $m{\to}\infty$.

The above analysis indicates that the weight of the elements in the factored pdf representation decay as m becomes large. Moreover, for $i \ll n$ and $n \ll k$, the parameters $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ will be much smaller than $\mathscr{a}_n(k|k)$ and $\mathscr{b}_n(k|k)$. This suggests that the latter could be neglected, truncating the elements of the pdf representation. However, before this truncation can be preformed, one has to examine the weighted sums where those parameters are used to compute the state estimate and the estimation error variance, Eqs. (42)-(44).

In Eq. (44), the elements in sum are

$$\begin{bmatrix} \frac{\mathscr{a}_i(k|k)\sigma_i(k|k)}{\omega_i(k|k)} & \frac{\mathscr{b}_i(k|k)}{\omega_i(k|k)} \end{bmatrix} \tag{59}$$

Using Eqs. (45)-(47) it is possible to derive a linear difference equation for the vector of Eq. (59). Following the convergence analysis steps for $\mathscr{a}_i(k|k)$ and $\mathscr{b}_i(k|k)$ presented above, identical convergence results are obtained also for the elements of the sum in Eq. (44).

Repeating the same procedure for the element of the sum in Eq. (42)

$$\begin{bmatrix} \frac{\mathscr{a}_i(k|k)\left(\sigma_i^2(k|k) - \omega_i^2(k|k)\right)}{\omega_i(k|k)} & \frac{\mathscr{b}_i(k|k)\sigma_i(k|k)}{\omega_i(k|k)} \end{bmatrix} \tag{60}$$

and then for those in Eq. (43)

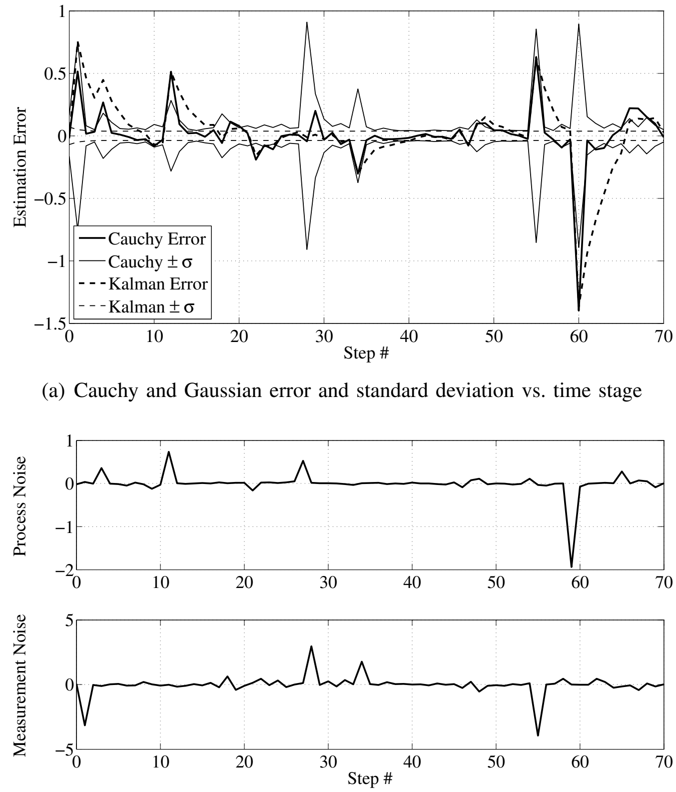
$$\begin{bmatrix} \frac{\mathscr{a}_i(k|k)\sigma_i(k|k)\left(\sigma_i^2(k|k) - 3\omega_i^2(k|k)\right)}{\omega_i(k|k)} \\ \frac{\mathscr{b}_i(k|k)\left(\sigma_i^2(k|k) - \omega_i^2(k|k)\right)}{\omega_i(k|k)} \end{bmatrix} \tag{61}$$

reveals also a similar decay of those sum elements.

