Robust Multiuser Detection in Non-Gaussian Channels

Xiaodong Wang, *Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

Abstract— In many wireless systems where multiuser detection techniques may be applied, the ambient channel noise is known through experimental measurements to be decidedly non-Gaussian, due largely to impulsive phenomena. The performance of many multiuser detectors can degrade substantially in the presence of such impulsive ambient noise. In this paper, we develop robust multiuser detection techniques for combating multiple-access interference and impulsive noise in CDMA communication systems. These techniques are based on the M-estimation method for robust regression. Analytical and simulation results show that the proposed robust techniques offer significant performance gain over linear multiuser detectors in impulsive noise, with little attendant increase in computational complexity.

We also develop a subspace-based technique for blind adaptive implementation of the robust multiuser detectors, which requires only the signature waveform and the timing of the desired user in order to demodulate that user's signal. The robust multiuser detection technique and its blind adaptive version can be applied to both synchronous and asynchronous CDMA channels.

Index Terms—Impulsive noise, M-estimation, multiuser detection, signal subspace tracking.

I. INTRODUCTION

RECENT years have seen a significant interest in advanced signal processing techniques for enhancing the performance of nonorthogonal signaling schemes for multiple-access communications. These techniques generally fall under the category of multiuser detection [25], which refers to optimum or near-optimum demodulation in such situations. One of the key theoretical contributions in this field has been to show that the use of multiuser detection (or derivative signal processing techniques) can return performance in multiuser channels to that of corresponding single-user channels, or at least to a situation in which performance is no longer limited by the multiple-access interference (MAI). By and large, the study of this problem has focused on the situation in which the ambient noise is additive white Gaussian noise (AWGN). This has been an appropriate model in the previous studies since the focus there has been on the mitigation of the most severe noise source—the MAI. However, as increasingly practical techniques for multiuser detection become available, such as adaptive and blind adaptive multiuser detection methods [11],

the situation in which practical multiple-access channels will be ambient-noise limited can be realistically envisioned.

In many physical channels, such as urban and indoor radio channels [4], [5], [17], [18], [20] and underwater acoustic channels [6], [19], the ambient noise is known through experimental measurements to be decidedly non-Gaussian due to the impulsive nature of man-made electromagnetic interference and a great deal of natural noise as well. [For recent measurement results of impulsive noise in outdoor/indoor mobile and portable radio communications, see [4], [5], and the references therein.] It is widely known in the single-user context that non-Gaussian noise can be quite detrimental to the performance of conventional systems based on the Gaussian assumption, whereas it can actually be beneficial to performance if appropriately modeled and ameliorated. Neither of these properties is surprising. The first is a result of the lack of robustness of linear and quadratic type signal processing procedures to many types of non-Gaussian statistical behavior [14]. The second is a manifestation of the well-known least favorability of Gaussian channels.

In view of the lack of realism of an AWGN model for ambient noise arising in many practical channels in which multiuser detection techniques may be applied, natural questions arise concerning the applicability, robustness, and performance of multiuser detection techniques for non-Gaussian multiple-access channels. Although performance indices such as mean-square-error (MSE) and signal-to-interference-plus-noise ratio (SINR) for linear multiuser detectors are not affected by the distribution of the noise (only the spectrum matters), the more crucial bit-error rate can depend heavily on the shape of the noise distribution. The results of an early study of error rates in non-Gaussian direct-sequence code-division multiple-access (DS-CDMA) channels are found in [1]–[3], in which the performance of the conventional and modified conventional detectors is shown to depend significantly on the shape of the ambient noise distribution. In particular, impulsive noise can severely degrade the error probability for a given level of ambient noise variance. In the context of multiple-access capability, this implies that fewer users can be supported with conventional detection in an impulsive channel than in a Gaussian channel. However, since non-Gaussian noise can, in fact, be beneficial to system performance if properly treated, the problem of joint mitigation of structured interference and non-Gaussian ambient noise is of interest [22]. An approach to this problem for narrowband interference (NBI) suppression in spread-spectrum systems is described in [9]. A recent study [23] has shown that the performance gains afforded

by maximum likelihood (ML) multiuser detection in impulsive noise can be substantial when compared with optimum multiuser detection based on a Gaussian noise assumption. However, the computational complexity of ML detection is quite high, and therefore, effective near-optimal multiuser detection techniques in non-Gaussian noise are needed. In this paper, we consider the MAI mitigation problem in DS-CDMA channels with non-Gaussian ambient noise.

This paper is organized as follows. In Section II, the signal model for a DS-CDMA communication system, as well as the impulsive channel noise model, is described. In Section III, we propose and analyze a robust technique for multiuser detection is non-Gaussian channels, which is essentially a robust version of the linear decorrelating multiuser detector. The robustification is based on the M-estimation method found in robust statistics. We show both analytically and through simulations that this robust multiuser detector offers significant performance gain over the linear decorrelating detector when the ambient channel noise in non-Gaussian, with little attendant increase in computational complexity. In Section IV, we develop a blind adaptive technique for implementing the robust multiuser detector, which requires the prior knowledge of only the signature waveform of the desired user, in order to robustly demodulate that user's signal. This technique is based on tracking the signal subspace components of the received signal. The analyses in Sections III and IV are based on a synchronous CDMA model. In Section V, we show that the robust and blind adaptive robust multiuser detection techniques developed in this paper can be applied to the more general asynchronous CDMA model as well. Section VI contains some concluding remarks.

II. SYSTEM MODEL

Consider a baseband digital DS-CDMA network operating with a coherent BPSK modulation format. The waveform received by a given terminal in such a network can be modeled as consisting of a set of superimposed modulated data signals observed in additive noise

$$r(t) = S(t) + n(t), \quad -\infty < t < \infty \tag{1}$$

where $S(t)$ and $n(t)$ represent the useful signal and the ambient channel noise, respectively. The ambient noise is assumed to be white. The useful signal $S(t)$ is comprised of the data signals of K active users in the channel and can be written as

$$S(t) = \sum_{k=1}^{K} A_k \sum_{i=0}^{M-1} b_k(i) s_k(t - iT - \tau_k) \tag{2}$$

where M is the number of data symbols per user in the data frame of interest, T is the symbol interval, and where A_k, τ_k, $\{b_k(i); \ i = 0, 1, \ldots, M-1\}$, and $\{s_k(t); \ 0 \le t \le T\}$ denote, respectively, the received amplitude, delay, symbol stream, and normalized signaling waveform of the kth user. It is assumed that $s_k(t)$ is supported only on the interval $[0, T]$ and has unit energy and that $\{b_k(i)\}$ is a collection of independent equiprobable ± 1 random variables. For the direct-sequence spread-spectrum (DS-SS) multiple-access format, the

user signaling waveforms are of the form

$$s_k(t) = \sum_{j=0}^{N-1} \beta_j^k \varphi(t - jT_c), \quad t \in [0, T] \tag{3}$$

where N is the processing gain, $(\beta_0^k, \beta_1^k, \ldots, \beta_{N-1}^k)$ is a signature sequence of ± 1's assigned to the kth user, and φ is a normalized chip waveform of duration T_c, where $NT_c = T$.

For the sake of simplicity of discussion, when developing the robust multiuser detection algorithms in the next two sections, we restrict our attention to the synchronous case of model (2), in which $\tau_1 = \tau_2 = \cdots = \tau_K = 0$. This does not incur any loss of generality. As will be shown in Section V, the robust techniques developed in Sections III and IV can be readily applied to asynchronous channels with a properly windowed received signal. For the synchronous case of model (2), to demodulate the ith symbols of the K users, $\{b_k(i)\}_{k=1}^{K}$, it is sufficient to consider the received signal during the ith signaling interval, i.e.,

$$r(t) = \sum_{k=1}^{K} A_k b_k(i) s_k(t - iT) + n(t), \quad t \in [iT, (i+1)T). \tag{4}$$

At the receiver, the received signal $r(t)$ is first filtered by a chip-matched filter and then sampled at the chip rate. The resulting discrete-time signal sample corresponding to the jth chip of the ith symbol is given by

$$r_j(i) \triangleq \int_{iT+jT_c}^{iT+(j+1)T_c} r(t) \varphi(t - iT - jT_c) \, dt$$

$$= \sum_{k=1}^{K} A_k b_k(i) s_j^k + n_j(i) \tag{5}$$

or in vector form

$$\underline{r}(i) = \sum_{k=1}^{K} A_k b_k(i) \underline{s}_k + \underline{n}(i) \tag{6}$$

where $\underline{s}_k \triangleq [s_0^k \ \cdots \ s_{N-1}^k]^T = \frac{1}{\sqrt{N}} [\beta_0^k \ \cdots \ \beta_{N-1}^k]^T$ is the normalized signature sequence of the kth user, and $\underline{n}(i) \triangleq [n_0(i) \ \cdots \ n_{N-1}(i)]^T$ is the channel ambient noise sample vector at the ith symbol interval. It is assumed that the sequence of noise samples $\{n_j(i)\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with a non-Gaussian distribution.

In this paper, we adopt the commonly used two-term Gaussian mixture model for the additive noise samples $\{n_j(i)\}$. The probability density function (pdf) of this noise model has the form

$$f = (1 - \epsilon) \mathcal{N}(0, \nu^2) + \epsilon \mathcal{N}(0, \kappa \nu^2) \tag{7}$$

with $\nu > 0$, $0 \le \epsilon \le 1$, and $\kappa \ge 1$. Here, the $\mathcal{N}(0, \nu^2)$ term represents the nominal background noise, and the $\mathcal{N}(0, \kappa \nu^2)$ term represents an impulsive component, with ϵ representing the probability that impulses occur. It is usually of interest to study the effects of variation in the shape of a distribution on

the performance of the system by varying the parameters ϵ and κ with fixed total noise variance

$$\sigma^2 \triangleq (1 - \epsilon)\nu^2 + \epsilon\kappa\nu^2. \qquad (8)$$

This model serves as an approximation to the more fundamental Middleton Class A noise model [19], [29] and has been used extensively to model physical noise arising in radio and acoustic channels.

III. MULTIUSER DETECTION VIA ROBUST REGRESSION

In this section, we propose and analyze some robust techniques for multiuser detection in non-Gaussian ambient noise CDMA channels, which are essentially robust versions of the linear decorrelating multiuser detector. In Section III-A, we consider the equivalence between linear decorrelating multiuser detection and least-square regression. In Section III-B, we propose robustifying the linear decorrelator via M-estimation-based robust regression techniques. In Section III-C, we analyze the asymptotic performance of the proposed robust multiuser detectors. In Section III-D, we discuss issues arising in the implementation of the robust multiuser detectors. In Section III-E, we provide some simulation examples to demonstrate the performance gains afforded by the proposed robust multiuser detection methods in non-Gaussian noise.

