A Stochastic Controller for Vector Linear Systems with Additive Cauchy Noises

Javier Fernández
Mechanical and Aerospace Engineering
University of California, Los Angeles
Los Angeles, California 90095–1597
Email: jhf@seas.ucla.edu

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

Moshe Idan
Aerospace Engineering
Technion - Israel Institute of Technology
Haifa, Israel, 32000
Email: moshe.idan@technion.ac.il

Abstract—An optimal predictive controller for linear, vector-state dynamic systems driven by Cauchy measurement and process noises is developed. For the vector-state system, the probability distribution function (pdf) of the state conditioned on the measurement history cannot be generated. However, the characteristic function of this pdf can be expressed in an analytic form. Consequently, the performance index is evaluated in the spectral domain using this characteristic function. By using an objective function that is a product of functions resembling Cauchy pdfs, the conditional performance index is obtained analytically in closed form by using Parseval’s equation and integrating over the spectral vector. This forms a non-convex function of the control signal, and must be optimized numerically at each time step. A two-state example is used to expose the interesting robustness characteristics of the proposed controller.

I. INTRODUCTION

Models in modern stochastic optimal control algorithms like the linear quadratic Gaussian (LQG) and the linear exponential Gaussian (LEG) assume linear dynamics and additive process and measurement noises described by the Gaussian probability density function (pdf). The Gaussian distribution function has very light tails, so that large deviations are essentially impossible. Therefore, the LQG and LEG algorithms do not perform well in the presence of heavy-tailed or impulsive uncertainties. In many practical applications, such as radar and sonar systems affected by atmospheric and underwater acoustic noises, more impulsive uncertainties are observed [1]. Impulsive behavior is also more effective at modeling adversarial motion, as is air turbulence, which is better described by distributions with heavier tails than the Gaussian [2].

Therefore, in this paper we propose a system model that assumes linear dynamics driven by additive process and measurement noises described by Cauchy pdfs. Both the Cauchy and Gaussian pdfs belong to a class of distributions called the symmetric $\alpha$-stable (So-S) class, whose members are described by their characteristic functions. A full treatment of the So-S class can be found in [3]. The Cauchy pdf is in a subset of this class whose members have infinite second moments. In addition, the mean of the Cauchy pdf is not well defined.

Algorithms for optimal estimation and control of scalar linear systems driven by Cauchy distributed process and measurement noises have been developed previously in [4, 5]. There, the conditional performance index for model predictive control is determined directly by taking the conditional expectation of the objective function using the probability density given the measurement history as presented in [4]. A dynamic programming algorithm is also developed in [5]. It is shown that the solution to the dynamic programming recursion is intractable because of the need to average over future measurements in determining the optimal return function.

In this paper, the Cauchy optimal control algorithm for scalar systems [5] is extended to systems with a vector state. For the vector case, the conditional pdf (cpdf) given the measurement history is not available. However the characteristic function of the cpdf can be recursively propagated [6, 7]. The significant contribution in this paper is evaluating, in closed form, the conditional performance index using the cpdf’s characteristic function instead of the cpdf itself, and integrating over the spectral variables instead of the state variables.

Although the cpdf is not available as a function of the state vector, the conditional expectation of the objective function, i.e. the conditional performance index, can be computed using the characteristic function of the cpdf, which is available as a function of the spectral vector [7]. The objective function is cast as a product of functions resembling Cauchy pdfs, which are easily transformed into a function of the spectral variables. Consequently, the conditional performance index is found in a closed form. Due to its complexity, the optimal control signal is determined by numerically optimizing this conditional performance index in a model predictive control setting.

The remainder of the paper is structured as follows. The controlled system model is presented in Section II. An appropriate, computable performance index for this problem is presented in Section III and subsequently transformed from the state variable to the spectral variable form. In Section IV the spectral integrations required to determine the conditional performance index are reduced to an integral formula that can be evaluated in closed form. Section V addresses a special case of systems with two states. Here, using an alternative, simplified form of the two state cpdf’s characteristic function, the conditional performance is determined in closed form. In Section VI numerical examples are given. Conclusions are given in Section VII.

This work was partially supported by Air Force Office of Scientific Research, Award No. FA9550-10-1-0570, and by the United States - Israel Binational Science Foundation, Grant 2008040.
II. DESCRIPTION OF THE MODEL

This paper deals with a discrete time, linear system described by
\[
    x(k+1) = \Phi x(k) + \Gamma w(k)
\]
where \( x(k) \in \mathbb{R}^n \) is the state vector, \( u(k) \) is a scalar deterministic input, \( z(k) \) is a scalar measurement, and \( w(k) \) and \( v(k) \) are scalar independent Cauchy distributed process and measurement noise inputs with medians at zero and scaling parameters of \( \beta \) and \( \gamma \), respectively, so that their pdfs are given by
\[
    f_{W_k}(w(k)) = \frac{\beta/\pi}{w^2(k) + \beta^2}, \quad f_{V_k}(v(k)) = \frac{\gamma/\pi}{v^2(k) + \gamma^2}. \quad (2)
\]
The characteristic functions of these pdfs are
\[
    \phi_{W_k}(\sigma) = e^{-\beta|\sigma|}, \quad \phi_{V_k}(\sigma) = e^{-\gamma|\sigma|}, \quad (3)
\]
where \( \sigma \) is the scalar spectral variable.

