AN EFFICIENT FILTER FOR MEASUREMENTS CORRUPTED WITH CAUCHY NOISE

James S. McCabe* **and Matthew J. Gualdoni***

This paper present a new sequential filter for state estimation using measurements corrupted with Cauchy noise. The new filter retains the familiar structure of the Kalman filter and is computationally efficient. In addition, it does not exhibit computational complexity which grows or varies in time like existing methods. These results are based upon a nearly 50 year old result by Masreliez in which the conditional mean estimator is approximated via linearization of the measurement predictive density. This work derives the new filter, provides discussion regarding practical implementation, and present Monte Carlo analyses to validate and assess the new filter's performance.

BACKGROUND & MOTIVATION

Let the n_x-dimensional state of interest at time t_k be denoted $\boldsymbol{x}_k$ and let it evolve according to

$$\boldsymbol{x}_k = \boldsymbol{F}_{x,k}\boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1}\,,$$

where $\boldsymbol{w}_{k-1}$ is a zero-mean, white, and Gaussian noise with known covariance $\boldsymbol{P}_{ww,k-1}$. Let scalar observations of the state be made according to

$$z_k = \boldsymbol{H}_{x,k}\boldsymbol{x}_k + v_k\,,$$

where $\boldsymbol{H}_{x,k}$ is a $1 \times n_x$ row vector and v_k is a Cauchy-distributed noise process corrupting the observations. Recall that the Cauchy distribution with location (or median or mode) parameter $m_v \in \mathbb{R}$ and scale parameter $\gamma > 0$ is given by

$$v_k \sim \mathcal{C}(v_k\,;\, m_v, \gamma) = \frac{1}{\pi}\left[\frac{\gamma}{(v_k - m_v)^2 + \gamma^2}\right]\,,$$

but note that the entirety of this paper selects $m_v = 0$. Additionally, let the collection of all measurements up to and including $\boldsymbol{z}_k$ be denoted as the set

$$\boldsymbol{Z}_{1:k} = \{z_1, \ldots, z_k\}\,.$$

A brief discussion regarding the Cauchy distribution is necessary. Firstly, it deserves note that the Cauchy distribution commonly serves as a standard example of a pathological distribution in that its first and second (central) moments are undefined, starkly contrasting it with the ubiquitous Gaussian distribution that is uniquely defined on those very moments. Additionally, linear combinations

*GN&C Autonomous Flight Systems Engineer, EG6, Aeroscience and Flight Mechanics Division, NASA Johnson Space Center, 2101 NASA Parkway

of Cauchy distributed random variables do not converge in the manner suggested by the standard Central Limit Theorem (to a Gaussian density) and, instead, converge to a result that is, again, Cauchy. Finally, the Cauchy density, due to its extreme heavy tails, is often referred to in practical filtering applications since real-world applications often exhibit heavy-tailed behavior. However, the Gaussian distribution dominates the published domain due to the analytical convenience of its related methods, and filters that directly utilize measurements corrupted by Cauchy noise are few and far between. The primary goal of this paper is to enable a result that is just as practical and easy to implement but operates directly with the true underlying Cauchy noise density.

The Cauchy Measurement Predictive Density

The methods employed later in this paper require explicit computation of the measurement predictive density $p(\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})$. As such, we will define the probabilistic structure of the prior estimate and the measurement noise. Then, the convolution of these two densities results in the likelihood. Accordingly, take the *a priori* distribution of $\boldsymbol{x}_k$, denoted as $p(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1})$, to be Gaussian with mean $\boldsymbol{m}_{x,k}^-$ and covariance $\boldsymbol{P}_{xx,k}^-$, i.e.,

$$p(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1}) = p_g(\boldsymbol{x}_k\,;\,\boldsymbol{m}_{x,k}^-, \boldsymbol{P}_{xx,k}^-)\,.$$

Then, the distribution of mapped estimate, $\boldsymbol{H}_{x,k}\boldsymbol{x}_k$, (i.e., the estimated measurement without the presence of noise) is also Gaussian of the form

$$p(\boldsymbol{H}_{x,k}\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1}) = p_g(\boldsymbol{H}_{x,k}\boldsymbol{x}_k\,;\,\boldsymbol{H}_{x,k}\boldsymbol{m}_{x,k}^-, W_k) \tag{1}$$

where (recalling that z_k is a scalar and thus $\boldsymbol{H}_{x,k}\boldsymbol{P}_{xx,k}^-\boldsymbol{H}_{x,k}^T$ is a scalar)

$$W_k \triangleq \boldsymbol{H}_{x,k}\boldsymbol{P}_{xx,k}^-\boldsymbol{H}_{x,k}^T\,. \tag{2}$$

The Cauchy noise distribution is given as

$$p(v_k|\boldsymbol{Z}_{1:k-1}) = p_v(v_k) = \mathcal{C}(v_k\,;\,m_v, \gamma)\,. \tag{3}$$

