A New Choice of Penalty Function for Robust Multiuser Detection Based on M-Estimation

Babak Seyfe and Shahrokh Valaee

***Abstract*—In this letter, we propose a new robust MUD, called α detector, for non-Gaussian noise. We consider the Gaussian-mixture model for non-Gaussian or impulsive noise. Our technique outperforms the decorrelator and the minimax detectors in highly impulsive noise. The proposed method uses a parametric cost function, where the parameter α is selected using the difference between the asymptotic variance of estimation error of the α detector and that of the minimax detector.**

***Index Terms*—Impulsive noise, M-estimation, minimax detector, robust multiuser detection.**

I. INTRODUCTION

RECENTLY, a robust multiuser detector (MUD) for non-Gaussian noise has been proposed by Wang and Poor [1]. This technique designs the optimum detector for the *worst-case* (least-favorable) model. The minimax MUD has a significant performance gain over the linear decorrelator in impulsive noise. Wang and Poor use the *minimax* approach of Huber [2] to design a MUD that has near-optimum performance for a limited degree of impulsiveness of noise.

Other robust methods also exist in the literature. Vikalo *et al.* [3] introduce a synthesis procedure to design finite impulse response (FIR) MUDs based on H^∞ and mixed H^2/H^∞ design techniques. In [4], the *signature waveform mismatch* is addressed via second-order cone programming. Spasojevic and Wang [5] propose a robust MUD technique based on the slowest-descent search through the minimization of the Huber penalty function. In [6], Tian *et al.* use multiple linear constraints to preserve the output energy that is scattered in multipath channels. Shahbazpanahi *et al.* [7] address the *blind* multiuser receiver based on uncertainties in the covariance matrices of the desired user signature and of the received data.

In this letter, we propose a novel nonlinear penalty function for the M-estimator. The proposed nonlinear penalty function generates a detector that outperforms the minimax detector in highly impulsive noise. We will refer to the proposed detector as the α *detector*.

II. SYSTEM MODEL

Consider a baseband-synchronous *direct-sequence code-division multiple-access* (DS-CDMA) system. At any time instant, the received signal comprises the waveform of K active users plus the ambient noise

$$\mathbf{r} = \mathbf{SAb} + \mathbf{n} \tag{1}$$

where $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K]$ is an $N \times K$ matrix with the columns $\mathbf{s}_k$, $k = 1, 2, \ldots, K$ being the normalized signature vector of the kth user, $\mathbf{A} = \mathrm{diag}(A_1, A_2, \ldots, A_K)$ is the diagonal matrix of the received signal amplitude, and $\mathbf{b} = [b_1, b_2, \ldots, b_K]^T$ is the user symbol vector ($b_k = \pm 1$, $k = 1, 2, \ldots, K$); the superscript T indicates transposition. In (1), $\mathbf{n}$ is a vector of independent and identically distributed (i.i.d.) random variables. Let $\mathbf{\Theta} = \mathbf{Ab}$. Then, our model will be

$$\mathbf{r} = \mathbf{S\Theta} + \mathbf{n}. \tag{2}$$

To estimate the vector $\mathbf{b}$, we need the sign of $\mathbf{\Theta}$.

A. Minimax Multiuser Detection

It is well known that even a slight deviation of the noise density function from the Gaussian distribution causes a substantial degradation of the least square (LS) estimate [1]. The robustness of an estimator refers to its insensitivity to small changes in the underlying statistical model [8]. The LS estimate can be made robust by using the class of M-estimators proposed by Huber [2]. His approach has recently been used by Wang and Poor in robust MUD [1].

Consider

$$\widehat{\mathbf{\Theta}} = \arg \min_{\mathbf{\Theta} \in R^K} \sum_{j=1}^{N} \rho \left(r_j - \sum_{k=1}^{K} s_j^k \theta_k \right) \tag{3}$$

where $\rho(.)$ is an increasing function of the residuals, s_j^k is the jth component of the kth user's signature waveform, θ_k is the kth component of $\mathbf{\Theta}$, and r_j is the jth component of $\mathbf{r}$. Let $\rho(x)$ have a derivative $\psi(x) = \rho'(x)$. Then, the solution of (3) satisfies

$$\mathbf{S}^T \boldsymbol{\psi}\left(\mathbf{r} - \mathbf{S}\widehat{\mathbf{\Theta}}\right) = \mathbf{0}_K \tag{4}$$

where $\boldsymbol{\psi}(\mathbf{x}) = [\psi(x_1), \psi(x_2), \ldots, \psi(x_N)]^T$ for any $\mathbf{x} \in R^N$, and $\mathbf{0}_K$ denotes the K-dimensional zero vector. An estimator defined by (3) is called an M-estimator. Note that for $\rho(x) = -\log f(x)$, the conventional maximum-likelihood (ML) estimator is obtained; hence, the name M-estimator.