The initial conditions are assumed to be independent Cauchy distributed random variables with the pdfs
\[
    f_{X_0}(x(0)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{(x_i(0) - \bar{x}_i(0))^2 + \alpha_i^2}. \quad (4a)
\]
Its characteristic function is given by
\[
    \phi_{X_0}(\nu) = \prod_{i=1}^{n} e^{-\alpha_i|\nu_i| + j\bar{x}_i(0)\nu_i}. \quad (4b)
\]
where \( \nu_i \) is an element of \( \nu \in \mathbb{R}^n \).

The stochastic system (1) can be decomposed into two systems, one driven by \( u(k) \) and one by \( w(k) \), by exploiting the linearity of the system. Let \( \bar{x}(k) \) and \( \tilde{z}(k) \) be the part of the system driven by the control \( u(k) \) only, and \( \bar{x}(k) \) and \( \tilde{z}(k) \) be the part of the system driven by the process noise \( w(k) \) only and contains all the underlying random variables. Then,
\[
    x(k) = \bar{x}(k) + \tilde{z}(k) \quad (5a)
\]
\[
    z(k) = \bar{z}(k) \quad (5b)
\]
The controlled part of the system is described by
\[
    \bar{x}(k) = \Phi \bar{x}(k-1) + \Lambda u(k-1)
\]
\[
    \bar{z}(k) = \bar{H}\bar{z}(k) \quad (6a)
\]
with initial condition \( \tilde{z}(0) \). The process noise driven part is given by
\[
    \tilde{x}(k) = \Phi \tilde{x}(k-1) + \Gamma w(k-1)
\]
\[
    \tilde{z}(k) = \tilde{H}\tilde{z}(k) + v(k). \quad (7a)
\]
The process and measurements noise pdfs were defined in (2), while the initial condition of this stochastic model is Cauchy distributed with a pdf given by
\[
    f_{\tilde{X}_0}(\tilde{x}(0)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{\tilde{x}_i^2(0) + \alpha_i^2}. \quad (8a)
\]
Its characteristic function is
\[
    \phi_{\tilde{X}_0}(\nu) = \prod_{i=1}^{n} e^{-\alpha_i|\nu_i|}. \quad (8b)
\]
The above decomposition will be used to derive the Cauchy controller.

Let the state, measurement, and control histories used in the control problem formulation be defined as
\[
    \mathcal{X}_k^m := \{x(\ell), \ldots, x(m)\}, \quad (9a)
\]
\[
    \tilde{Z}_k := \{\tilde{z}(0), \ldots, \tilde{z}(k)\}, \quad (9b)
\]
\[
    \mathcal{U}_k^m := \{u(\ell), \ldots, u(m)\}, \quad \mathcal{U}_k^m \in \mathcal{F} \quad (9c)
\]
where \( \mathcal{F} \) is the class of piecewise continuous functions adapted to the \( \sigma \)-algebra \( \sigma_k \) generated by the measurement history, i.e. the control is a random variable that is measurable with respect to events in \( \sigma_k \) [8]. Moreover, in [9] it is shown that \( u_k \) is adapted to the \( \sigma \)-algebra \( \sigma_k \) generated by \( \tilde{Z}_k \), which means that the control is measurable on events generated by \( \tilde{Z}_k \) only.

III. DERIVATION OF THE COST USING CHARACTERISTIC FUNCTIONS

Our proposed controller is an \( m \)-step model predictive controller [10] that uses current and past measurements, and averages over future process noise. At each time step, the conditional performance index is computed. Since the performance index will be shown to be a nonconvex function of the control sequence, it is maximized numerically. Once the optimal control sequence of length \( m \) is computed, only the first control in that sequence is applied. At the next step, a new measurement is taken and the process is repeated, producing a new optimal control sequence and applying only the first one.

In this paper, we study the optimal stochastic state regulation problem, noting that the tracking problem can be handled in a similar fashion. Our regulation problem will have a finite horizon of length \( m \) such that the terminal state occurs at time-step \( p = k + m \).

Similar to the scalar control problem presented in [5], the control objective function is chosen as a product of Cauchy-like functions given by
\[
    \psi \left( \mathcal{X}_k^p, \mathcal{U}_k^{p-1} \right) = \prod_{i=k}^{p-1} \left( \frac{\zeta_i/\pi}{u^2(i) + \zeta_i^2} \right) \prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{\alpha_{i+1,r}^2(i + 1) + \eta_{i+1,r}^2} \quad (10)
\]
Then, the the performance index conditioned on the current measurement history and averaged over future process noises is given by
\[
    J_{k,p} = \max_{\mathcal{U}_k^{p-1} \in \mathcal{F}} E \left[ \psi \left( \mathcal{X}_k^p, \mathcal{U}_k^{p-1} \right) \right]
\]
\[
= \max_{\mathcal{U}_k^{p-1} \in \mathcal{F}} E \left[ \psi \left( \mathcal{X}_k^p, \mathcal{U}_k^{p-1} \right) \right] \tilde{Z}_k
\]
\[
= E \left[ \max_{\mathcal{U}_k^{p-1} \in \mathcal{F}} E \left[ \psi \left( \mathcal{X}_k^p, \mathcal{U}_k^{p-1} \right) \tilde{Z}_k \right] \right] \triangleq E \left[ J_{\tilde{Z}_k} \right], \quad (11)
\]
where the interchange of the maximum and expectation operations is due to the fundamental theorem in [11].