A. Least Squares Regression and the Linear Decorrelator

Consider the synchronous signal model (6). For simplicity, we drop the symbol index i and denote $\theta_k \triangleq A_k b_k$. Then, (6) can be rewritten as

$$r_j = \sum_{k=1}^{K} s_j^k \theta_k + n_j, \quad j = 1, \ldots, N \qquad (9)$$

or in matrix notation

$$\underline{r} = \underline{S}\,\underline{\theta} + \underline{n} \qquad (10)$$

where $\underline{S} \triangleq [\underline{s}_1 \ \underline{s}_2 \ \cdots \ \underline{s}_K]$, and $\underline{\theta} \triangleq [\theta_1 \ \theta_2 \ \cdots \ \theta_K]^T$. Consider the linear regression problem of estimating the K unknown parameters $\theta_1, \theta_2, \ldots, \theta_K$ from the N observations $r_1, r_2, \ldots, r_N$ in (9). Classically, this problem can be solved by minimizing the sum of squared errors, i.e., through the least-squares (LS) method

$$\hat{\underline{\theta}}_{\mathrm{LS}} = \arg\min_{\underline{\theta}} \sum_{j=1}^{N} \left(r_j - \sum_{k=1}^{K} s_j^k \theta_k \right)^2$$
$$= \arg\min_{\underline{\theta}} \|\underline{r} - \underline{S}\,\underline{\theta}\|^2. \qquad (11)$$

If $n_j \sim \mathcal{N}(0, \sigma^2)$, then the pdf of the received signal $\underline{r}$ under the true parameters $\underline{\theta}$ is given by

$$f_{\underline{\theta}}(\underline{r}) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{\|\underline{r} - \underline{S}\,\underline{\theta}\|^2}{2\sigma^2} \right). \qquad (12)$$

It is easily seen from (12) that the maximum likelihood estimate of $\underline{\theta}$ under the i.i.d. Gaussian noise samples is given

by the LS solution $\hat{\underline{\theta}}_{\mathrm{LS}}$ in (11). Upon differentiating (11), $\hat{\underline{\theta}}_{\mathrm{LS}}$ is the solution to the linear system of equations

$$\sum_{j=1}^{N} \left(r_j - \sum_{l=1}^{K} s_j^l \theta_l \right) s_j^k = 0, \quad k = 1, \ldots, K \qquad (13)$$

or in matrix form

$$\underline{S}^T \underline{S}\,\underline{\theta} = \underline{S}^T \underline{r}. \qquad (14)$$

Define the cross-correlation matrix of the signature waveforms of all users as $\underline{R} \triangleq \underline{S}^T \underline{S}$. Assuming that the user signature waveforms are linearly independent, i.e., $\underline{S}$ has a full rank K, then $\underline{R}$ is invertible, and the LS solution to (13) or (14) is given by

$$\hat{\underline{\theta}}_{\mathrm{LS}} = (\underline{S}^T \underline{S})^{-1} \underline{S}^T \underline{r}$$
$$= \underline{R}^{-1} \underline{S}^T \underline{r}. \qquad (15)$$

We observe from (15) that the LS estimate $\hat{\underline{\theta}}_{\mathrm{LS}}$ is exactly the output of the linear decorrelating multiuser detector for the K users [16]. This is not surprising since the linear decorrelating detector gives the maximum likelihood estimate of the product of the amplitude and the data bit $\theta_k = A_k b_k$ in Gaussian noise [16]. Given the estimate $\hat{\theta}_k$, the estimated amplitude and the data bit are then determined by

$$\hat{A}_k = |\hat{\theta}_k|, \qquad (16)$$
$$\hat{b}_k = \mathrm{sgn}(\hat{\theta}_k). \qquad (17)$$

B. Robust Multiuser Detection via M-Estimation-Based Regression

It is well known that the LS estimate is very sensitive to the tail behavior of the noise density. Its performance depends significantly on the Gaussian assumption, and even a slight deviation of the noise density from the Gaussian distribution can, in principle, cause a substantial degradation of the LS estimate. Since the linear decorrelating multiuser detector is in the form of the LS solution to a linear regression problem, it can be concluded that its performance is also sensitive to the tail behavior of the noise distribution. As will be demonstrated in Section III-E, the performance of the linear decorrelating detector degrades substantially if the ambient noise deviates even slightly from a Gaussian distribution. In this section, we consider some robust versions of the decorrelating multiuser detector, which are nonlinear in nature. (Recall that robustness of an estimator refers to its insensitivity to small changes in the underlying statistical model.)

The LS estimate corresponding to (11) and (13) can be robustified by using the class of M-estimators proposed by Huber [13]. Instead of minimizing a sum of squared residuals as in (11), Huber proposed to minimize a sum of a less rapidly increasing function ρ of the residuals

$$\hat{\underline{\theta}} = \arg\min_{\underline{\theta} \in \mathcal{R}^K} \sum_{j=1}^{N} \rho\left(r_j - \sum_{k=1}^{K} s_j^k \theta_k \right). \qquad (18)$$

Suppose that ρ has a derivative $\psi = \rho'$; then, the solution to (18) satisfies the implicit equation

$$\sum_{j=1}^{N} \psi \left(r_j - \sum_{l=1}^{K} s_j^l \theta_l \right) s_j^k = 0, \quad k = 1, \ldots, K \tag{19}$$

or in vector form

$$\underline{S}^T \psi(\underline{r} - \underline{S}\,\underline{\theta}) = \underline{0}_K, \tag{20}$$

where $\psi(\underline{x}) \triangleq [\psi(x_1), \ldots, \psi(x_K)]^T$ for any $\underline{x} \in \mathcal{R}^K$, and $\underline{0}_K$ denotes a K-dimensional zero vector. An estimator defined by (18) or (19) is called an *M-estimator*. The name "*M-estimator*" comes from "maximum likelihood type estimator" [13] since the choice of $\rho(x) = -\log f(x)$ gives the ordinary maximum likelihood estimate. If ρ is convex, then (18) and (19) are equivalent; otherwise, (19) is still very useful in searching for the solution to (18). To achieve robustness, it is necessary that ψ be bounded and continuous. Usually, to achieve high efficiency at the Gaussian case, we require that $\psi(x) \approx x$ for x small. Consistency of the estimate requires that $E\{\psi(n_j)\} = 0$. Hence, for symmetric noise densities, ψ is usually odd symmetric. We next consider some specific choices of the penalty function ρ and the corresponding derivative ψ.

1) Linear Decorrelating Detector

The linear decorrelating detector, which is simply the LS estimator, corresponds to having the penalty function and its derivative, respectively, as

$$\rho_{\text{LS}}(x) = \frac{x^2}{2\gamma} \tag{21}$$

$$\psi_{\text{LS}}(x) = \frac{x}{\gamma} \tag{22}$$

where γ is any positive constant. Notice that the linear decorrelating detector is scale invariant.

2) Maximum Likelihood Decorrelating Detector

Assume that the i.i.d. noise samples have a pdf f. Then, the likelihood function of the received signal $\underline{r}$ under the true parameters $\underline{\theta}$ is given by

$$\begin{aligned} L_\theta(\underline{r}; f) &\triangleq -\log \prod_{j=1}^{N} f \left(r_j - \sum_{k=1}^{K} s_j^k \theta_k \right) \\ &= -\sum_{j=1}^{N} \log f \left(r_j - \sum_{k=1}^{K} s_j^k \theta_k \right). \end{aligned} \tag{23}$$

Therefore, the maximum likelihood decorrelating detector in non-Gaussian noise with pdf f [in the sense that it is a maximum likelihood estimate of the product of the amplitude and data bit $\theta_k \triangleq A_k b_k$] is given by the M-estimator with penalty function and its derivative, respectively, as

$$\rho_{\text{ML}}(x) = -\log f(x) \tag{24}$$

$$\psi_{\text{ML}}(x) = -\frac{f'(x)}{f(x)}. \tag{25}$$

3) Minimax Decorrelating Detector

We next consider a robust decorrelating detector in a minimax sense based on Huber's minimax M-estimator [13]. Huber considered the robust location estimator problem. Suppose we have one-dimensional (1-D) i.i.d. observations $X_1, \ldots, X_n$. The observations belong to some sample space $\mathcal{X}$, which is a subset of the real line $\mathcal{R}$. A parametric model consists of a family of probability distributions F_θ on the sample space, where the unknown parameter θ belongs to some parameter space Θ. When estimating location in the model $\mathcal{X} = \mathcal{R}$, $\Theta = \mathcal{R}$, $F_\theta(x) = F(x - \theta)$, the M-estimator is determined by a ψ-function of the type $\psi(x, \theta) = \psi(x - \theta)$, i.e., the M-estimate of the location parameter θ is given by the solution to the equation

$$\sum_{i=1}^{n} \psi(x_i - \theta) = 0. \tag{26}$$

Assuming that the sample distribution belongs to the set of ϵ-*contaminated Gaussian models* given by

$$\begin{aligned} \mathcal{P}_\epsilon \triangleq \{ &(1-\epsilon)\mathcal{N}(0, \nu^2) + \epsilon H; \\ &H \text{ is a symmetric distribution} \} \end{aligned} \tag{27}$$

where $0 < \epsilon < 1$ is fixed, and ν^2 is the variance of the nominal Gaussian distribution. It can be shown that, within mild regularity, the asymptotic variance of an M-estimator of the location θ defined by (26) at a distribution $F \in \mathcal{P}_\epsilon$ is given by [13]

$$V(\psi; F) = \frac{\int \psi^2 \, dF}{\left(\int \psi' \, dF \right)^2}. \tag{28}$$

Huber's idea was to minimize the maximal asymptotic variance over $\mathcal{P}_\epsilon$, that is, to find an M-estimator ψ_0 that satisfies

$$\sup_{F \in \mathcal{P}_\epsilon} V(\psi_0; F) = \inf_\psi \sup_{F \in \mathcal{P}_\epsilon} V(\psi; F). \tag{29}$$

This is achieved by finding the least favorable distribution F_0, i.e., the distribution that minimizes the Fisher information

$$I(F) = \int \left(\frac{F''}{F'} \right)^2 dF \tag{30}$$

over all $F \in \mathcal{P}_\epsilon$. Then, $\psi_0 = -\frac{F''}{F'}$ is the maximum likelihood estimator for this least favorable distribution. Huber showed that the Fisher information is minimized by [i.e., the least favorable pdf in $\mathcal{P}_\epsilon$]

$$f_0(x) = \begin{cases} \frac{1-\epsilon}{\sqrt{2\pi}\nu} \exp\left(\frac{x^2}{2\nu^2} \right), & \text{for } |x| \le k\nu^2 \\ \frac{1-\epsilon}{\sqrt{2\pi}\nu} \exp\left(\frac{k^2\nu^2}{2} - k|x| \right), & \text{for } |x| > k\nu^2 \end{cases} \tag{31}$$

where k, ϵ, and ν are connected through

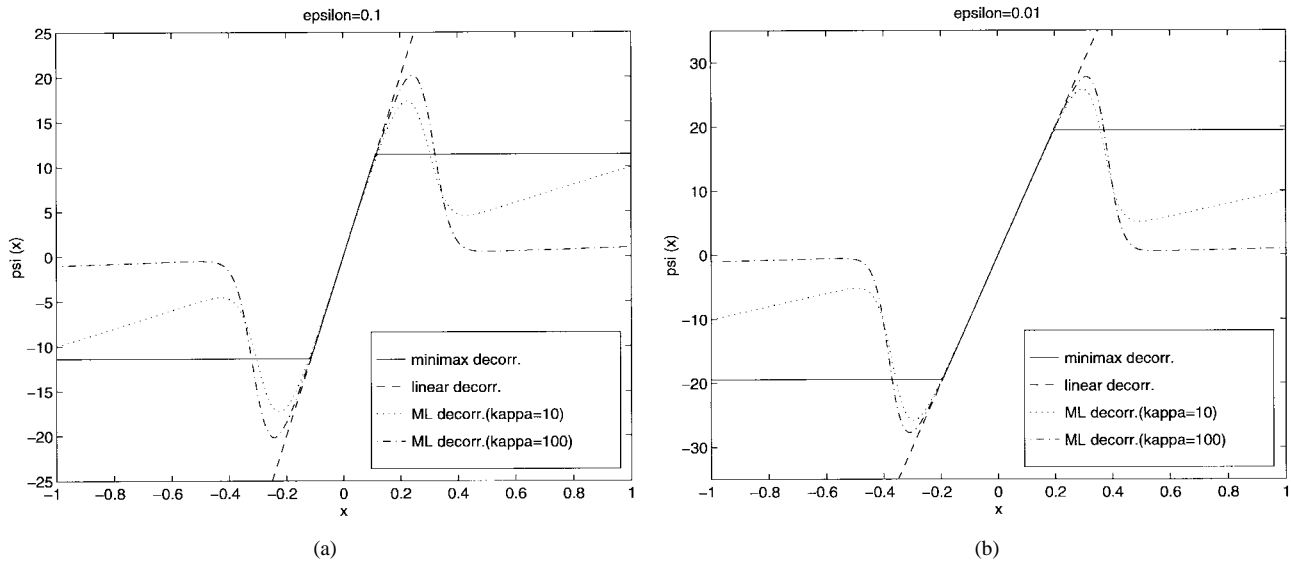$$\frac{\phi(k\nu)}{k\nu} - Q(k\nu) = \frac{\epsilon}{2(1-\epsilon)} \tag{32}$$