The robust minimax MUD, as suggested in [1], uses the following derivative of the penalty function:

$$\psi_{A,H}(x) = \begin{cases} \frac{x}{\sigma^2}, & \text{if } |x| \le \gamma\sigma^2 \\ \gamma \mathrm{sgn}(x), & \text{if } |x| > \gamma\sigma^2 \end{cases} \tag{5}$$

Paper approved by X. Wang, the Editor for Modulation, Detection, and Equalization of the IEEE Communications Society. Manuscript received August 15, 2003; revised March 24, 2004 and August 16, 2004.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (email: valaee@comm.utoronto.ca).

Digital Object Identifier 10.1109/TCOMM.2004.842001

where $\gamma = 1.5/\sigma$ and σ^2 is the noise variance. It has been shown that for a noise with moderate degree of impulsiveness, the bit-error rate (BER) of the proposed minimax detector is close to that of the ML detector [1]. In the following section, we will propose a new penalty function for the M-estimator that has a better performance than the minimax detector for highly impulsive noise.

III. The α Detector

The minimax MUD is designed for the worst-case (least-favorable) density function of noise [1]. Huber [2] states that it might be worthwhile to increase the maximum risk slightly beyond its minimax value in order to gain a better performance at very long-tailed distributions. Hampel [8] shows that, to achieve robustness, it is necessary for the $\psi(x)$ to be bounded and continuous. Most M-estimators use a monotone increasing function for $\psi(x)$. However, it has been shown that there exists a family of nonmonotone increasing functions that begets a very good regression estimator [9]. It can be shown that in most cases, nonmonotone $\psi(x)$ functions have the same behavior as the monotone increasing functions [10]. In this section, we devise a new robust MUD for non-Gaussian noise using a nonmonotone $\psi(x)$.

We start with using the Gaussian-mixture model of noise that is defined as a noise with the probability density function (PDF)

$$f_{GG}(x) = (1-\epsilon)\frac{1}{\sqrt{2\pi\nu^2}}e^{-x^2/2\nu^2} + \epsilon\frac{1}{\sqrt{2\pi\kappa\nu^2}}e^{-x^2/2\kappa\nu^2} \tag{6}$$

where $0 \le \epsilon \le 1$ indicates the probability that impulses occur, and $\kappa > 1$ is the variance factor of the impulsive component. The Gaussian-mixture model (6) serves as an approximation to the more fundamental Middleton Class-A noise, and has been used extensively to model physical noise in radio and acoustic channels [1].

We propose the following derivative of the penalty function that exponentially suppresses the large values of noise:

$$\psi_\alpha(x) = xe^{-\alpha x^2}, \quad \alpha > 0. \tag{7}$$

We note that for large values of noise, $\psi_\alpha(x)$ is exponentially decreasing. Therefore, we expect that the proposed detector will substantially suppress excessive noise amplitudes.

Using (7), we get the following penalty function for our detector:

$$\rho_\alpha(x) = \frac{-1}{2\alpha}e^{-\alpha x^2} + C, \quad \alpha > 0 \tag{8}$$

where α is a parameter of design and C is a constant. Note that the performance of the detector does not depend on C. Therefore, C can be chosen arbitrarily. $\rho_\alpha(x)$ is a function that increases less rapidly than x^2. We call the detector that is generated by this nonlinearity the α detector. Fig. 1 illustrates the nonlinearity of the ML estimator ($-\log f(x)$) for Gaussian-mixture noise, and also the penalty function of the α detector for two values of $\alpha = 0.1$ and $\alpha = 0.2$ for $C = 1/2\alpha + 1$. In this figure, for the Gaussian-mixture model, we have $\epsilon = 0.01, \nu = 1, \kappa = 100$, and $\epsilon = 0.1, \nu = 1, \kappa = 100$. Note that at each case, the penalty function of the α detector can be viewed as an approximation to the penalty function of the corresponding noise model.