We are now concerned with determining the analytic form for the conditional performance index \( J_{\tilde{Z}_k} \). Using (10), it
For now, let us only consider weighting on the terminal state, \( x(p) \), and on the \( m \) scalar control inputs. The control weighting functions \( \mathcal{M}_{it} \) depend on \( \int \eta_{i+1,r}/\pi \) can come out of the integral. Then, the product over \( i \) inside the integrand in (12) has only the term for \( i = p \) and, the integral is only over \( \{\tilde{x}_1(p), \ldots, \tilde{x}_n(p)\} \). Thus, for notational convenience we can drop the time-step index in this integral and write it over \( \{\tilde{x}_1, \ldots, \tilde{x}_n\} \) as

\[
J_{\tilde{z}_k} = \mathcal{M}_{it} \int_{-\infty}^{\infty} \left( \prod_{r=1}^{p-1} \frac{\eta_{i,r}/\pi}{\pi} \right) \times f_{\tilde{X}_k|\tilde{z}_k}(\tilde{x}(k+1)|\tilde{z}(k)) d\tilde{x}_1(k+1) \ldots d\tilde{x}_n(k+1) \times d\tilde{x}_1(k+2) \ldots d\tilde{x}_n(k+2) \ldots d\tilde{x}_1(p) \ldots d\tilde{x}_n(p) \quad (12)
\]

The pdf \( f_{\tilde{X}_k|\tilde{z}_k}(\tilde{x}|\tilde{Z}_k) \) can be evaluated in closed form for scalar systems [4]. However, for vector state systems it is the characteristic function of the pdf, \( \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) \), that is evaluated in closed form. Therefore, when computing the conditional performance index, we need to be able to integrate over the spectral variable \( \nu \) instead of the pdf variable \( \tilde{x} \).

Define the product over \( r \) in the integral in (13) as \( \ell_x \) and its Fourier transform as \( L_x \).

\[
\ell_x(\tilde{x} + \tilde{x}) = \prod_{r=1}^{n} \left( \frac{\eta_{i+1,r}/\pi}{\pi} \right) \times f_{\tilde{X}_k|\tilde{z}_k}(\tilde{x}(k+1)|\tilde{z}(k)) d\tilde{x}_1 \ldots d\tilde{x}_n \quad (13a)
\]

\[
L_x(\nu) = \prod_{r=1}^{n} e^{-\eta_{i,r}|\nu_r|} = \prod_{r=1}^{n} e^{-j\eta_{i,r}|\nu_r|} \quad (13b)
\]

Using these definitions, we can apply Parseval’s equation over each variable in (13) to express the conditional performance index as an integral over the spectral variable \( \nu \).

\[
J_{\tilde{z}_k} = \frac{\mathcal{M}_{it}}{(2\pi)^n} \int_{-\infty}^{\infty} L_x^* (\nu) |\phi_{\tilde{X}_k|\tilde{z}_k}(\nu) d\nu_1 \ldots d\nu_n
\]

\[
= \frac{\mathcal{M}_{it}}{(2\pi)^n} \int_{-\infty}^{\infty} \left( \prod_{r=1}^{n} e^{-\eta_{i,r}|\nu_r|} + j\tilde{x}_r(p)|\nu_r| \right)
\times \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) d\nu_1 \ldots d\nu_n, \quad (15)
\]

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

### IV. The Conditional Performance Index

Consider the integral over \( \nu_n \) in (15),

\[
I_n = \int_{-\infty}^{\infty} \left( \prod_{r=1}^{n} e^{-\eta_{i,r}|\nu_r|} \right) e^{j\tilde{z}_r(p)|\nu_r|} \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) d\nu_n
\]

\[
= \int_{-\infty}^{\infty} e^{j\tilde{z}_r(p)|\nu_r|} \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) d\nu_n
\]

The pdf for the state \( \tilde{x}(k) \) is denoted as \( f_{\tilde{X}_k|\tilde{z}_k} \). The normalized pdf (ucpdf) is denoted as \( \tilde{f}_{\tilde{X}_k|\tilde{z}_k} = f_{\tilde{X}_k|\tilde{z}_k} / \int f_{\tilde{X}_k|\tilde{z}_k} \) where \( \tilde{f}_{\tilde{X}_k|\tilde{z}_k} \) is the pdf of the measurement history and has a known value. In [12, 13], the characteristic function of the ucpdf \( \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) \) is recursively propagated; the characteristic function of the normalized pdf is \( \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) = \tilde{f}_{\tilde{X}_k|\tilde{z}_k}(\nu) / \int \tilde{f}_{\tilde{X}_k|\tilde{z}_k} \) for \( \nu = 0 \). From [12, 13] the form of the characteristic function of the ucpdf at time \( k \) is shown to be

\[
\phi_{\tilde{X}_k|\tilde{z}_k}(\nu) = \sum_{i=1}^{n} g_{i,k}^{j,k}(y_{\nu_i}(\nu)) e^{\nu_i y_{\nu_i}(\nu)} \quad (17a)
\]

where

\[
y_{\nu_i}(\nu) = \sum_{i=1}^{n} g_{i,k}^{j,k}(y_{\nu_i}(\nu)) e^{\nu_i y_{\nu_i}(\nu)} \quad (17b)
\]

and the parameters \( g_{i,k}^{j,k}, \nu_{\nu_i}, \nu_{k.i}, p_{i,k}, \beta_{i,k}, b_{i,k} \) are generated sequentially from \( k = 0 \).