Fig. 1. ψ functions for the linear decorrelator, the maximum likelihood decorrelator, and the minimax decorrelator, under the Gaussian mixture noise model. The variance of the nominal Gaussian distribution is $\nu^2 = 0.01$. (a) $\epsilon = 0.1$. The cutoff point for the Huber estimator is obtained by solving (32), resulting in $k = 11.40$. (b) $\epsilon = 0.01$, $k = 19.45$.

where $\phi(x) \triangleq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, and $Q(t) \triangleq \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-\frac{x^2}{2}} dx$. The corresponding minimax M-estimator is then determined by the Huber penalty function and its derivative, which is given, respectively, by

$$\rho_H(x) = \begin{cases} \frac{x^2}{2\nu^2}, & \text{for } |x| \le k\nu^2 \\ \frac{k^2\nu^2}{2} - k|x|, & \text{for } |x| > k\nu^2 \end{cases} \tag{33}$$

and

$$\psi_H(x) = \begin{cases} \frac{x}{\nu^2}, & \text{for } |x| \le k\nu^2 \\ k\,\text{sgn}(x), & \text{for } |x| > k\nu^2. \end{cases} \tag{34}$$

The minimax robust decorrelating detector is obtained by substituting ρ_H and ψ_H into (18) and (19).

Assuming that the noise distribution has the ϵ-mixture density (7), in Fig. 1, we plot the ψ functions for the three types of decorrelating detectors discussed above for the cases $\epsilon = 0.1$ and $\epsilon = 0.01$, respectively. Note that for small measurement x, both $\psi_{\text{ML}}(x)$ and $\psi_H(x)$ are essentially linear, and they coincide with $\psi_{\text{LS}}(x)$; for large measurement x, $\psi_{\text{ML}}(x)$ approximates a blanker, whereas $\psi_H(x)$ acts as a clipper. Thus, the action of the nonlinear function ψ in the nonlinear decorrelators defined by (19) relative to the linear decorrelator defined by (13) is clear in this case. Namely, the linear decorrelator incorporates the residual linearly into the signal estimate, whereas the nonlinear decorrelators incorporates small residuals linearly, but blank or clip larger residuals that are likely to be the result of noise impulses.

C. Asymptotic Performance of Robust Multiuser Detector

1) Influence Function The influence function (IF) introduced by Hampel [8], [13] is an important tool used to study robust estimators. It measures the influence of a vanishingly small contamination of the underlying distribution on the estimator. It is assumed that the estimator can be defined

as a functional T operating on the empirical distribution function F_n, $T = T(F_n)$ and that the estimator is consistent as $n \to \infty$, i.e., $T(F) = \lim_{n\to\infty} T(F_n)$, where F is the underlying distribution. It is defined as

$$\text{IF}(x; T, F) \triangleq \lim_{t\to 0} \frac{T((1-t)F + t\Delta_x) - T(F)}{t} \tag{35}$$

where Δ_x is the distribution that puts a unit mass at x. Roughly speaking, the influence function $\text{IF}(x; T, F)$ is the first derivative of the statistic T at an underlying distribution F and at the coordinate x. We next compute the influence function of the nonlinear decorrelating multiuser detectors defined by (19).

Denote the jth row of the matrix $\underline{S}$ by $\underline{\xi}_j^T$, i.e., $\underline{\xi}_j^T \triangleq [s_j^1 \cdots s_j^K]$. Assume that the signature waveforms of all users are random, and let $K(\underline{\xi})$ be the distribution function of $\underline{\xi}_j$ with density $k(\underline{\xi})$. Assume further that the noise distribution has density f. Denote the joint distribution of the received signal r_j and the chip samples of the K users $\underline{\xi}_j$ under the true parameter $\underline{\theta}$ by $G_{\underline{\theta}}(r, \underline{\xi})$ with density

$$g_{\underline{\theta}}(r, \underline{\xi}) = f(r - \underline{\xi}^T\underline{\theta})k(\underline{\xi}). \tag{36}$$

If G_n is the empirical distribution function generated by the signal samples $\{r_j, \underline{\xi}_j\}_{j=1}^n$, then the solution $\hat{\underline{\theta}}_n$ to (19) can also be written as $\hat{\underline{\theta}}(G_n)$, where $\hat{\underline{\theta}}$ is the K-dimensional functional determined by

$$\int \psi(r - \underline{\xi}^T\hat{\underline{\theta}}(G))\underline{\xi}\, dG(r, \underline{\xi}) = \underline{0}_K \tag{37}$$

for all distributions G for which the integral is defined. Let the distribution be $G = (1-t)G_{\underline{\theta}} + t\Delta_{r,\underline{\xi}}$. Substituting this into (37), differentiating with respect to t, and evaluating it

at $t = 0$, we get

$$
\begin{aligned}
\underline{0}_K &= \int \psi(r - \underline{\xi}^T \hat{\underline{\theta}}(G_{\underline{\theta}})) \underline{\xi}\, d(\Delta_{r,\xi} - G_{\underline{\theta}}) \\
&\quad - \int \psi'(r - \underline{\xi}^T \hat{\underline{\theta}}(G_{\underline{\theta}})) \underline{\xi}\, \underline{\xi}^T f(r - \underline{\xi}^T \underline{\theta}) \\
&\quad \times k(\underline{\xi})\, d\underline{\xi}\, dr \cdot \frac{\partial}{\partial t}[\hat{\underline{\theta}}(G)]|_{t=0} \\
&= \psi(r - \underline{\xi}^T \hat{\underline{\theta}}(G_{\underline{\theta}})) \underline{\xi} - \int \psi(r - \underline{\xi}^T \hat{\underline{\theta}}(G_{\underline{\theta}})) \underline{\xi}\, dG_{\underline{\theta}} \\
&\quad - \int \psi'(r - \underline{\xi}^T \hat{\underline{\theta}}(G_{\underline{\theta}})) f(r - \underline{\xi}^T \underline{\theta}) \underline{\xi}\, \underline{\xi}^T \\
&\quad \times k(\underline{\xi})\, d\underline{\xi}\, dr \cdot \mathrm{IF}(r, \underline{\xi}; \psi, G_{\underline{\theta}})
\end{aligned}
\tag{38}
$$

where, by definition

$$
\begin{aligned}
\frac{\partial}{\partial t}[\hat{\underline{\theta}}(G)]|_{t=0} &\triangleq \lim_{t \to 0} \frac{\hat{\underline{\theta}}((1-t)G_{\underline{\theta}} + t\Delta_{r,\xi}) - \hat{\underline{\theta}}(G_{\underline{\theta}})}{t} \\
&\triangleq \mathrm{IF}(r, \underline{\xi}; \hat{\underline{\theta}}, G_{\underline{\theta}}).
\end{aligned}
\tag{39}
$$

Note that the second term on the right-hand side of (38) $\int \psi(r - \underline{\xi}^T \hat{\underline{\theta}}(G_{\underline{\theta}})) \underline{\xi}\, dG_{\underline{\theta}}$ is $\underline{0}_K$ by (37). Now, assume that the functional $\hat{\underline{\theta}}$ is Fisher consistent [8], i.e., $\hat{\underline{\theta}}(G_{\underline{\theta}}) = \underline{\theta}$, which means that at the model, the estimator $\{\hat{\underline{\theta}}_n : n \geq 1\}$ asymptotically measures the right quantity. We proceed with (38) to obtain

$$
\begin{aligned}
\underline{0}_K &= \psi(r - \underline{\xi}^T \underline{\theta}) \underline{\xi} - \int \psi'(r - \underline{\xi}^T \underline{\theta}) f(r - \underline{\xi}^T \underline{\theta}) \underline{\xi}\, \underline{\xi}^T \\
&\quad \times k(\underline{\xi})\, d\underline{\xi}\, dr \cdot \mathrm{IF}(r, \underline{\xi}; \psi, G_{\underline{\theta}}) \\
&= \psi(r - \underline{\xi}^T \underline{\theta}) \underline{\xi} - \int \psi'(u) f(u)\, du \cdot \underline{R}^* \cdot \mathrm{IF}(r, \underline{\xi}; \hat{\underline{\theta}}, G_{\underline{\theta}})
\end{aligned}
\tag{40}
$$

where

$$
\underline{R}^* \triangleq \int \underline{\xi}\underline{\xi}^T k(\underline{\xi})\, d\underline{\xi}
\tag{41}
$$

is the cross-correlation matrix of the random infinite-length signature waveforms of the K users. From (40), we obtain the influence function of the nonlinear decorrelating multiuser detectors determined by (19) as

$$
\mathrm{IF}(r, \underline{\xi}; \hat{\underline{\theta}}, G_{\underline{\theta}}) = \frac{\psi(r - \underline{\xi}^T \underline{\theta})}{\int \psi'(u) f(u)\, du} \underline{R}^{*-1} \underline{\xi}.
\tag{42}
$$

2) Asymptotic Probability of Error: Under certain regularity conditions, the M-estimators defined by (18) or (19) are consistent and asymptotically normal [8], i.e.,

$$
\sqrt{n}(\hat{\underline{\theta}}_n - \underline{\theta}) \sim \mathcal{N}(\underline{0}, \underline{V}(\hat{\underline{\theta}}, G_{\underline{\theta}})), \quad \text{as } n \to \infty
\tag{43}
$$

where the asymptotic covariance matrix is given by

$$
\begin{aligned}
\underline{V}(\hat{\underline{\theta}}, G_{\underline{\theta}}) &= \int \mathrm{IF}(r, \underline{\xi}; \hat{\underline{\theta}}, G_{\underline{\theta}}) \cdot \mathrm{IF}(r, \underline{\xi}; \hat{\underline{\theta}}, G_{\underline{\theta}})^T\, dG_{\underline{\theta}}(r, \underline{\xi}) \\
&= \frac{\int \psi^2(u) f(u)\, du}{\left[\int \psi'(u) f(u)\, du\right]^2} \cdot \underline{R}^{*-1}.
\end{aligned}
\tag{44}
$$