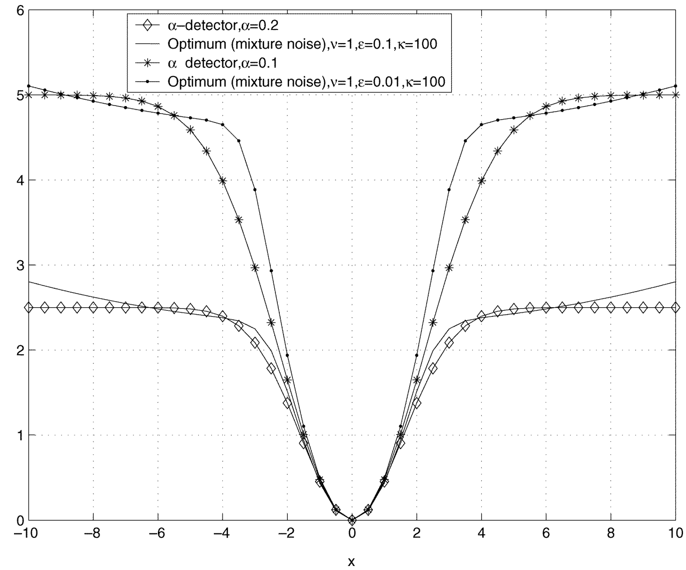

Fig. 1. Penalty function $\rho(x)$ of the ML estimator of the Gaussian-mixture model, and the penalty function of the α detector for $\alpha = 0.1$ and $\alpha = 0.2$.

Swami and Sadler [11] propose a nonlinearity which is similar to the α detector, but not identical. Unlike the α detector, their proposed nonlinearity has two breakpoints and some control parameters. A nonlinearity similar to (8), with a minor difference, was used by Holland and Welsch [9], [12] for a different application. As indicated by Zhang [12], it is very difficult to select a penalty function for general use. It seems that each penalty function has superior performance in some applications. Our study shows that our proposed detector has a better performance, as compared with minimax and decorrelator, for MUD in highly impulsive noise.

For $C = 1/2\alpha + 1$ and small values of α, our detector tends to the linear case (decorrelator). To show this, note that

$$\lim_{\alpha \to 0} \rho_\alpha(\mathbf{x}) = \rho_0(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2 + 1, \quad \mathbf{x} \in R^N \tag{9}$$

where $\|\mathbf{x}\|$ is the Euclidean norm of vector $\mathbf{x}$. Then, $\rho_0(x)$ is the decorrelator detector [1]. Also note that for $|x| \to \infty$, we have $\rho_\alpha(x) \to 1/2\alpha$. Therefore, for input samples with large amplitude, the penalty function of the α detector does not increase. However, the penalty function of the ML estimator increases with $|x|$, and therefore, the ML estimator is not robust [8]. The influence function (IF) (or influence curve) is the most useful heuristic tool of robust statistics. By definition, an estimator is robust if its IF is bounded [8]. Since $|\psi_\alpha(x)|$ is bounded and has continuous derivatives, the α detector has a bounded IF, and hence, is robust.

We use an iterative Newton-type algorithm to estimate the data vector $\mathbf{\Theta}$. Let $\mathbf{\Theta}^{(\mathbf{i})}$ be the estimate of $\mathbf{\Theta}$ at the ith step. Then, the new estimate of $\mathbf{\Theta}$ is given by

$$\mathbf{\Theta}^{(\mathbf{i+1})} = \mathbf{\Theta}^{(\mathbf{i})} + \frac{1}{\mu}(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\psi_\alpha\left(\mathbf{r} - \mathbf{S}\mathbf{\Theta}^{(\mathbf{i})}\right) \tag{10}$$

where μ is a constant step size and $\psi_\alpha\left(\mathbf{r} - \mathbf{S}\mathbf{\Theta}^{(\mathbf{i})}\right)$ indicates the member-wise application of $\psi_\alpha(x)$ to the elements of the vector $\mathbf{r} - \mathbf{S}\mathbf{\Theta}^{(\mathbf{i})}$. The algorithm (10) converges if we use the ML estimation of $\mathbf{\Theta}$ in Gaussian noise for $\mathbf{\Theta}^{(\mathbf{0})}$ [9].