For the MPC algorithm, the characteristic function of the ucpdf is to be propagated through the stochastic dynamics to time \( k + m = p \). This characteristic function \( \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) \) is

\[
\phi_{\tilde{X}_k|\tilde{z}_k}(\nu) = \phi_{\tilde{X}_k|\tilde{z}_k}(\nu) |\phi_{W}(\Phi^{m-1}\Gamma)\nu) \times \cdots \times \phi_{W}(\Phi\Gamma\nu) \]

\[
= \sum_{i=1}^{n} g_{i,k}^{j,k}(y_{\nu_i}(\Phi^{m-1}\Gamma)\nu) e^{\nu_i y_{\nu_i}(\Phi^{m-1}\Gamma)\nu) \times \exp \left\{ -\beta |\phi^{m-1}\Gamma, \nu) \right| - \cdots - \beta |\phi\Gamma, \nu) - \beta |\Gamma, \nu) \right| \}
\]

(18)

In (18) we add \( m \) terms to the sum in \( y_{\nu_i}(\nu) \) of (17c). By combining the exponent in (16) with that in (18), the combined exponent in the integrand of (16) has in total of \( n_{\nu_i} + m + n \) real terms, and the imaginary part is composed of two components. Define the following terms

\[
\tilde{p}_i = p_i, \quad \tilde{a}_{i,k} = \Phi^{m-1}\tilde{a}_{i,k} \quad \text{for } i = 1, \ldots, n_{\nu_i}
\]

\[
\tilde{p}_i = \beta, \quad \tilde{a}_{i,k} = \Phi\Gamma \quad \text{for } i=1, \ldots, n_{\nu_i}
\]

\[
\tilde{p}_i = \eta_i, \quad \tilde{a}_{i,k} = \epsilon_\nu \quad \text{for } i=1, \ldots, n_{\nu_i}
\]

\[
\tilde{b}_{i,k} = b_{i,k}^{k,k} + \tilde{x}_p, \quad n_{\nu_i} = n_{\nu_i} + m + n \quad \text{for } i = n_{\nu_i} + 1, \ldots, n_{\nu_i}
\]

(19a)
Using these definitions, the integrand in (16) becomes

\[ \psi(\nu) = \sum_{i=1}^{n_k} \psi_i(\nu) = \sum_{i=1}^{n_k} g_{i,k}^k (\tilde{\psi}_i(\nu)) \cdot e^{\tilde{\alpha}_{i,k}(\nu)} \]  

(20a)

where

\[ \tilde{\psi}_i(\nu) = \sum_{\ell=1}^{n_{\tilde{\nu}}} q_{\ell,i,k} \sgn \left( \tilde{\alpha}_{i,k}(\nu) \right) \]  

(20b)

\[ \tilde{\alpha}_{i,k}(\nu) = \sum_{\ell=1}^{n_{\tilde{\nu}}} \tilde{g}_{\ell,i,k} \left| \tilde{\alpha}_{i,k}(\nu) \right| + j \tilde{\beta}_{\ell,i,k} \]  

(20c)

The integration in (16) is performed for each element of \( \nu \) in turn. Beginning with \( \nu_n \), decompose \( \nu = [\tilde{\nu} \ \nu_n] \) where \( \tilde{\nu} \in \mathbb{R}^{n-1} \). Then,

\[ I_n = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{i=1}^{n_k} \psi_i(\nu) d\nu = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[ \sum_{i=1}^{n_k} \psi_i(\nu, \tilde{\nu}) \right] d\nu \]  

(21)

The objective is to reduce the inner integral in (21) to a form that is obtained in closed form using the integral formula developed in [12, 13]. First, since \( a_{i,k}^{k,k} \) multiplies \( \nu \) in the sign function in (20b) and the absolute value function in (20c), they are decomposed as \( \tilde{a}_{i,k}^{k,k} = \left[ a_{i,k}^{k,k} \right] \tilde{a}_{i,k}^{k,k} \) \( \tilde{a}_{i,k}^{k,k} \) is a scalar and \( \tilde{a}_{i,k}^{k,k} \) is a vector. Therefore, the inner products in (20b) and (20c) become

\[ \left\langle \tilde{a}_{i,k}^{k,k}, \nu \right\rangle = \left\langle \tilde{a}_{i,k}^{k,k}, \tilde{\nu} \right\rangle - \left\langle -\tilde{a}_{i,k}^{k,k}, \nu_n \right\rangle \]  

(22)

In order to rewrite (22) in a form consistent with the integral formula in [12, 13], \( \tilde{a}_{i,k}^{k,k} \) is divided out of the second term. If \( \tilde{a}_{i,k}^{k,k} = 0 \), then the term \( e^{\tilde{\alpha}_{i,k}^{k,k} \tilde{\nu}} \) loses dependence on \( \nu_n \) and it is removed from the inner integral in (21). Therefore, only \( \tilde{a}_{i,k}^{k,k} \neq 0 \) needs to be considered. Let (22) be rewritten as

\[ \left\langle \tilde{a}_{i,k}^{k,k}, \nu \right\rangle = \left\langle \tilde{a}_{i,k}^{k,k}, \nu \right\rangle + \left\langle -\tilde{a}_{i,k}^{k,k}, \nu_n \right\rangle \]  

(23a)

where

\[ \mu_{i,k}^{k,k} = \left\langle \tilde{a}_{i,k}^{k,k}, \nu \right\rangle \]  