We can also compute the Fisher information matrix for the parameters $\underline{\theta}$ at the underlying noise distribution. Define the likelihood score vector as

$$
\begin{aligned}
\underline{s}(r, \underline{\xi}; \underline{\theta}) &\triangleq \frac{\partial}{\partial \underline{\theta}} \ln g_{\underline{\theta}}(r, \underline{\xi}) \\
&= \frac{\partial}{\partial \underline{\theta}} \ln f(r - \underline{\xi}^T \underline{\theta}) \\
&= \frac{f'(r - \underline{\xi}^T \underline{\theta})}{f(r - \underline{\xi}^T \underline{\theta})} \cdot \underline{\xi}.
\end{aligned}
\tag{45}
$$

The Fisher information matrix is then given by

$$
\begin{aligned}
\underline{J}(\underline{\theta}) &\triangleq \int \underline{s}(r, \underline{\xi}; \underline{\theta}) \underline{s}(r, \underline{\xi}; \underline{\theta})^T g_{\underline{\theta}}(r, \underline{\xi})\, dr\, d\underline{\xi} \\
&= \int \frac{f'(u)^2}{f(u)}\, du \cdot \underline{R}^*.
\end{aligned}
\tag{46}
$$

It is known that the maximum likelihood estimate based on i.i.d. samples is asymptotically unbiased and that the asymptotic covariance matrix is $\underline{J}(\underline{\theta})^{-1}$ [21]. As discussed earlier, the maximum likelihood estimate of $\underline{\theta}$ corresponds to having $\psi(x) = -\frac{f'(x)}{f(x)}$. Hence, we can deduce that the asymptotic covariance matrix $\underline{V}(\hat{\underline{\theta}}, G_{\underline{\theta}}) = \underline{J}(\underline{\theta})^{-1}$ when $\psi(x) = -\frac{f'(x)}{f(x)}$. To verify this, substituting $\psi(x) = -\frac{f'(x)}{f(x)}$ into (44), we obtain

$$
\begin{aligned}
\underline{V}(\hat{\underline{\theta}}, G_{\underline{\theta}}) &= \frac{\int \frac{f'(u)^2}{f(u)}\, du}{\left[\int \frac{f'(u)^2}{f(u)}\, du - \int f''(u)\, du\right]^2} \cdot \underline{R}^{*-1} \\
&= \left[\int \frac{f'(u)^2}{f(u)}\, du\right]^{-1} \cdot \underline{R}^{*-1} \\
&= \underline{J}(\underline{\theta})^{-1}
\end{aligned}
\tag{47}
$$

where we have assumed that under regularity conditions, the differentiation and integration can be interchanged so that $\int f''(u)\, du = (\int f(u)\, du)'' = 1'' = 0$.

Next, we consider the asymptotic probability of error for the class of decorrelating detectors defined by (19) for large processing gain $N \to \infty$. Using the asymptotic normality condition (43), $\hat{\underline{\theta}}_N \sim \mathcal{N}(\underline{\theta}, \underline{V})$. The asymptotic probability of error for the kth user is then given by

$$
\begin{aligned}
P_e^k &\triangleq \Pr(\hat{\theta}_k < 0 \mid \theta_k > 0) \\
&\tilde{Q}\left(\frac{A_k}{\upsilon\sqrt{[\underline{R}^{*-1}]_{kk}}}\right)
\end{aligned}
\tag{48}
$$

where υ is the asymptotic variance given by

$$
\upsilon^2 \triangleq \frac{\int \psi^2(u) f(u)\, du}{\left[\int \psi'(u) f(u)\, du\right]^2}.
\tag{49}
$$

Hence, for the class of M-decorrelators defined by (19), their asymptotic probabilities of detection error are connected through the parameter υ. We next compute υ for the three decorrelating detectors discussed in Section III-B under the Gaussian mixture noise model (7).

1) Linear Decorrelating Detector

The asymptotic variance of the linear decorrelator is given by

$$v_{\mathrm{LS}}^2 = \int u^2 f(u)\, du = \mathrm{Var}(n_j)$$
$$= (1 - \epsilon)\nu^2 + \epsilon\kappa\nu^2. \tag{50}$$

That is, asymptotically, the performance of the linear decorrelating detector is completely determined by the noise variance independent of the noise distribution. However, as will be seen in the next subsection, the noise distribution does affect substantially the finite sample performance of the linear decorrelating detector.

2) Maximum Likelihood Decorrelating Detector

The maximum likelihood decorrelating detector achieves the Fisher information covariance matrix, and we have

$$v_{\mathrm{ML}}^2 = \left[\int \frac{f'(u)^2}{f(u)}\, du\right]^{-1}. \tag{51}$$

Then, (51) gives the minimum achievable v^2. To see this, we use the Cauchy–Schwarz inequality to yield

$$\int \psi(u)^2 f(u)\, du \cdot \int \frac{f'(u)^2}{f(u)}\, du$$
$$\geq \left(\int |\psi(u)f'(u)|\, du\right)^2$$
$$\geq \left(\int \psi(u)f'(u)\, du\right)^2$$
$$= \left(\psi(u)f(u)|_{-\infty}^{+\infty} - \int \psi'(u)f(u)\, du\right)^2$$
$$= \left(\int \psi'(u)f(u)\, du\right)^2 \tag{52}$$

where the last equality follows from the fact that $\psi(u)f(u) \to 0$, as $|u| \to \infty$. To see this, we use (7) and (25) to obtain

$$f(u)\psi(u) \triangleq -f(u)\frac{f'(u)}{f(u)} = -f'(u)$$
$$= \frac{u}{\nu^2\sqrt{2\pi\nu^2}}\left[(1 - \epsilon)\exp\left(-\frac{u^2}{2\nu^2}\right)\right.$$
$$\left. + \frac{\epsilon}{\kappa\sqrt{\kappa}}\exp\left(-\frac{u^2}{2\kappa\nu^2}\right)\right]$$
$$\to 0, \quad \text{as } |u| \to \infty. \tag{53}$$

Hence, it follows from (52) that

$$\frac{\int \psi^2(u)f(u)\, du}{\left[\int \psi'(u)f(u)\, du\right]^2} \geq \left[\int \frac{f'(u)^2}{f(u)}\, du\right]^{-1}. \tag{54}$$

3) Minimax Decorrelating Detector

For the minimax decorrelating detector, we have $\psi(x) = \frac{x}{\nu^2} \cdot 1_{\{|x| \leq k\nu^2\}} + k\,\mathrm{sgn}(x) \cdot 1_{\{|x| > k\nu^2\}}$, and $\psi'(x) = \frac{1}{\nu^2} \cdot 1_{\{|x| \leq k\nu^2\}} + k\delta_{-k\nu^2} - k\delta_{k\nu^2}$, where $1_\Omega(x)$ is the indicator

function of the set Ω, and δ_x denotes the Dirac delta function at x. After some algebra, we obtain

$$\int \psi^2(u)f(u)\, du$$
$$= \frac{2}{\nu^2}\left[\frac{1 + (\kappa - 1)\epsilon}{2} + (1 - \epsilon)(k^2\nu^2 - 1)Q(k\nu)\right.$$
$$+ \epsilon(k^2\nu^2 - \kappa)Q\left(\frac{k\nu}{\sqrt{\kappa}}\right) - \frac{(1 - \epsilon)k\nu}{\sqrt{2\pi}}\exp\left(-\frac{k^2\nu^2}{2}\right)$$
$$\left. - \sqrt{\frac{\kappa}{2\pi}}\epsilon k\nu\exp\left(-\frac{k^2\nu^2}{2\kappa}\right)\right] \tag{55}$$
$$\int \psi'(u)f(u)\, du = \frac{2}{\nu^2}\left[\frac{1}{2} - (1 - \epsilon)Q(k\nu) - \epsilon Q\left(\frac{k\nu}{\sqrt{\kappa}}\right)\right]. \tag{56}$$

The asymptotic variance v_H^2 of the minimax decorrelating detector is obtained by substituting (55) and (56) into (49).

In Fig. 2, we plot the asymptotic variance v^2 of the maximum likelihood decorrelator and the minimax robust decorrelator as a function of ϵ and κ, under the Gaussian mixture noise model (7). The total noise variance is kept constant as ϵ and κ vary, i.e., $\sigma^2 \triangleq (1 - \epsilon)\nu^2 + \epsilon\kappa\nu^2 = (0.1)^2$. From the two plots, we see that the two nonlinear detectors have very similar asymptotic performance. Moreover, in this case, the asymptotic variance v^2 is a decreasing function of either ϵ or κ when one of them is fixed. The asymptotic variance of both nonlinear decorrelators are strictly less than that of the linear decorrelator, which corresponds to a plane at $v^2 = \sigma^2 = (0.1)^2$. In Fig. 3, we plot the asymptotic variance v^2 of the three decorrelating detectors as a function of κ with fixed ϵ, and in Fig. 4, we plot the asymptotic variance v^2 of the three decorrelating detectors as a function of ϵ with fixed κ. As before, the total variance of the noise for both figures is fixed at $\sigma^2 = (0.1)^2$. From these figures, we see that the asymptotic variance of the minimax decorrelator is very close to that of the maximum likelihood decorrelator for the cases of small contamination, e.g., $\epsilon \leq 0.1$, whereas both of the detectors can outperform the linear detector by a substantial margin.

D. Computational Issues

In this section, we discuss computational procedures for obtaining the output of the nonlinear decorrelating multiuser detectors, i.e., the solution to (19). Assume that the penalty function $\rho(x)$ in (18) has a bounded second-order derivative, i.e., $|\rho''(x)| = |\psi'(x)| \leq \mu$, for some $\mu > 0$. Then, (19) can be solved iteratively by the following modified residual method [13]. Let $\underline{\theta}^l$ be the estimate at the lth step; then, it is updated according to

$$\underline{z}^l \triangleq \psi(\underline{r} - \underline{S}^T\underline{\theta}^l) \tag{57}$$
$$\underline{\theta}^{l+1} = \underline{\theta}^l + \frac{1}{\mu}(\underline{S}^T\underline{S})^{-1}\underline{S}^T\underline{z}^l \tag{58}$$