The above analysis implies that the pdf elements i with $i \ll n$ and $n \ll k$ can be neglected, leading to an approximate pdf with a *finite* number of elements. Due to the linear relation between the numerator parameters of the conditional and joint pdfs, the suggested truncation is valid for both approaches. An estimator with a finite number of elements, i.e., a finite dimensional filter is desirable for any practical implementation. Based on the above conclusion it is tantalizing to suggest that there could be a finite dimensional structure for the Cauchy estimator. This structure is yet to be revealed!

V. Numerical Example

The proposed recursive Cauchy estimation scheme is evaluated using a numerical simulation. The parameters used in this simulation are as follows. For the system model we chose $\Phi{=}0.9$, $\Gamma u(k){=}1$, $k{=}0, 1, \ldots, 70$, i.e., a constant input, and $H{=}2$. The initial condition and noise parameters are: $\bar{x}{=}5$, $\alpha{=}0.5$, $\beta{=}0.02$, and $\gamma{=}0.1$. We also compare the performance of the Cauchy and Gaussian (Kalman) filters when the additive noises are Cauchy and then when they are Gaussian. For the Cauchy noise case, the Kalman filter parameters were chosen to least square fit the respective Cauchy pdf, which leads to setting the initial condition, process and measurement noise standard deviations to $\sigma_{x_0}{=}1.4\alpha$, $\sigma_w{=}1.4\beta$, and $\sigma_v{=}1.4\gamma$, respectively. For the Gaussian noise simulations, the latter values were used to also generate the simulation data.



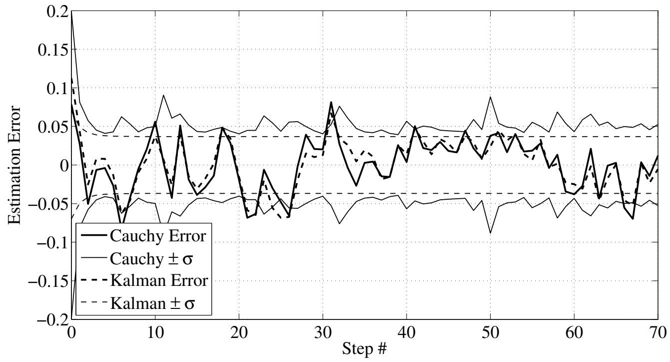
(a) Cauchy and Gaussian error and standard deviation vs. time stage

(b) Cauchy process and measurement noises

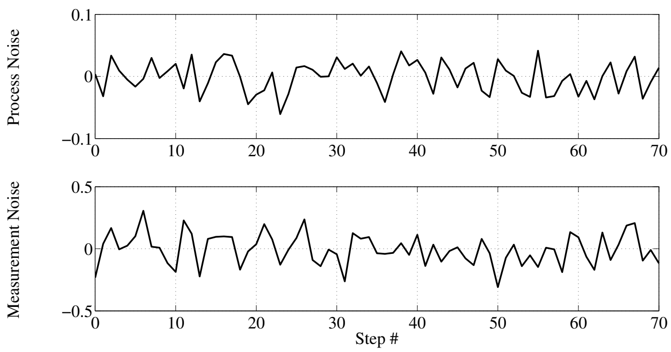
Fig. 2. Performance of the Cauchy and Kalman filters with Cauchy noises.

The simulation results for 70 steps are presented and analyzed next. The estimation errors of the Cauchy and Kalman estimators are plotted in Fig. 2(a) together with the computed estimation error standard deviations for Cauchy process and measurement noise sequences of Figs. 2(b). Here we can clearly see how the impulsive nature of the data, seen in the noise sequences of Figs. 2(b), increase the estimation inaccuracy. However, contrary to the Kalman filter, the Cauchy standard deviation of the error also increases, indicating the abnormal state deviations (due to drastic process noise input) or measurement deviations. The conditional standard deviation of the error for the Cauchy filter, plotted in Fig. 2(a), is minimal. However, the standard deviation of the error plotted for the Kalman filter is calculated assuming Gaussian noise variances and thus is neither related to the actual estimation errors nor it is minimal.