Therefore, the elements in (20b) and (20c) are

\[ q_{i,k}^{k,k} \sgn \left( \mu_{i,k}^{k,k} - \nu_n \right) = \tilde{q}_{i,k}^{k,k} \sgn \left( \mu_{i,k}^{k,k} - \nu_n \right) \]  

(23b)

\[ p_{i,k}^{k,k} \sgn \left( \mu_{i,k}^{k,k} - \nu_n \right) = \tilde{p}_{i,k}^{k,k} \sgn \left( \mu_{i,k}^{k,k} - \nu_n \right) \]  

(23c)

where \( q_{i,k}^{k,k} = \tilde{q}_{i,k}^{k,k} \sgn \left( -\tilde{p}_{i,k}^{k,k} \right) \) and \( p_{i,k}^{k,k} = \tilde{p}_{i,k}^{k,k} \tilde{a}_{i,k}^{k,k} \). Using these definitions, the inner integral of (21) is of the form

\[ \int_{-\infty}^{\infty} \psi_i(\nu(n)) d\nu_n = e^{j \tilde{\alpha}_i(\nu(n))} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[ \sum_{i=1}^{n_k} \psi_i(\nu(n)) \right] d\nu_n \]  

(24)

The convolution integral in (24) is shown in [12, 13] to have a closed form solution composed as a sum with \( n_{\tilde{\nu}} \) \( n_k \) terms, each of which is structurally similar to the terms in \( \psi_i(\nu) \). That is, there will be a new \( g \) function which is a function of signs of inner products of \( \tilde{\nu} \).

Therefore, this integration process can be repeated until all of the integrals are taken, and a closed form solution of the conditional performance index is determined. This pattern will be seen in the next section, where the conditional performance index for the two-state system is explicitly obtained.

V. THE CONDITIONAL PERFORMANCE INDEX FOR A SECOND ORDER SYSTEM

Now, let us limit our discussion to a second order system in order to use the structure for the cpdf’s characteristic function presented in [7]. We use this alternate, two state structure for the characteristic function of the cpdf in order to make the subsequent derivations and computation more tractable. This is due to a simpler structure for the exponential argument in (17c), which produces fewer terms in the sum, as well as a simpler, closed form representation for the \( g_{i,k}^{k,k} \) coefficients in (17a).

The structure for the characteristic function for the ucpdf is given by

\[ \tilde{\phi}_{X_\nu | Z_k}(\nu) = \sum_{i=1}^{N_k} G_i(\nu) \times \exp \left( -\sum_{\ell=1}^{L_k} P^l_{k,i} |B^l_k | \nu + j \sum_{\ell=1}^{2} \left( Z^l_{k,i} |B^l_k | \nu \right) \right) \]  

(25)

which consists of a sum of \( N_k \) similar terms. Each of these terms has a coefficient \( G_i(\nu) \) and an exponential whose argument involves a sum of absolute values equivalent to (17c). In the exponential argument: there is a sum of \( L_k \) absolute value terms; the \( P^l_{k,i} \) and \( Z^l_{k,i} \) values represent real, scalar constants where \( P^l_{k,i} > 0 \); the \( B^l_k \)’s are \( 1 \times 2 \) row vectors, called the fundamental directions; and the \( M^l_k \)’s represent integers that index the fundamental directions that multiply \( \nu \).

All of these parameters correspond with those of the structure for the cpdf’s characteristic function presented in Section IV. The \( P^l_{k,i} \) and \( Z^l_{k,i} \) parameters correspond to \( p_{i,k}^{k,k} \) and \( b_{i,k}^{k,k} \), respectively, and the fundamental directions \( B^l_k \) are the same as the \( a_{i,k}^{k,k} \) vectors.

The coefficients \( G_i(\nu) \) equal the \( g_{i,k}^{k,k} \) from (17a), and are rational polynomials of sums of sign functions, given by

\[ G_i(\nu) = \frac{1}{(2\pi)^n} \left\{ a_i + b_i \sgn \left( B^M_k \nu \right) \sgn \left( B^M_{k,\nu} \right) \right. \]  

\[ + j c_i \sgn \left( B^M_k \nu \right) + j d_i \sgn \left( B^M_{k,\nu} \right) \]  

\[ \times \prod_{\ell=1}^{L_k - 2} \left\{ \frac{1}{j \theta_{k,-\ell} + c + S_{k,-\ell}(B_k \nu)} - \frac{1}{j \theta_{k,-\ell} - c - S_{k,-\ell}(B_k \nu)} \right\} \]  

(26a)
where
\[ S_{k-\ell}(B_k \nu) = \sum_{r=1}^{L_{k-\ell}^+ - \ell - 2} P_{k,i}^\ell \left( B_{k-\ell}^{M_{k,i}} H^T \right) \text{sgn} \left( B_{k-\ell}^{M_{k,i}} \nu \right), \] (26b)
\[ \theta_{k-\ell}^i = z_{k-\ell} - Z_k^1 B_{k-\ell}^{M_{k,i}} H^T - Z_k^2 B_{k-\ell}^{M_{k,i}} H^T. \] (26c)

The arguments \( y_{g_{i,k}}^k(\nu) \) of \( g_{i,k} \), given in (17b), correspond to the \( S_{k-\ell}(B_k \nu) \) defined above.