where μ is a step-size parameter. Denote the cost function in (18) by $C(\underline{\theta}) \triangleq \sum_{j=1}^N \rho(r_j - \underline{\xi}_j^T\underline{\theta})$. We have the following result regarding the convergence behavior of the above iterative procedure.
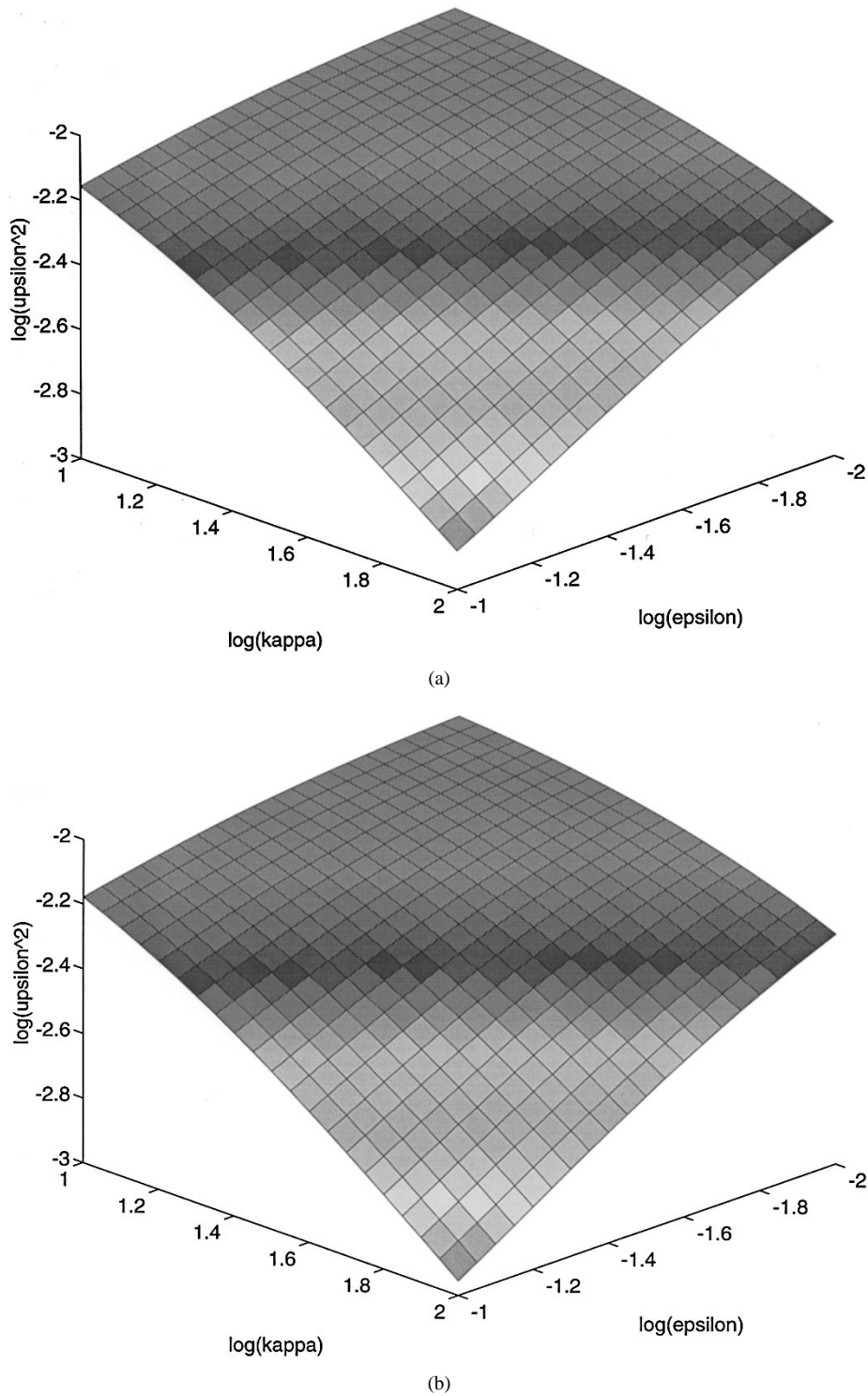
(a)

(b)

Fig. 2. Asymptotic variance v^2 of (a) the minimax robust decorrelating detector and (b) the maximum likelihood decorrelating detector as a function of ϵ and κ under the Gaussian mixture noise model with variance of the noise fixed at $\sigma^2 \triangleq (1 - \epsilon)v^2 + \epsilon \kappa v^2 = (0.1)^2$.

Proposition 1: If $|\psi'(x)| \leq \mu$, then the iterative estimates defined by (57) and (58) satisfy

$$C(\underline{\theta}^l) - C(\underline{\theta}^{l+1}) \geq \frac{\mu}{2}(\underline{\theta}^l - \underline{\theta}^{l+1})^T \underline{R}(\underline{\theta}^l - \underline{\theta}^{l+1})$$

$$= \frac{1}{2\mu} \underline{z}(\underline{\theta}^l)^T \underline{S} \, \underline{R}^{-1} \underline{S}^T \underline{z}(\underline{\theta}^l) \qquad (59)$$

where $\underline{R} \triangleq \underline{S}^T \underline{S}$ is assumed to be positive definite, and

$\underline{z}(\underline{\theta}) \triangleq \psi(\underline{r} - \underline{S}^T \underline{\theta})$. Furthermore, if $\rho(x)$ is convex and bounded from below, then $\underline{\theta}^l \to \underline{\theta}^*$ as $l \to \infty$, where $\underline{\theta}^*$ is the unique minimum point of the cost function $C(\underline{\theta})$, i.e., the unique solution to (19). ∎

The proof of this result is found in Appendix A. Notice that for the minimax robust decorrelating detector, the Huber penalty function $\rho_H(x)$ does not have second-order derivatives at the two "corner" points, i.e., $x = \pm k v^2$. In principle, this
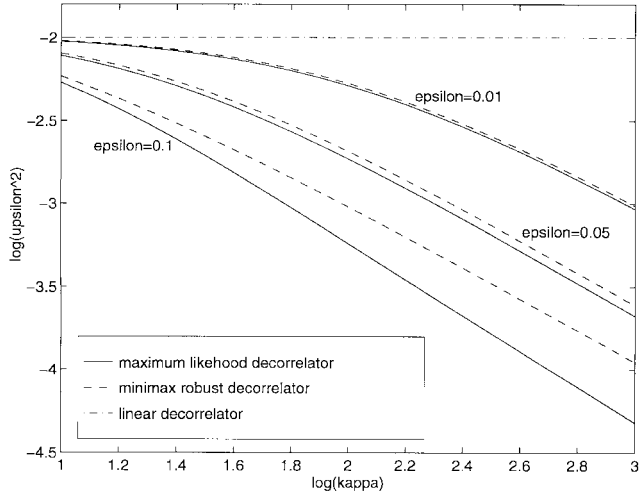
Fig. 3. Asymptotic variance v^2 of the three decorrelating detectors as a function of κ with fixed parameter ϵ. The variance of the noise is fixed $\sigma^2 \triangleq (1-\epsilon)\nu^2 + \epsilon\kappa\nu^2 = (0.1)^2$.
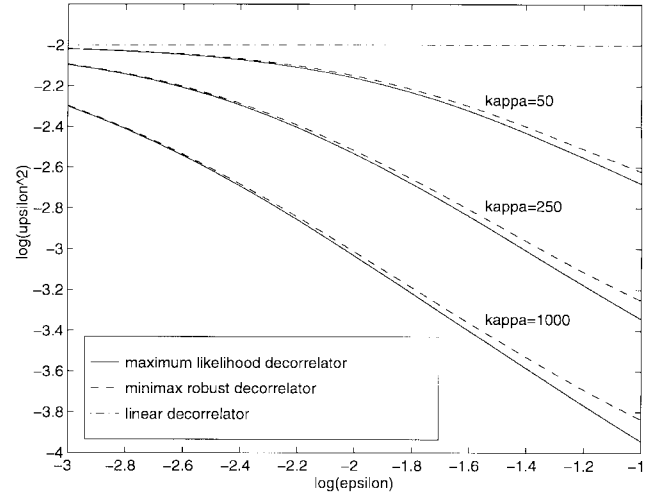

Fig. 4. Asymptotic variance v^2 of the three decorrelating detectors as a function of ϵ with fixed parameter κ. The variance of the noise is fixed at $\sigma^2 \triangleq (1-\epsilon)\nu^2 + \epsilon\kappa\nu^2 = (0.1)^2$.

can be resolved by defining a smoothed version of the Huber penalty function, for example, as (60), shown at the bottom of the page, where η is a small number. The first- and second-order derivatives of this smoothed Huber penalty function are given, respectively, by (61) and (62), shown at the bottom of the page. We can then apply the iterative procedure (57)–(58) using this smoothed penalty function and the step size $\frac{1}{\mu} = \nu^2$. In practice, however, convergence can always be achieved even if the nonsmooth nonlinearity $\psi_H(x)$ is used.

Notice that matrix $\frac{1}{\mu}(\underline{S}^T\underline{S})^{-1}\underline{S}^T$ in (58) can be computed offline, and the major computation involved at each iteration is the product of this $(K \times K)$ matrix with a K-vector $\underline{z}^l$. For the initial estimate $\underline{\theta}^0$, we can take the least squares solution, i.e.,

$$\underline{\theta}^0 = \frac{1}{\mu}(\underline{S}^T\underline{S})^{-1}\underline{S}^T\underline{r}. \tag{63}$$

The iteration is stopped if $\|\underline{\theta}^l - \underline{\theta}^{l-1}\| \le \varepsilon$ for some small number ε. Numerical experiments show that on average, it takes fewer than ten iterations for the algorithm to converge.

The operations of the M-decorrelating multiuser detector are depicted in Fig. 5. It is evident that it is essentially a robust version of the linear decorrelating detector. At each iteration, the residual signal, which is the difference between the received signal $\underline{r}$ and the remodulated signal $\underline{S}\,\underline{\theta}^l$, is passed through the nonlinearity $\psi(\cdot)$. Then, the modified residual $\underline{z}^l$ is passed through the linear decorrelating filter to get the modification on the previous estimate.

E. Simulation Examples

In this section, we provide some simulation examples to demonstrate the performance of the nonlinear M-decorrelating multiuser detectors against multiple-access interference and non-Gaussian additive noise. We consider a synchronous system with $K = 6$ users. The spreading sequence of each user is a shifted version of an m-sequence of length $N = 31$.

We first demonstrate the performance degradation of linear multiuser detectors in non-Gaussian ambient noise. Two

$$\rho_H^s = \begin{cases} \frac{x^2}{2\nu^2}, & \text{if } |x| \le (k-\eta)\nu^2 \\ (k-\eta)x + \eta^2\nu^2\ln\cosh\left(\frac{x-(k-\eta)\nu^2}{\eta\nu^2}\right), & \text{if } x > (k-\eta)\nu^2 \\ -(k-\eta)x + \eta^2\nu^2\ln\cosh\left(\frac{x+(k-\eta)\nu^2}{\eta\nu^2}\right), & \text{if } x < -(k-\eta)\nu^2 \end{cases} \tag{60}$$

$$\psi_H^s \triangleq \rho_H^{s\prime} = \begin{cases} \frac{x}{\nu^2}, & \text{if } |x| \le (k-\eta)\nu^2 \\ k-\eta + \eta\tanh\left(\frac{x-(k-\eta)\nu^2}{\eta\nu^2}\right), & \text{if } x > (k-\eta)\nu^2 \\ -(k-\eta) + \eta\tanh\left(\frac{x+(k-\eta)\nu^2}{\eta\nu^2}\right), & \text{if } x < -(k-\eta)\nu^2 \end{cases} \tag{61}$$

$$\psi_H^{s\prime} = \rho_H^{s\prime\prime} = \begin{cases} \frac{1}{\nu^2}, & \text{if } |x| \le (k-\eta)\nu^2 \\ \frac{1}{\nu^2}\left[1 - \tanh^2\left(\frac{x-(k-\eta)\nu^2}{\eta\nu^2}\right)\right], & \text{if } x > (k-\eta)\nu^2 \\ \frac{1}{\nu^2}\left[1 - \tanh^2\left(\frac{x+(k-\eta)\nu^2}{\eta\nu^2}\right)\right], & \text{if } x < -(k-\eta)\nu^2 \end{cases}$$
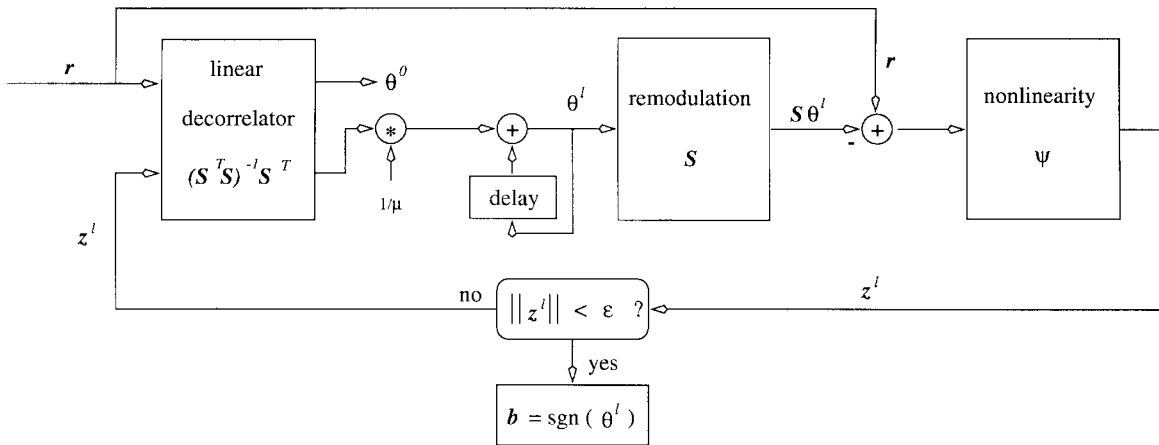
$$\le \frac{1}{\nu^2}. \tag{62}$$