A. Selecting α and Computing the Performance

In this subsection, we devise a method to select an appropriate value for the parameter α in the proposed detector. We select α by comparing the asymptotic variance of the estimation error of the proposed detector with that of the minimax detector. We use the minimax detector in the Gaussian-mixture noise model, and then set α for the desired range of noise impulsiveness.

It can be shown that, within mild regularity conditions, the asymptotic variance of estimation error for an M-estimator of $\boldsymbol{\Theta}$, at the noise probability distribution function F, is given by [2]

$$V(\psi; F) = \frac{\int \psi^2 dF}{\left[\int \psi' dF\right]^2}. \tag{11}$$

For the α detector and in Gaussian-mixture noise, after algebraic manipulations, we have

$$V_\alpha(\epsilon, \kappa) = \nu^2 \frac{\left[\frac{1-\epsilon}{(1+4\alpha\nu^2)^{3/2}} + \frac{\epsilon\kappa}{(1+4\kappa\alpha\nu^2)^{3/2}}\right]}{\left[\frac{1-\epsilon}{(1+2\alpha\nu^2)^{3/2}} + \frac{\epsilon}{(1+2\kappa\alpha\nu^2)^{3/2}}\right]^2}. \tag{12}$$

Wang and Poor [1] compute the asymptotic variance of estimation error, $V_H(\epsilon, \kappa)$, of the minimax MUD for Gaussian-mixture noise.

To choose the value of α, we compare the asymptotic variance of estimation error of the α detector with that of the minimax (as a standard suboptimum) detector. We define a distance measure as

$$D_H(\alpha) = \int_{\epsilon_1}^{\epsilon_2} \int_{\kappa_1}^{\kappa_2} \left[V_\alpha(\epsilon, \kappa) - V_H(\epsilon, \kappa)\right]^2 d\kappa d\epsilon \tag{13}$$

where the limits (ϵ_1, ϵ_2) and (κ_1, κ_2) are selected so as to span the range of variation of the noise-model parameters. Our observations show that $D_H(\alpha)$ is a positive function of α and has a unique minimum, therefore, it is an appropriate distance metric. Indeed, (13) is the ℓ_2-norm in the space of integrable functions on the rectangle $[\epsilon_1, \epsilon_2] \times [\kappa_1, \kappa_2]$.

Fig. 2 illustrates $D_H(\alpha)$ as a function of the control parameter α for $0.01 \leq \epsilon \leq 0.1$ and $1 \leq \kappa \leq 100$. It shows that the minimum distance between the variance of the two detectors, for the given range of ϵ and κ, is located at $\alpha^* = 0.15$. Note that α^* is a function of the range of variations of ϵ and κ. For each application, the range of variations of ϵ and κ should be measured and used to select an appropriate α. Since we are interested in $0.01 \leq \epsilon \leq 0.1$ and $1 \leq \kappa \leq 100$, in the following, we will use $\alpha = 0.15$.

We study the performance of the α detector by comparing its asymptotic variance with that of the ML estimator. The penalty function of the ML detector for any noise model $f(x)$ is $\rho(x) = -\log f(x)$. Then, the asymptotic variance of estimation error for the Gaussian-mixture noise will be the inverse of the Fisher information [2]. We find the asymptotic variance of the estimation error of the ML estimator for Gaussian-mixture noise and use it to get the relative efficiency (RE), defined as

$$\mathrm{RE}_{\alpha,\mathrm{ML}}(\epsilon, \kappa) = \log \frac{V_{\mathrm{ML}}(\epsilon, \kappa)}{V_\alpha(\epsilon, \kappa)}. \tag{14}$$

Fig. 3 illustrates $\mathrm{RE}_{\alpha,\mathrm{ML}}(\epsilon, \kappa)$ as a function of ϵ and κ. Here, we have assumed $\alpha = 0.15, 1$ and $(1-\epsilon)\nu^2 + \epsilon\kappa\nu^2 = 1$. Note