Next, the transformed objective function in the performance index is given by
\[ L_{x,p}(\nu) = e^{-\eta_1 |\nu_1| + j \bar{x}_1 \nu_1 - \eta_2 |\nu_2| + j \bar{x}_2 \nu_2} \] (27)
where, since \( k \) will be a constant throughout this process, the time subscript of \( \bar{x}(p) \) is replaced with element subscripts as \( \bar{x}(p) = [\bar{x}_1 \ \bar{x}_2] \).

The integrand is continuous. Since \( G_i(\nu) \) is piecewise-constant, its discontinuities lie on the boundaries of these.

The integral \( I_2 \) is over the \( \nu_2 \) variable, but it also contains \( \nu_1 \) in the absolute value terms in the argument of the exponential. In order to solve it, we need to use the integral of absolute values method presented in [7], which requires writing the integral in the form given in
\[ I_2 = \int G_i(\nu_1, \nu_2) \left( \sum_{i=1}^{N_k} \rho_i \text{sgn}(\mu_{i,\ell} - \sigma) \right) \times \exp \left( -\sum_{i=1}^{N_k} \rho_i (\mu_{i,\ell} - \sigma) + j \xi_{i,\ell} \sigma \right) d\sigma, \] (32)
as was done in obtaining (24). That method involves defining a set of scalar constants \( \rho_i \) and \( \xi_{i,\ell} \), as well as scalar variables \( \mu_{i,\ell} \) that depend on \( \nu_1 \), and thus are constants in this integration.

In this derivation, we assume that all the \( B_{k,2}^{M_{k,i}} \neq 0 \). Then, we can construct a \( \{\mu_{i,\ell}\}_{1}^{L_{k}^{1}+1} \) set that transforms \( I_2 \) into the integral of absolute values structure. This assumption does not affect generality, because if one of the \( B_{k,2}^{M_{k,i}} \) does equal zero for some \( i \), then it would multiply the \( \nu_2 \) variable by 0 and thus, that absolute value term would not be a function of \( \nu_2 \) and would come out of \( I_2 \) and integrated later.

The set of variables, denoted \( \{\mu_{i,\ell}\} \), is constructed as
\[ B_{k,1}^{M_{k,i}} \nu_1 + B_{k,2}^{M_{k,i}} \nu_2 = -B_{k,1}^{M_{k,i}} \nu_1 - B_{k,2}^{M_{k,i}} \nu_2 \]
\[ = -B_{k,2}^{M_{k,i}} (\mu_{i,\ell} - \nu_2) \] (33a)
where
\[ \mu_{i,\ell} = \begin{cases} \frac{-B_{k,1}^{M_{k,i}} \nu_1}{B_{k,2}^{M_{k,i}}} & \ell \in \{1, \ldots, L_{k}^1\}, i \in \{1, \ldots, N_{k-1}\} \\ 0 & \ell = L_{k}^1 + 1. \end{cases} \] (33b)

Similarly, for the argument of the exponential, we can construct a set of \( \{\rho_{i,\ell}\}_{1}^{L_{k}^{1}+1} \) as
\[ \rho_{i,\ell} = \left( \frac{P_{k}^{i}}{B_{k,2}^{M_{k,i}}} \right) \ell \in \{1, \ldots, L_{k}^1\}, i \in \{1, \ldots, N_{k-1}\} \]
\[ \ell = L_{k}^1 + 1 \] (33c)
and a scalar number \( \xi_{i,\ell} \) as
\[ \xi_{k} = \bar{x}_2 + \sum_{i=1}^{L_{k}^1} Z_k^1 B_{k,2}^{M_{k,i}}. \] (33d)

The solution to an integral of an exponent of absolute values requires dividing the domain of integration into regions in which the integrand is continuous. Since \( G_i(\nu) \) is piecewise-constant, its discontinuities lie on the boundaries of these.
regions, and hence $G_i(\nu)$ is treated as a constant in each integral. In order for the $G_i(\nu)$ coefficients to be consistent with the form in (32), use the tilde, bar, and hat substitutions similar to the measurement update process described in [7] in order to write $G_i$ in (31) as $\hat{G}_i$:

$$G_i(\nu_1, \nu_2) \rightarrow \hat{G}_i \left( \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \sigma), \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \sigma), \right),$$

$$\sum_{\ell=1}^{L_k^i+1} (L_k^i-2) \tilde{\rho}_\ell \sgn(\mu_\ell - \sigma), \ldots, \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \sigma) \right) \rightarrow \hat{G}_i \left( \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \sigma) \right). \quad (34)$$

The sums in the rational polynomial $G_i(\nu)$ contain different $\tilde{\rho}$ constants but the same set of $\sgn(\mu_\ell - \sigma)$. Hence, it is written as $\hat{G}_i$ for shorthand.