Fig. 5. Diagram of the M-decorrelating multiuser detector, which is a robust version of the linear decorrelating multiuser detector.

popular linear multiuser detectors are the linear decorrelating detector and the linear MMSE detector. The performance of the linear decorrelating detector in several different ϵ-mixture channels is depicted in Fig. 6. In this figure, we plot the error probability versus the signal-to-noise ratio (SNR) corresponding to the first user under perfect power control. The performance of the linear MMSE multiuser detector is indistinguishable in this case from that of the decorrelating detector. It is seen that the impulsive character of the channel ambient noise can substantially degrade the performance of both linear multiuser detectors. Similar situations have been observed for the conventional matched filter receiver in [1]. In [23], it is observed that non-Gaussian-based optimal detection can achieve significant performance gain (more than 10 dB in some cases) over Gaussian-based optimal detection in multiple-access channels when the ambient noise is impulsive. However, this gain is obtained with a significant penalty on complexity. The robust techniques developed in this paper constitute some low-complexity multiuser detectors that account for non-Gaussian ambient noise. We next demonstrate the performance gain afforded by this non-Gaussian-based suboptimal detection technique over its Gaussian-based counterpart, i.e., linear multiuser detectors.

The next example demonstrates the performance gains achieved by the minimax robust decorrelating detector over the linear decorrelator in impulsive noise. The noise distribution parameters are $\epsilon = 0.01$ and $\kappa = 100$. The bit error rate versus SNR for the two detectors is plotted in Fig. 7. The performance of an "approximate" minimax decorrelating detector is also shown in this figure, in which the nonlinearity $\psi(\cdot)$ is taken as

$$\psi(x) = \begin{cases} \frac{x}{\sigma^2}, & \text{for } |x| \leq k\sigma^2 \\ k\,\text{sgn}(x), & \text{for } |x| > k\sigma^2 \end{cases} \qquad (64)$$

where the parameter k is taken as $k = \frac{1.5}{\sigma}$, and the step size parameter μ in the modified residual method (58) is set as $\mu = \sigma^2$. The reason for studying such an approximate robust detector is that in practice, it is unlikely that the exact parameters ϵ and ν in the noise model (7) are known to the receiver. However, the total noise variance σ^2 can be estimated from the received signal [as will be discussed in the next section]. Hence, if we could set some simple rule for
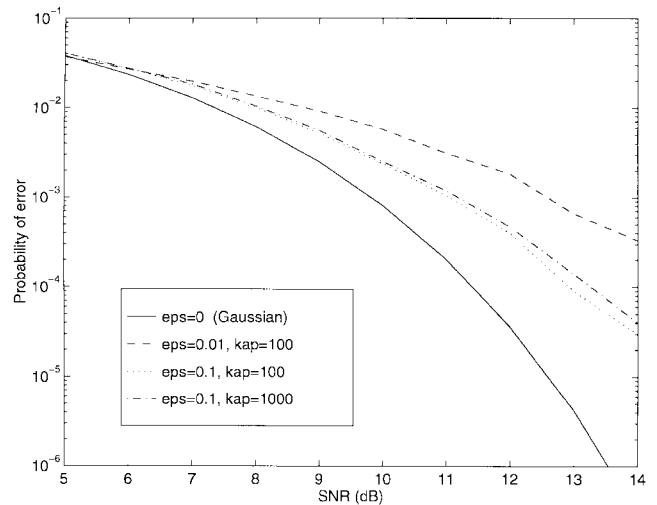

Fig. 6. Probability of error versus signal-to-noise ratio (SNR) of the linear decorrelating detector for user 1 in a synchronous CDMA channel with Gaussian and ϵ-mixture ambient noise. $N = 31$, $K = 6$. Perfect power control is assumed.
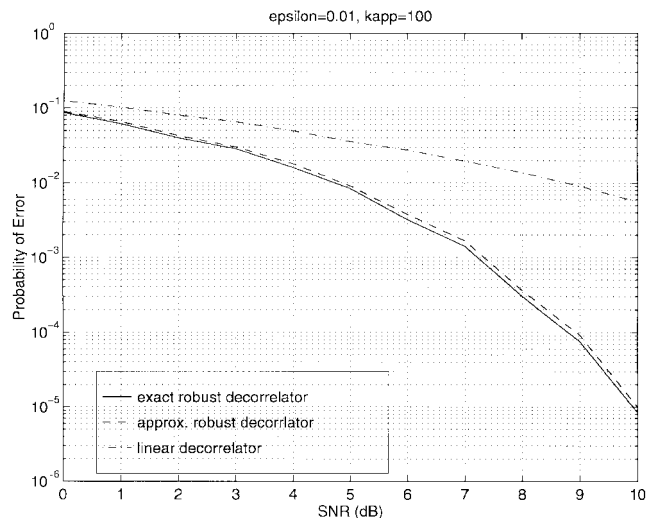

Fig. 7. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the exact minimax decorrelating detector, an approximate minimax decorrelating detector and the linear decorrelating detector, in a synchronous CDMA channel with impulsive noise. $N = 31$, $K = 6$, $\epsilon = 0.01$, $\kappa = 100$. The powers of the interferers are 10 dB above the power of user 1.
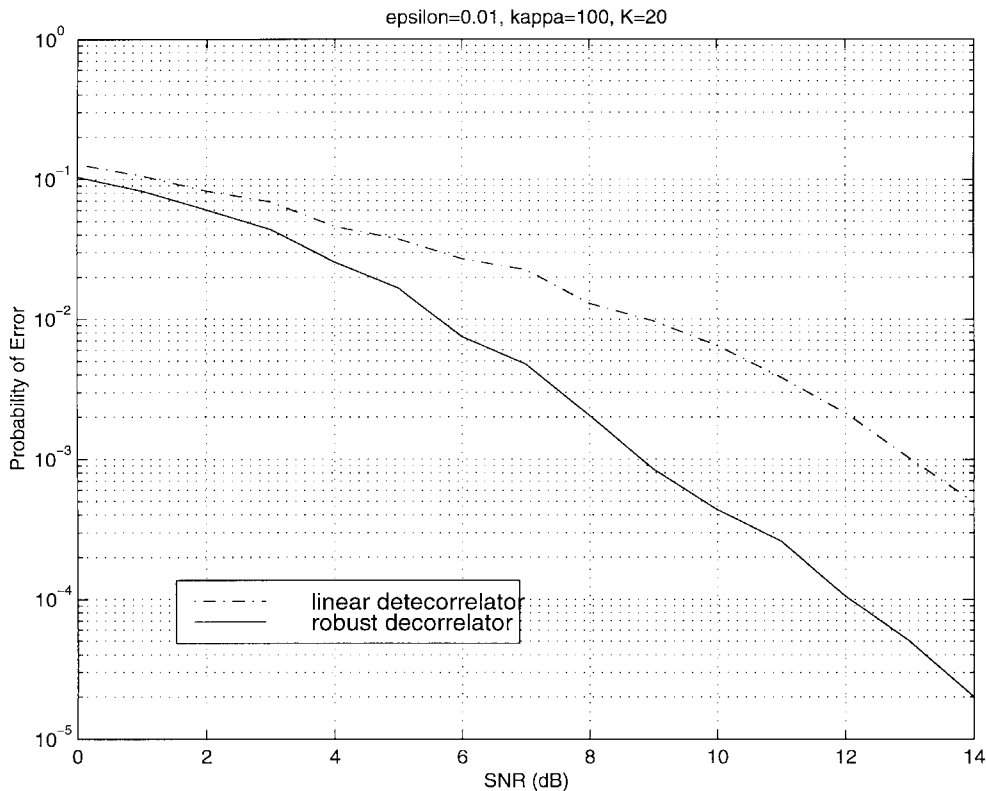
Fig. 8. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the approximate minimax decorrelating detector and the linear decorrelating detector in a synchronous CDMA channel with impulsive noise. $N = 31$, $K = 20$, $\epsilon = 0.01$, $\kappa = 100$. All users have the same power.

choosing the nonlinearity $\psi(\cdot)$ and μ, then this approximate robust detector is much easier to implement in practice than the exact one. It is seen from Fig. 7 that the robust decorrelating multiuser detector offers significant performance gains over the linear decorrelating detector. Moreover, this performance gain increases as the SNR increases. Another important observation is that the performance of the robust multiuser detector is insensitive to the parameters ϵ and κ in the noise model, which is evidenced by the fact that the performance of the approximate robust detector is very close to that of the exact robust detector. We next consider a synchronous system with 20 users $(K = 20)$. The spreading sequence of each user is still a shifted version of the m-sequence of length $N = 31$. The performance of the approximate robust decorrelator and that of the linear decorrelator is shown in Fig. 8. Again, it is seen that the robust detector offers substantial performance gains over the linear detector.

In the third example, we consider the performance of the approximate robust decorrelator in Gaussian noise. Bit error rate curves for the robust decorrelator and the linear decorrelator in a six-user system $(K = 6)$ are shown in Fig. 9. It is seen that there is only a very slight performance degradation by the robust decorrelator in Gassian channels, relative to the linear decorrelator, which is the optimal decorrelating detector in Gaussian noise. Comparing the bit-error curves of the robust decorrelator in Figs. 7 and 9, it is seen that the robust detector performs better in impulsive noise than in Gaussian noise with the same noise variance. This is because in an impulsive environment, a portion of the total noise
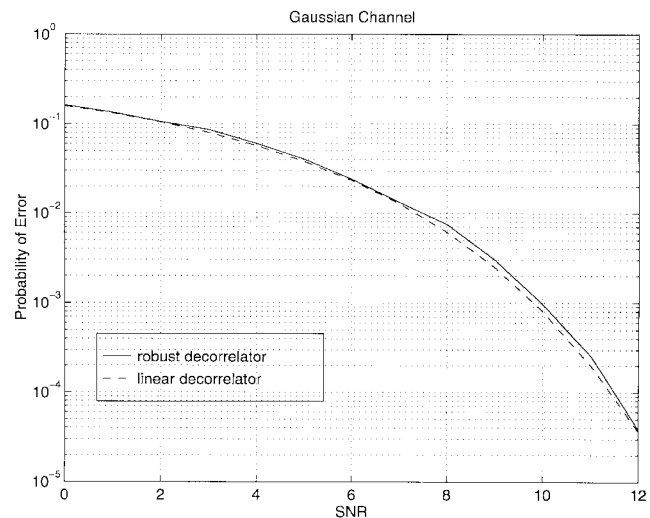

Fig. 9. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the robust decorrelating detector and the linear decorrelating detector in a synchronous CDMA channel with Gaussian noise. $N = 31$, $K = 6$. The powers of the interferers are 10 dB above the power of user 1.

variance is due to impulses, which have large amplitudes. Such impulses are clipped by the nonlinearity in the detector. Therefore, the effective noise variance at the output of the robust detector is smaller than the input total noise variance. In fact, the asymptotic performance gain of the robust detector in impulsive noise over Gaussian noise is quantified by the asymptotic variance v^2 in (49) [cf., Figs. 2–4].