The convolution of Eqs. (1) and (3) results in the predictive density

$$p(z_k|\boldsymbol{Z}_{1:k-1}) = \int p_g(\boldsymbol{H}_{x,k}\boldsymbol{x}_k\,;\,\boldsymbol{H}_{x,k}\boldsymbol{m}_{x,k}^-, W_k)\mathcal{C}(v_k - \xi\,;\,m_v, \gamma)\mathrm{d}\xi\,.$$

The solution to this function, known as "the Voigt profile", named for German physicist Woldemar Voigt, has been a focus of study within physics, engineering, and mathematics for more than a century, and, as its origins were in spectroscopy, it seems mere coincidence that it appears in the present application. The Voigt profile has yet to be shown to have an analytic solution, though several closed-form solution in terms of non-analytic functions exist. Seemingly, the earliest published closed-form solution was presented by Kendall in 1938,[2] though the most common reference appears to be that of Fadeeva in 1961.[3*] Abramowitz and Stegun published an identical result to Fadeeva's in 1964 with Eq. (7.1.3) in Ref. 4 without attribution to Fadeeva, but it seems unlikely they would have been aware of Fadeeva's work.

[*]It has been claimed that the original, untranslated version of these results were published in Russia in 1954, but the authors have been unable to verify this.

These solutions are closed-form but non-analytic in that they contain the error function in their representation. Many modern looks into the Voight profile have been in search of analytic approximations, such as those by Liu[5] and Rocco,[6] though their own works document their limitations. Other modern approaches to this have been, instead, in further generalizing the historical results, with Pagnini producing, perhaps, the most general form known to-date with their Fox H-Function representation in Ref. 7. This work proceeds with a related work, by Mainardi, that shows the Voigt profile to be given as[8]

$$p(z_k|\boldsymbol{Z}_{1:k-1}) = \frac{1}{4\sqrt{\pi a}}\left[\varphi(z_k) + \bar{\varphi}(z_k)\right] , \tag{4}$$

where

$$a \triangleq \frac{1}{2}W_k \tag{5}$$

$$A(z_k) \triangleq (\gamma + i[z_k - \boldsymbol{H}_{x,k}\boldsymbol{m}^-_{x,k}])/\sqrt{a} \tag{6}$$

$$\varphi(z_k) \triangleq \exp\left\{\frac{1}{4}A^2(z_k)\right\}\operatorname{erfc}\left\{\frac{1}{2}A(z_k)\right\} . \tag{7}$$

Note that the error function is result is present and, thus, this predictive density is non-analytic. However, this form proves sufficient for the present application in terms of differentiability and structure. Throughout this paper, the following abbreviations are employed

$$A(z_k) \overset{\text{abbr.}}{=} A$$
$$\varphi(z_k) \overset{\text{abbr.}}{=} \varphi ,$$

but it is important to remember that both A and φ depend on the measurement, z_k.

The Masreliez Filter

In Ref. 9, Masreliez presented a novel strategy for approximate optimal estimation in the presence of non-Gaussian measurement noises, and the resulting filter, herein referred to as the Masreliez filter, was demonstrated to produce optimal estimates (matching those generated by the Gaussian mixture filter[10]). They presented two filters, referred to as "the first kind" and "the second kind." The second kind considers cases where dynamical process noises are non-Gaussian and the measurement noises are Gaussian. This work proceeds with the filter of the first kind, wherein dynamical process noises are taken to be Gaussian and the measurement noises are non-Gaussian. The strategies presented here can surely be adapted to the filter of the second kind but are disregarded here.

Assuming the *a priori* state distribution is approximately Gaussian, the Masreliez filter's mean and covariance update are given as

$$\boldsymbol{m}^+_{x,k} = \boldsymbol{m}^-_{x,k} + \boldsymbol{P}^-_{xx,k}\boldsymbol{H}^T_{x,k}\boldsymbol{g}_k(\boldsymbol{z}_k) \tag{8}$$

$$\boldsymbol{P}^+_{xx,k} = \boldsymbol{P}^-_{xx,k} - \boldsymbol{P}^-_{xx,k}\boldsymbol{H}^T_{x,k}\boldsymbol{G}_k(\boldsymbol{z}_k)\boldsymbol{H}_{x,k}\boldsymbol{P}^-_{xx,k} \tag{9}$$

where the mean update quantity is given by the vector

$$\boldsymbol{g}_k(\boldsymbol{z}_k) = \begin{bmatrix} g_{1,k}(\boldsymbol{z}_k) & \cdots & g_{n_z,k}(\boldsymbol{z}_k) \end{bmatrix}^T$$

with elements defined as

$$g_{i,k}(\boldsymbol{z}_k) = -\left[\frac{\partial p(\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})}{\partial z_{i,k}}\right]\frac{1}{p(\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})} \tag{10}$$

and the matrix $\boldsymbol{G}_k(\boldsymbol{z}_k)$ has entries

$$[\boldsymbol{G}_k(\boldsymbol{z}_k)]_{i,j} = \frac{\partial g_{i,k}(\boldsymbol{z}_k)}{\partial z_{j,k}}\,, \tag{11}$$

where $[\boldsymbol{A}]_{i,j}$ denotes the $(i,j)^{\text{th}}$ entry of $\boldsymbol{A}$.