Then, let $\sigma = \nu_2$ in order to write the one-dimensional integral in (32). The solution is given as a sum of $L_k^i + 1$ terms as

$$I_2(\nu_1) = \sum_{m=1}^{L_k^i+1} \exp \left( - \sum_{\ell \neq m}^{L_k^i+1} \rho_\ell |\mu_\ell - \mu_m| + j\hat{\xi}_k^i \mu_m \right)$$

$$\times \left\{ \hat{G}_i \left( \bar{\tilde{\rho}}_m + \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \mu_m) \right) + j\hat{\xi}_k^i + \rho_m + \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \mu_m) \right\}$$

$$\times \left\{ \hat{G}_i \left( \bar{\tilde{\rho}}_m + \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \mu_m) \right) + j\hat{\xi}_k^i + \rho_m + \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \mu_m) \right\}$$

$$- \hat{G}_i \left( \bar{\tilde{\rho}}_m + \sum_{\ell=1}^{L_k^i+1} \tilde{\rho}_\ell \sgn(\mu_\ell - \mu_m) \right) \right\}. \quad (35)$$

This complicated looking (35) simplifies readily into a simple, double-sided scalar integral over $\nu_1$,

$$I_2(\nu_1) = \sum_{m=1}^{L_k^i+1} \{ a_{i,m} + j d_{i,m} \sgn(\nu_1) \} \cdot e^{-D_m \cdot |\nu_1|} + j\hat{\xi}_k^i \nu_1. \quad (36)$$

This simplification is based on algebraic relations used in the estimator’s measurement update process in [7], as well as constants $D_{i,m}$, $\hat{D}_{i,m}$, and $\tilde{D}_{i,m}$ defined as

$$D_{i,m} \cdot |\nu_1| = \sum_{\ell=1,\ell \neq m}^{L_k^i+1} \rho_\ell |\mu_\ell - \mu_m| \quad (37a)$$

$$\hat{D}_{i,m} \cdot \sgn(\nu_1) = \sum_{\ell=1,\ell \neq m}^{L_k^i+1} \rho_\ell \sgn(\mu_\ell - \mu_m) \quad (37b)$$

$$\tilde{D}_{i,m} \cdot \sgn(\nu_1) = \sum_{\ell=1,\ell \neq m}^{L_k^i+1} \rho_\ell \sgn(\mu_\ell - \mu_m) \quad (37c)$$

The index is non-convex and depends on the control input sequence $\{u(k),\ldots,u(p-1)\}$ in a complex way; specifically, the

$$\hat{\xi}_k^i \cdot \nu_1 = \xi_k^i - \mu_m. \quad (37d)$$

Denote the outer integral with respect to $\nu_1$ in (30) by

$$I_1 = \sum_{m=1}^{L_k^i+1} \int \{ a_{i,m} + j d_{i,m} \sgn(\nu_1) \}$$

$$\times \exp \left( -D_m \cdot |\nu_1| + j\hat{\xi}_k^i \nu_1 \right) \exp \left( -\eta_1 |\nu_1| + j\bar{x}_1 \nu_1 \right) \right) \times \exp \left( j \left( 2 \sum_{\ell=1}^{M_k^e} \hat{Z}_{k,i}^\ell B_{k,1}^M \right) \right) d\nu_1$$

$$= \sum_{m=1}^{L_k^i+1} \int \{ a_{i,m} + j d_{i,m} \sgn(\nu_1) \}$$

$$\times \exp \left( -D_m + \eta_1 \right) |\nu_1|$$

$$+ j \left[ \hat{\xi}_k^i \bar{x}_1 + \sum_{\ell=1}^{2} Z_{k,i}^\ell B_{k,1}^M \right] \nu_1 \right) d\nu_1. \quad (38)$$

This integral has a form identical to the measurement update process for a scalar system [4]. Its solution is given by

$$I_1 = \sum_{m=1}^{L_k^i+1} j \left[ \hat{\xi}_k^i \bar{x}_1 + \sum_{\ell=1}^{2} Z_{k,i}^\ell B_{k,1}^M \right]$$

$$\times \left[ a_{i,m} - j d_{i,m} \right] \left[ D_m + \eta_1 \right]$$

$$- \left[ \hat{\xi}_k^i \bar{x}_1 + \sum_{\ell=1}^{2} Z_{k,i}^\ell B_{k,1}^M \right] - \left[ D + \eta_1 \right]$$

$$= \sum_{m=1}^{L_k^i+1} j \left[ \hat{\xi}_k^i \bar{x}_1 + \sum_{\ell=1}^{2} Z_{k,i}^\ell B_{k,1}^M \right]$$

$$\times \left[ a_{i,m} + \eta_1 \right] - \left[ d_{i,m} \bar{x}_1 + \sum_{\ell=1}^{2} Z_{k,i}^\ell B_{k,1}^M \right]$$

$$\times \left[ D_m + \eta_1 \right]^2 \right) \right). \quad (39)$$

Finally, the conditional performance index in (30) is given by

$$J_{\tilde{f}_k} = \frac{1}{2\pi^2} \left( \prod_{i=k}^{p-1} \frac{\gamma_i}{\pi} \right) \frac{D_{i,m} \cdot |\nu_1|}{\sum_{i=1}^{N_k} \sum_{m=1}^{L_k^i+1} \left[ a_{i,m} + \eta_1 \right] - \left[ d_{i,m} \bar{x}_1 + \sum_{\ell=1}^{2} Z_{k,i}^\ell B_{k,1}^M \right]}$$

$$\times \left[ D_m + \eta_1 \right]^2 \right) \right). \quad (40)$$

This closed form conditional performance index is non-convex and depends on the control input sequence $\{u(k),\ldots,u(p-1)\}$ in a complex way; specifically, the
parameters $a_{i,m}$, $d_{1,m}$, and $\bar{x}_1$ depend on the control sequence. Thus, we maximize (40) numerically using the accelerated gradient search method [14]. The optimization is done in two steps: first, the global optimum of the double sum term in (40) without the control weighting terms is optimized with respect to $[\bar{x}_1, \bar{x}_2] = \bar{x}(p)$; then, that final state is used to generate a control sequence as an initial guess for the second, local accelerated gradient search optimization step.