In summary, we have seen that the performance of the linear decorrelating detector degrades substantially when the distribution of the ambient channel noise deviates even slightly from Gaussian. By using the robust decorrelating detector, such performance loss is prevented, and it offers significant performance gains over the linear detectors, which translates into a channel capacity increase in multiple-access channels. On the other hand, even when the ambient noise distribution is indeed Gaussian, the robust detector incurs only negligible performance loss relative to the linear detectors.

IV. ROBUST BLIND MULTIUSER DETECTION

The robust multiuser detection procedure developed in the previous section offers substantial performance gain over the linear multiuser detectors when the ambient channel noise becomes impulsive. In the previous sections, we have assumed that the signature waveforms of all users, as well as the distribution of the ambient noise, are known to the receiver in order to implement the robust multiuser detectors. The requirement on the knowledge of the exact noise distribution can be alleviated since, as demonstrated in Section III-E, little performance degradation is incurred if we simply adopt in the robust multiuser detector a nonlinearity ψ, which depends only on the total noise variance but not the shape of the distribution. In this section, we develop a technique to alleviate the requirement on the knowledge of all users' signatures. One remarkable feature of the linear multiuser detectors is that there exist *blind* techniques that can be used to adapt these detectors, which allow us to use a linear multiuser detector for a given user with no knowledge beyond that required for implementation of the conventional matched-filter detector for that user. In this section, we show that the robust multiuser detector can also be implemented blindly, i.e., with the prior knowledge of the signature waveform of only the user of interest.

There are two major approaches to blind adaptive multiuser detection. In the first approach, the received signal is passed through a linear filter, which is chosen to minimize, within a constraint, the mean-square value of its output [10]. Adaptation algorithms such as least-mean-squares (LMS) or recursive-least-squares (RLS) can be applied for updating the filter weights. Ideally, the adaptation will lead the filter converge to the linear MMSE multiuser detector, irrespective of the noise distribution. [In practice, the impulsiveness of the noise will slow down the convergence.] Therefore, this approach can not adapt the robust multiuser detector.

Another approach to blind multiuser detection is the subspace-based method proposed in [27] through which both the linear decorrelating detector and the linear MMSE detector can be obtained blindly. As will be discussed in this section, this approach is more fruitful in leading to a blind adaptive robust multiuser detection method.

A. Subspace-Based Blind Robust Multiuser Detection

Denote $\underline{S} \triangleq [\underline{s}_1 \ \underline{s}_2 \ \cdots \ \underline{s}_K]$ and $\underline{A} \triangleq \mathrm{diag}(A_1^2, \ldots, A_K^2)$. Since the data bits of K users $\{b_k(i)\}$ are independent ± 1 random variables, and they are independent of the noise

samples $\{\underline{n}(i)\}$, the autocorrelation matrix of the received signal $\underline{r}(i)$ in (6) is then given by

$$\underline{C} \triangleq E\{\underline{r}(i)\underline{r}(i)^T\} = \sum_{k=1}^{K} A_k^2 \underline{s}_k \underline{s}_k^T + \sigma^2 \underline{I}_N$$
$$= \underline{S}\,\underline{A}\,\underline{S}^T + \sigma^2 \underline{I}_N. \qquad (65)$$

By performing an eigendecomposition of the matrix $\underline{C}$, we can write

$$\underline{C} = \underline{U}\,\underline{\Lambda}\,\underline{U}^T = [\underline{U}_s \underline{U}_n] \begin{bmatrix} \underline{\Lambda}_s & \\ & \underline{\Lambda}_n \end{bmatrix} \begin{bmatrix} \underline{U}_s^T \\ \underline{U}_n^T \end{bmatrix} \qquad (66)$$

where $\underline{U} = [\underline{U}_s \underline{U}_n]$, $\underline{\Lambda} = \mathrm{diag}(\underline{\Lambda}_s, \underline{\Lambda}_n)$, $\underline{\Lambda}_s = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$ contains the K largest eigenvalues of $\underline{C}$ in descending order, $\underline{U}_s = [\underline{u}_1 \ \cdots \ \underline{u}_K]$ contains the corresponding orthonormal eigenvectors $\underline{\Lambda}_n = \sigma^2 \underline{I}_{N-K}$, and $\underline{U}_n = [\underline{u}_{K+1} \ \cdots \ \underline{u}_N]$ contains the $(N - K)$ orthonormal eigenvectors that correspond to the eigenvalue σ^2. It is easy to see that $\mathrm{range}(\underline{S}) = \mathrm{range}(\underline{U}_s)$. The range space of $\underline{U}_s$ is called the *signal subspace*, and its orthogonal complement, called the *noise subspace*, is spanned by $\underline{U}_n$. The following result is instrumental to developing the subspace-based blind robust multiuser detector.

Proposition 2: Given the eigendecomposition (66) of the autocorrelation matrix $\underline{C}$, suppose that

$$\sum_{k=1}^{K} \theta_k \underline{s}_k = \sum_{j=1}^{K} \zeta_j \underline{u}_j, \quad \theta_k \in \mathcal{R}, \quad \zeta_j \in \mathcal{R}. \qquad (67)$$

Then, we have

$$\theta_k = \alpha_k \sum_{j=1}^{K} \frac{\underline{u}_j^T \underline{s}_k}{\lambda_j - \sigma^2} \zeta_j, \quad k = 1, \ldots, K \qquad (68)$$

where α_k is a positive constant given by $\alpha_k \triangleq A_k^2 = [\sum_{j=1}^{K} \frac{(\underline{u}_j^T \underline{s}_k)^2}{\lambda_j - \sigma^2}]^{-1}$. ∎

The proof of this result is found in Appendix B.

The above result leads to a subspace-based blind robust multiuser detection technique as follows. From the received data $\{\underline{r}(i)\}$, we can estimate the signal subspace components, i.e., $\underline{\hat{\Lambda}}_s = \mathrm{diag}(\hat{\lambda}_1, \ldots, \hat{\lambda}_K)$, and $\underline{\hat{U}}_s = [\underline{\hat{u}}_1 \ \cdots \ \underline{\hat{u}}_K]$. The received signal $\underline{r}$ in (10) can be expressed as

$$\underline{r} = \underline{S}\,\underline{\theta} + \underline{n} = \underline{U}_s \underline{\zeta} + \underline{n} \qquad (69)$$

where $\underline{\zeta} \triangleq [\zeta_1, \ldots, \zeta_K]^T$. Now, instead of robustly estimating the parameters $\underline{\theta}$ using the known signature waveforms $\underline{S}$ of all users, as is done in the previous section, we can robustly estimate the parameters $\underline{\zeta}$ using the estimated signal subspace coordinates $\underline{\hat{U}}_s$. Denote such a robust estimate as $\underline{\hat{\zeta}} \triangleq [\hat{\zeta}_1, \ldots, \hat{\zeta}_K]^T$. Finally, we compute the estimate of the parameter θ_k of the desired user (up to a positive scaling

Fig. 10. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the blind robust detector and the blind linear detector, using batch-SVD, in a synchronous CDMA channel with non-Gaussian noise. $N = 31$, $K = 6$. The powers of all interferers are 10 dB above the power of user 1.
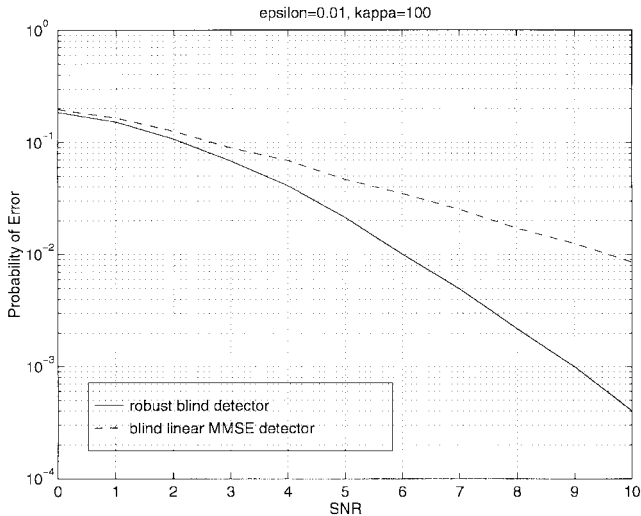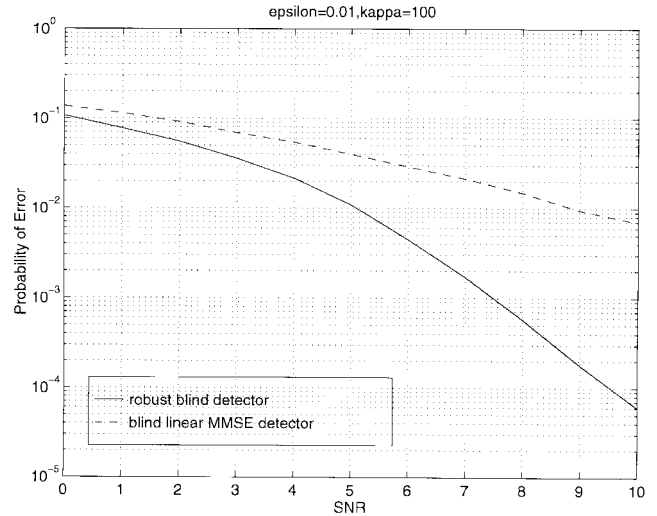

Fig. 11. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the blind robust detector and the blind linear detector, using subspace tracking in a synchronous CDMA channel with non-Gaussian noise. $N = 31$, $K = 6$. The powers of the interferers are 10 dB above the power of user 1.

factor) according to

$$\hat{\theta}_k = \sum_{j=1}^{K} \frac{\hat{\underline{u}}_j^T \underline{s}_k}{\hat{\lambda}_j} \hat{\zeta}_j, \quad k = 1, \dots, K. \tag{70}$$

Note that in the estimator (70), the denomiators are $\hat{\lambda}_j$ instead of $(\hat{\lambda}_j - \hat{\sigma}^2)$, as in (68). This makes the estimator (70) less sensitive to the estimation errors in the estimated quantities $\hat{\underline{u}}_j$, $\hat{\lambda}_j$ and $\hat{\zeta}_j$. Moreover, if the noise $\underline{n}$ has a Gaussian distribution, then (70) corresponds to the output of a linear MMSE multiuser detector [27]. Hence, the above robust multiuser detection procedure can be viewed as a robust version of the linear MMSE multiuser detector. Notice that using this method to demodulate the kth user's data bit $b_k(i)$, the only prior knowledge required at the receiver is the signature waveform of this user and, thus, the term *blind* robust multiuser detector. Notice also that since the columns of $\hat{\underline{U}}_s$ are orthonormal, the modified residual method for updating the robust estimate of $\underline{\zeta}$ is given by

$$\underline{z}^l \triangleq \psi(\underline{r} - \hat{\underline{U}}_s \underline{\zeta}^l) \tag{71}$$

$$\underline{\zeta}^{l+1} = \underline{\zeta}^l + \frac{1}{\mu} \hat{\underline{U}}_s^T \underline{z}^l. \tag{72}$$

There are two approaches to implementing subspace-based blind robust multiuser detectors. In the classical approach, a block of (say, M) received data samples $\{\underline{r}(i)\}_{i=0}^{M-1}$ is collected first. The signal subspace components $\underline{\Lambda}_s$, $\underline{U}_s$, σ are then computed through batch eigenvalue decomposition (ED) of the sample autocorrelation matrix or batch singular value decomposition (SVD) of the data matrix. We then apply robust procedure (70)–(72) to demodulate the kth user's data bits $\{b_k(i)\}_{i=0}^{M-1}$. This approach is computationally expensive, and it incurs a large delay in data demodulation.