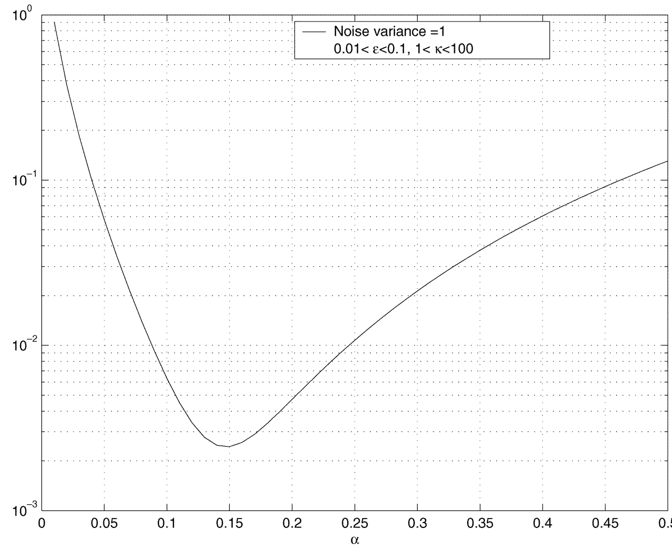

Fig. 2. Distance between the asymptotic variance of the estimation error of the α detector and that of the minimax detector for $0.01 \leq \epsilon \leq 0.1$ and $1 \leq \kappa \leq 100$.

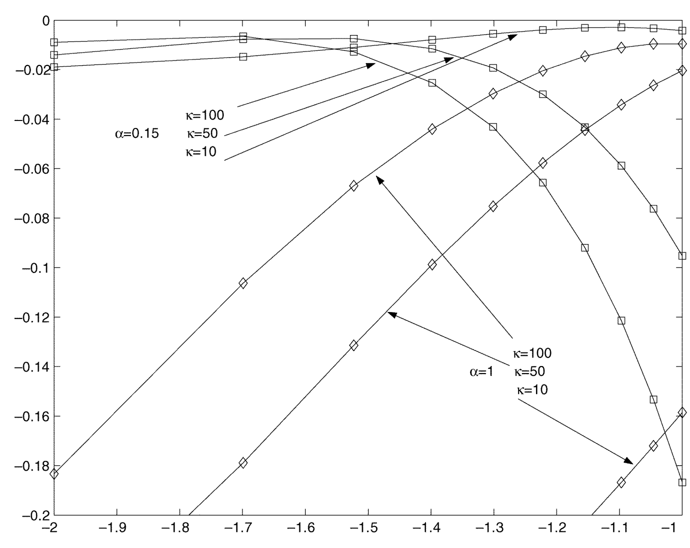

Fig. 3. RE of the α detector to the ML detector as a function of ϵ for $\kappa = 10, 50, 100, \alpha = 0.15, 1$, and the noise variance $(1-\epsilon)\nu^2 + \epsilon\kappa\nu^2 = 1$.

that for $\alpha = 1$ and large values of ϵ and κ (highly impulsive noise), the asymptotic variance of the estimation error of the α detector approaches that of the ML detector. It means that the α detector has a good performance in highly impulsive noise for $\alpha = 1$. This figure also shows that for $\alpha = 0.15$, the proposed detector has a good performance in impulsive noise. For instance, the α detector has a performance near the ML performance for $\alpha = 0.15$, $0.01 \leq \epsilon \leq 0.04$, and $10 \leq \kappa \leq 100$.

IV. Simulation Results

Consider a synchronous DS-CDMA system in which the spreading sequence for each user is a shifted m-sequence [1]. Noting that α-stable distributions are the well-known models for impulsive noise, we use the Gaussian-mixture noise model that is a very good approximation of α-stable distributions [5]. The signal-to-noise ratio (SNR) is defined as the ratio of