Figure 3(a) presents the estimation errors together with the computed estimation error standard deviations for the Cauchy and Kalman estimator when the process and measurement noise sequences shown in Figs. 3(b) are Gaussian. The error variance for the Kalman filter is now the minimal, since the process and measurement noise sequence are Gaussian. It is remarkable how close the standard deviation of the estimation error generated from the Cauchy conditional pdf approximates



(a) Cauchy and Gaussian error and standard deviation vs. time stage



(b) Gaussian process and measurement noises

Fig. 3. Performance of the Cauchy and Kalman filters with Gaussian noises.

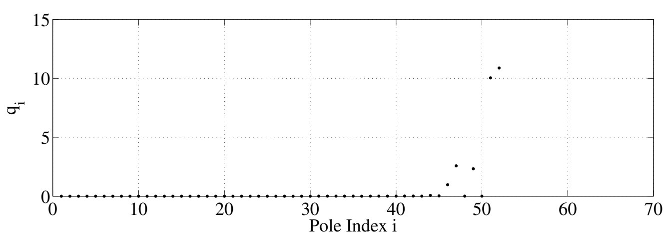


Fig. 4. Weights of the terms in the pdf sum at time step 50.

that computed from the Kalman filter, demonstrating the robustness of the former.

Next we examine the characteristics of the growing number of terms in the conditional pdfs of Eq. (28) for the Cauchy noise case. Identical results would have been obtained by analyzing the joint pdf formulation of Eq. (41).

To evaluate the relative importance of the different elements in the pdf sums of Eq. (28), analyzed in section IV-E, we define the weight of the $i\text{-}th$ term as $q_i = \sqrt{a_i^2(k|k) + b_i^2(k|k)}$. Those weights are plotted for a sample time step number 50 in Fig. 4. It is obvious that out of the 52 terms only the last few have a significant weight, while the effect of the others is negligible. Here, the last twelve are significant while the next term has a weight four orders of magnitude smaller than the biggest term. If eighteen largest terms were considered, the next term is 12 orders of magnitude smaller.

This result is in agreement with the analysis carried out in section IV-E, since for this stable system with the chosen parameters the convergence conditions in Eq. (54) are satisfied. This suggests that the pdf sum can be trimmed, with a negligible affect on the estimation accuracy. This result was verified numerically: the estimation results with a trimmed pdf were undistinguishable from those presented in Fig. 2(a).

VI. Conclusion

A new estimator for scalar linear discrete system with additive process and measurement noises characterized by Cauchy probability density functions (pdf) has been presented. This Cauchy estimator is characterized by a functional form of the conditional pdf that remains fixed through time propagation and measurement updates. Although this rational pdf has growing order, its parameters are propagated through linear difference equations with stochastic coefficients. The functional form of the pdf enables an efficient computation of the state estimate and its error variance, as is demonstrated in a numerical example. Sufficient conditions and the numerical study also reveal that the majority of terms in the growing-in-dimension conditional pdf are negligible, suggesting efficient, finite-order numerical approximation.

The proposed Cauchy estimator is probably the only known simple recursive scheme for non-Gaussian pdfs. Moreover, the proposed methodology and pdf factorization presented in this work can be easily adopted to a wide range of pdfs that can be expressed as a class of rational functions of the random variable, Cauchy pdf being a member of this class. Although only the scalar dynamic estimation case is considered in the current work, the extension to the multi-variable case appears to have the same mathematical structure. Therefore, the techniques developed for the scalar problem are preserved in the more general multivariate setting, accompanied though with a growth in the algorithmic complexity.

Acknowledgment

The authors thank Prof. Gil Iosilevskii from the Faculty of Aerospace Engineering, Technion, Haifa, Israel, for his advise on some of the mathematical derivations in this work.

References

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

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

[3] P. Tsakalides and C. L. Nikias, *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, R. J. Adler, R. E. Feldman, and M. S. Taqqu, Eds. Birkhauser, 1998.

[4] G. A. Tsihrintzis, *Statistical Modeling and Receiver Design for Multi-User Communication Networks;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, R. J. Adler, R. E. Feldman, and M. S. Taqqu, Eds. Birkhauser, 1998.

[5] L. N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. John Wiley & Sons, Inc., 1994, vol. 1, ch. 16.

[6] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[7] G. F. Carrier, M. Krook, and C. E. Pearson, *Functions of a Complex Variable: Theory and Technique*. Academic Press, 1966.