VI. NUMERICAL EXAMPLES

Here, we present two sets of examples, the first of which shows the optimal control $u(0)^*$ versus the measurement $z(0)$ for the first measurement update only, and the second set shows two multi-step examples. All of these examples use a two-step horizon, i.e. $m = 2$, so that there exists a control sequence that can drive our two-state system to the origin over this horizon length. However, as we are using model predictive control, only the first control input of this sequence is applied at that time step.

All of our examples compare our Cauchy optimal model predictive controller with a similar LEG model predictive optimal controller. The LEG estimator assumes that the stochastic inputs are described by the Gaussian pdfs that are closest, in a least squared sense, to the given Cauchy pdfs; and the LEG controller assumes that its objective functions of the state and control resemble scaled Gaussian pdfs that are closest, in the least-squared sense, to the scaled Cauchy pdfs in (10). The LEG controllers’ responses are shown in dashed lines in the figures.

The first set of examples are shown in Fig. 1. These figures show the applied optimal control input at the first time step, $k = 0$, given the first measurement. In the two cases presented, all the systems parameters are the same, except in Fig. 1(a) $\gamma > \alpha_1 = \alpha_2$ (i.e. more measurement than state uncertainty), and in Fig. 1(b) $\alpha_1 = \alpha_2 > \gamma$ (i.e. more state than measurement uncertainty).

The example in Fig. 1(a) shows that the Cauchy controller reduces its control effort to zero as the measurement deviations become large. This is in contrast to the LEG controller, which is linear and thus responds strongly to large measurement deviations. This behavior in the Cauchy controller occurs when the measurement uncertainty is larger than the state uncertainty. In the opposite case shown in Fig. 1(b), the measurement has less uncertainty than the state. In this case, the Cauchy controller’s response closely matches that of the LEG in a neighborhood of the origin, and in fact responds even more strongly than the LEG for large measurement deviations.

The three different curves in both of these figures represent the response for three different control weights: no control weight, $\zeta_0 = 10$, and $\zeta_0 = 5$. As expected, heavier control weights (i.e. smaller $\zeta$) reduce the control effort, but even without any control weighting, the response in Fig. 1(a) goes to zero for large measurement deviations. The fact that this behavior is seen when there is no control weighting implies that the attenuation of the control signal for large measurement deviations is due to the cpdf and not the objective function.

The complexity of evaluating the cost grows as the number of terms increases across time steps, as indicated in (40). For implementable control, this growth needs to be arrested. The full information characteristic function of the upcdf (25) is approximated by a characteristic function of a upcdf conditioned on a fixed sliding window of the most recent measurements, taken here to have length eight. The relative error in the approximation appears to be $10^{-6}$ or smaller.

The second set of examples are shown in Fig. 2. All of the plots show the state, control, and noise histories for the given simulations. The difference between Fig. 2(a) and Fig. 2(b) is that the process noise $\beta$ and the measurement noise $\gamma$ parameter values are interchanged.

It is interesting to compare Fig. 2(a) with Fig. 2(b) in light of Fig. 1. In Fig. 2(a), there is more uncertainty in the state process noise than in the measurement noise. When large measurement deviations occur (such as at $k = 52$), the Cauchy controller’s effort is very small. In contrast, the LEG controller responds with a large control effort that drives the states from their regulated state of zero. However, when large process noise inputs occur, the state deviates and the Cauchy controller applies a larger control effort than the LEG, thus regulating the state more effectively. This suggests that when the measurement noise density parameter dominates the process noise density parameter in constructing the Cauchy controller, the effect of measurement outliers is mitigated, while still responding to state deviations due to process noise.

On the other hand, in Fig. 2(b) there is more uncertainty in the state than in the measurement, and the Cauchy controller behaves very much like the linear LEG controller. The state trajectories and control inputs of the Cauchy and LEG controllers appear equal, but actually their differences are much smaller than the scale of the axis and cannot be seen. This suggests that, when the stochastic parameters allow it, the Cauchy controller follows the measurement more. It responds in a more linear fashion to the measurements, imitating the performance of the LEG controller in that setting.

This behavior is seen again when both controllers face Gaussian noises, as in Figure 2(c). Here, the Cauchy controller closely follows both the control and state trajectories of the LEG, which is the true optimal solution. Hence, the Cauchy is robust under non-impulsive noise environments, as it closely approximates the true optimal solution given by the LEG.

VII. CONCLUSIONS

An optimal stochastic controller was derived for vector-state, linear, discrete-time systems with additive process and measurement Cauchy distributed noises. Since the Cauchy distribution has an undefined mean and an infinite second moment, we cannot use standard objective function, e.g., the expected value of a quadratic function of state and control variables. Therefore, a new and computable objective function was defined. Opposed to previous work, the characteristic function of the cpdf of the state given the measurement history and the Parseval’s equation are used to express the
Fig. 1. Parameters used are: $\eta = [1 \ 1]$, $\beta = 0.02$, $T^T = [1 \ 1]$, $A^T = [1 \ 1]$, and the eigenvalues of $\Phi$ are $0.2455 \pm j0.1523$.

conditional performance index in a closed form. This closed-
form conditional performance index is optimized numerically using an accelerated gradient search. Examples are presented that show how our vector state Cauchy controller compares against an equivalent LEG controller, demonstrating the Cauchy controller’s performance and improved robustness over its Gaussian counterpart.

REFERENCES