The method relies upon computing the derivatives in Eqs. (10) and (11), a procedure which, in general, can require intensive effort or may even be impossible depending on the nature of the noise $\boldsymbol{v}_k$ and resulting measurement predictive density $p(\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})$. This, of course, even assumes the convolution necessary to produce $p(\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})$ can be computed in closed form! In fact, Masreliez themselves calculated the necessary convolution and subsequent derivatives numerically in their demonstrations, deeming the necessary partials too "hard to implement".[9]

Main Result: A New Filter For Cauchy Noise

A new filter for estimation of systems subject to Cauchy measurement noises is now presented. Note that, since z_k is a scalar, both $g_k(z_k)$ and $G_k(z_k)$ are scalars as well. As such, the mean and covariance updates in Eqs. (8) and (9) become

$$\boldsymbol{m}^+_{x,k} = \boldsymbol{m}^-_{x,k} + g_k(z_k)\boldsymbol{P}^-_{xx,k}\boldsymbol{H}^T_{x,k} \tag{12}$$

$$\boldsymbol{P}^+_{xx,k} = \boldsymbol{P}^-_{xx,k} - G_k(z_k)\boldsymbol{P}^-_{xx,k}\boldsymbol{H}^T_{x,k}\boldsymbol{H}_{x,k}\boldsymbol{P}^-_{xx,k} \tag{13}$$

where

$$g_k(z_k) = \frac{1}{2\sqrt{a}}\frac{\text{Im}[A\varphi]}{\text{Re}[\varphi]} \tag{14}$$

$$G_k(z_k) = g_k^2(z_k) + \frac{\text{Re}\left[(\frac{1}{2}A^2+1)\varphi - A/\sqrt{\pi}\right]}{2a\text{Re}[\varphi]}\,. \tag{15}$$

Equations (12)–(15), paired with Eqs. (5)–(7), fully define the new filter and provide (approximately) optimal state estimates for systems corrupted with Cauchy measurement noise. Here, "approximately optimal" applies in the sense of Masreliez as discussed in Ref. 9, and the numerical studies later in this paper support this claim. For subsequent time updates, all standard mean and covariance propagation equations can be applied.

WORK TO PRESENT IN FINAL PAPER

Given these theoretical findings (and additional explicit proof and appropriate theoretical justification for these results), more discussion is to be given regarding the implementation considerations for this new filter. Then, numerical studies, namely Monte Carlo analyses, are to be presented to illustrate its performance in comparison to other published techniques, including the Kalman filter and the Idan-Speyer Cauchy Estimator (ISCE). Conclusions will be presented, considering performance, implementation feasibility, and applicability to real-world problems.

REFERENCES

[1] A. Gelb, *Applied Optimal Estimation*. M.I.T. Press, 1974. Ch. 4.

[2] D. G. Kendall, "The effect of radiation damping and Doppler broadening on the atomic absorption coefficient.," *Zeitschrift fr Astrophysik*, Vol. 16, Jan. 1938, p. 308.

[3] V. N. Fadeeva and T. N. M. Stegun, *Tables of values of the function* $w(z) = exp(-z^2)(1 + 2i/\sqrt{\pi}\int_0^z exp(t^2)dt)$ *for complex argument*. New York: Pergamon Press, 1961.

[4] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, ninth printing ed., 1964.

[5] Y. Liu, J. Lin, G. Huang, Y. Guo, and C. Duan, "Simple empirical analytical approximation to the Voigt profile," *J. Opt. Soc. Am. B*, Vol. 18, May 2001, pp. 666–672, doi:10.1364/JOSAB.18.000666.

[6] H. O. D. Rocco and A. Cruzado, "The Voigt Profile as a Sum of a Gaussian and a Lorentzian Functions, when the Weight Coefficient Depends Only on the Widths Ratio," *Acta Physica Polonica A*, Vol. 122, 2012, pp. 666–669.

[7] G. Pagnini and R. K. Saxena, "A note on the Voigt profile function," 2008, doi:10.48550/arXiv.0805.2274.

[8] F. Mainardi, G. Pagnini, *et al.*, "Mellin-Barnes integrals for stable distributions and their convolutions," *Fractional Calculus & Applied Analysis*, Vol. 11, No. 4, 2008, pp. 443–456.

[9] C. Masreliez, "Approximate non-Gaussian filtering with Linear State and Observation Relations," *IEEE Transactions on Automatic Control*, Vol. 20, No. 1, 1975, pp. 107–110, doi:10.1109/TAC.1975.1100882.

[10] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian Estimation using Gaussian Sum Approximations," *IEEE Transactions on Automatic Control*, Vol. AC-17, August 1972, pp. 439–448, doi:10.1109/TAC.1972.1100034.