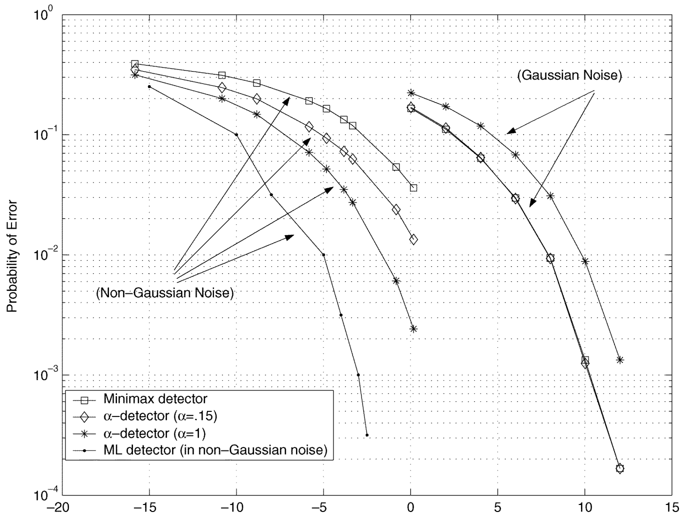

Fig. 4. BER versus SNR in a synchronous DS-CDMA system with Gaussian and impulsive noise for the α detector ($\alpha = 0.15, 1$), the minimax detector, and the ML detector. Here, $N = 31$, $K = 6$, $\epsilon = 0.1$, $\kappa = 100$, the power of the desired user (the first user) is $P_1 = P$, and the powers of the interferer users are $P_2 = P$, $P_3 = P_4 = 2P$, and $P_5 = P_6 = 4P$.

the received signal power to the noise variance. As the initial condition in the recursion (10), we use the LS solution of $\mathbf{\Theta}$.

Fig. 4 illustrates the performance of three detectors for the Gaussian and impulsive noises. The processing gain of the DS-CDMA signal is $N = 31$, and the number of users is $K = 6$. The impulsive noise parameters are $\epsilon = 0.1$ and $\kappa = 100$. Here, the second user has the same power as the desired user (the first user). The third and the fourth users have 3 dB, and the fifth and the sixth users have 6 dB more power than the first user. The figure shows that in Gaussian noise, the α detector for $\alpha = 0.15$ has a performance similar to the minimax detector, but for $\alpha = 1$, it performs worse than the minimax detector. The figure also shows that the α detector has a considerable gain over the minimax detector in highly impulsive noise. As noticed, the α detector can easily handle the near–far problem in CDMA networks.

V. CONCLUSIONS

In this letter, we have proposed a new penalty function for robust MUDs based on M-estimation. The proposed method is called the α detector. We have proposed a metric to select the parameter α in the desired range of noise impulsiveness. We have shown that for non-Gaussian noise, the α detector outperforms the minimax and the decorrelator detectors.

REFERENCES

[1] X. Wang and H. V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Trans. Signal Process.*, vol. 47, pp. 289–305, Feb. 1999.
[2] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
[3] H. Vikalo, B. Hassibi, and T. Kailath, "On robust multiuser detection," in *Proc. 34th Asilomar Conf. Signals, Syst., Computers*, vol. 2, Nov. 2000, pp. 1168–1172.
[4] S. Cui, Z.-Q. Luo, and Z. Ding, "Robust CDMA signal detection in the presence of user and interference signature mismatch," in *Proc. IEEE SPAWC Workshop*, Mar. 2001, pp. 221–224.
[5] P. Spasojevic and X. Wang, "Improved robust multiuser detection in non-Gaussian noise," *IEEE Signal Process. Lett.*, vol. 8, pp. 83–86, Jan. 2001.
[6] Z. Tian, K. L. Bell, and H. L. Van Trees, "Robust constrained linear receivers for CDMA wireless systems," *IEEE Trans. Signal Process.*, vol. 49, pp. 83–86, Jul. 2001.
[7] S. Shahbazpanahi and A. B. Gershman, "Robust blind multiuser detection for synchronous CDMA systems," in *Proc. IEEE ICASSP*, Apr. 2003, pp. 53–56.
[8] F. R. Hampel, E. M. Ronchetti, P. J. Rousseuw, and W. A. Stahl, *Robust Statistics: The Approach Based on Influence Functions*: Wiley, 1986.
[9] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least squares," *Commun. Statist.*, pt. A, vol. 6, pp. 813–827, 1977.
[10] P. J. Bickel, "One-step Huber estimates in the linear model," *J. Amer. Statist. Assoc.*, vol. 70, pp. 428–434, 1975.
[11] A. Swami and B. Sadler, "On some detection and estimation problems in heavy-tailed noise," *Signal Process.*, vol. 82, pp. 1829–1846, 2002.
[12] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image, Vis. Comput.*, vol. 15, pp. 59–76, 1997.