A more attractive approach is to use computationally efficient sequential eigendecomposition (subspace tracking) al-

gorithms to adaptively update the estimated signal subspace components. For example, in [27], the projection approximation subspace tracking with deflation (PASTd) algorithm [28] is used to implement subspace-based blind multiuser detectors. This algorithm sequentially updates the K most dominant eigenvalues and the corresponding eigenvectors at each iteration, using a recursive least-squares type of algorithm. The computational complexity of this algorithm at each iteration is $O(NK)$. It is straightforward to use this subspace tracking algorithm for the blind robust multiuser detector; at the ith symbol interval, after receiving the ith data vector $\underline{r}(i)$, the signal subspace components are updated by the PASTd algorithm. Then, the robust procedure in (68), (71), and (72) is invoked to demodulate the kth user's data bit $b_k(i)$. In what follows, we provide some simulation examples to demonstrate the performance of the blind robust multiuser detection algorithms.

B. Simulation Examples

As before, we consider the synchronous system with $K = 6$ users and spreading gain $N = 31$. First, we illustrate the performance of the blind robust multiuser detector based on batch SVD. The size of the data block is $M = 200$. The noise distribution parameters are $\epsilon = 0.01$ and $\kappa = 100$. The bit error rate versus the signal to noise ratio for user 1 is plotted in Fig. 10 for both the blind linear MMSE detector and the blind robust detector. The powers of all interferers are 10 dB above that of user 1. The performance of the blind adaptive robust multiuser detector based on subspace tracking is shown in Fig. 11, where the PASTd algorithm from [28] is used for tracking the signal subspace parameters. The forgetting factor in this algorithm is 0.999. It is seen from these two figures that as in the nonadaptive case, the robust multiuser detector offers significant performance gain over the linear multiuser detector in impulsive noise. Furthermore, in this example, the adaptive version of the blind robust detector based on subspace tracking

outperforms the batch SVD-based approach. [This is because with a forgetting factor 0.999, the effective data window size is $1/(1 - 0.999) = 1000$, whereas the window size in the batch method is $M = 200$, while it has a practical computational complexity and incurs no delay in data demodulation.]

V. EXTENSION TO ASYNCHRONOUS CHANNELS

A. Robust Muliuser Detection in Asynchronous Channels

Thus far, we have assumed that the users are symbol and chip synchronous. In this section, we consider the more general case where the users are symbol and chip asynchronous. We show that in this case, the techniques developed in the previous sections can be applied on a signal with larger dimension, resulting from an enlarged window on the received signal.

Consider the asynchronous signal model (2). Define $m_k \triangleq \lfloor \frac{\tau_k}{T_c} \rfloor$ and $\pi_k \triangleq \frac{\tau_k}{T_c} - m_k$. At the receiver, the received continuous-time signal $r(t)$ is filtered by a chip-matched filter and sampled at the chip rate. Denote the jth chip-matched sample at the ith symbol interval by $r_j(i)$, i.e.,

$$r_j(i) \triangleq \int_{iT+jT_c}^{iT+(j+1)T_c} r(t)\varphi(t - iT - jT_c)\, dt$$
$$= \sum_{k=1}^{K} A_k y_j^k(i) + n_j(i), \quad j = 0, \ldots, N-1 \quad (73)$$

where $n_j(i) \triangleq \int_{iT+jT_c}^{iT+(j+1)T_c} n(t)\varphi(t - iT - jT_c)\, dt$, and

$$y_j^k(i) \triangleq \int_{iT+jT_c}^{iT+(j+1)T_c} \left[\sum_{l=1}^{M-1} b_k(l)s_k(t - lT - \tau_k) \right]$$
$$\times \varphi(t - iT - jT_c)\, dt$$
$$= b_k(i)h_j^k(0) + b_k(i-1)h_j^k(1) \quad (74)$$

where $h_j^k(0)$ and $h_j^k(1)$ are parameters determined by the kth user's signature sequence and delay, which are given later in (77). Define the quantities

$$\underline{r}(i) \triangleq \begin{bmatrix} r_0(i) \\ \vdots \\ r_{N-1}(i) \end{bmatrix}_{N \times 1} \quad \underline{n}(i) \triangleq \begin{bmatrix} n_0(i) \\ \vdots \\ n_{N-1}(i) \end{bmatrix}_{N \times 1}$$

$$\underline{b}(i) \triangleq \begin{bmatrix} b_1(i) \\ \vdots \\ b_K(i) \end{bmatrix}_{K \times 1}.$$

By stacking m successive data samples, we further define the quantities

$$\mathbf{r}_m(i) \triangleq \begin{bmatrix} \underline{r}(i) \\ \vdots \\ \underline{r}(i+m-1) \end{bmatrix}_{Nm \times 1}$$

$$\mathbf{n}_m(i) \triangleq \begin{bmatrix} \underline{n}(i) \\ \vdots \\ \underline{n}(i+m-1) \end{bmatrix}_{Nm \times 1}$$

and

$$\mathbf{b}_m(i) \triangleq \begin{bmatrix} \underline{b}(i-1) \\ \vdots \\ \underline{b}(i+m-1) \end{bmatrix}_{K(m+1) \times 1}.$$

Then, from (73) and (74), we can write

$$\mathbf{r}_m(i) = \mathcal{H}_m \mathcal{A}_m \mathbf{b}_m(i) + \mathbf{n}_m(i) \quad (75)$$

where $\mathcal{A}_m$ is an $[K(m+1) \times K(m+1)]$ diagonal matrix given by $\mathcal{A}_m \triangleq \underline{A} \otimes \underline{I}_{m+1}$ [$\otimes$ denotes Kronecker product], and $\mathcal{H}_m$ is a generalized Sylvester matrix of the form

$$\mathcal{H}_m \triangleq \begin{bmatrix} \underline{H}(1) & \underline{H}(0) & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \underline{H}(1) & \underline{H}(0) \end{bmatrix}_{Nm \times K(m+1)}$$

where the matrix $\underline{H}(0)$ and $\underline{H}(1)$ are $N \times K$ matrices defined as

$$\begin{bmatrix} \underline{H}(0) \\ \underline{H}(1) \end{bmatrix} \triangleq [\underline{h}_1 \quad \cdots \quad \underline{h}_K] \quad (76)$$

with

$$\underline{h}_k \triangleq (1-\pi_k) \begin{bmatrix} \underline{0}_{m_k} \\ \underline{s}_k \\ \underline{0}_{N-m_k} \end{bmatrix} + \pi_k \begin{bmatrix} \underline{0}_{m_k+1} \\ \underline{s}_k \\ \underline{0}_{N-m_k-1} \end{bmatrix}. \quad (77)$$

The parameter m is called the smoothing factor and is chosen such that the Sylvester matrix $\mathcal{H}_m$ has full column rank. It is shown in [15] that under the conditions that a) $\underline{H}(1)$ is of full column rank (K) and b) the Z-transform of $\{\underline{H}(0), \underline{H}(1)\}$, $H(z) \triangleq \underline{H}(0) + \underline{H}(1)z^{-1}$ is of full column rank for all $z \in \mathcal{C}$, then, for all $m \geq \lfloor \frac{K}{N-K} \rfloor + 1$, $\mathcal{H}_m$ has full column rank.

Assuming that $\mathcal{H}_m$ has full column rank, we can then apply the robust technique developed in the previous section for detection of the ith data bits $\underline{b}(i)$ from the data vector $\mathbf{r}_m(i)$. For example, if the signature waveforms of all users as well as the relative delays are known to the receiver, then the Sylvester matrix $\mathcal{H}_m$ can be constructed. Now, similarly to (57) and (58), we can robustly estimate the parameters $\Theta(i) \triangleq \mathcal{A}_m \mathbf{b}_m(i)$ as

$$\Theta^0 = \frac{1}{\mu} (\mathcal{H}_m^T \mathcal{H}_m)^{-1} \mathcal{H}_m^T \mathbf{r}_m \quad (78)$$

$$\mathbf{z}^l \triangleq \psi(\mathbf{r}_m - \mathcal{H}_m^T \Theta^l) \quad (79)$$

$$\Theta^{l+1} = \Theta^l + \frac{1}{\mu} (\mathcal{H}_m^T \mathcal{H}_m)^{-1} \mathcal{H}_m^T \mathbf{z}^l. \quad (80)$$

The ith data bits of K users are then demodulated according to $\underline{b}(i) = \text{sgn}(\Theta[K+1 : 2K])$, where $\Theta[K+1 : 2K]$ denotes the subvector of Θ consisting of the elements indexed from $K+1$ through $2K$.

On the other hand, to do blind robust multiuser detection in asynchronous channels, we first compute the signal subspace components of the covariance matrix of the signal $\{\mathbf{r}_m(i)\}$

$$\mathbf{C} \triangleq E\{\mathbf{r}_m(i)\mathbf{r}_m(i)^T\} = \mathcal{H}_m \mathcal{A}_m^2 \mathcal{H}_m^T + \sigma^2 \underline{I}_{Nm}$$
$$= \mathbf{U}_s \Lambda_s \mathbf{U}_s^T + \sigma^2 \mathbf{U}_n \mathbf{U}_n^T. \quad (81)$$
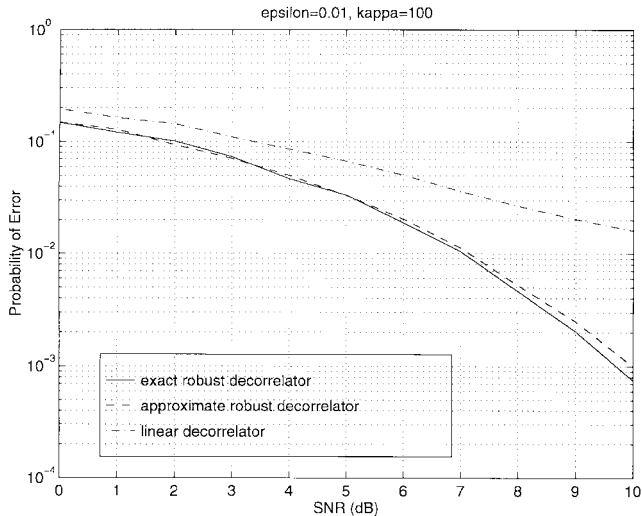
Fig. 12. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the robust detector and the linear detector in an asynchronous CDMA channel with non-Gaussian noise. $N = 31$, $K = 6$. The powers of the interferers are 10 dB above the power of user 1.
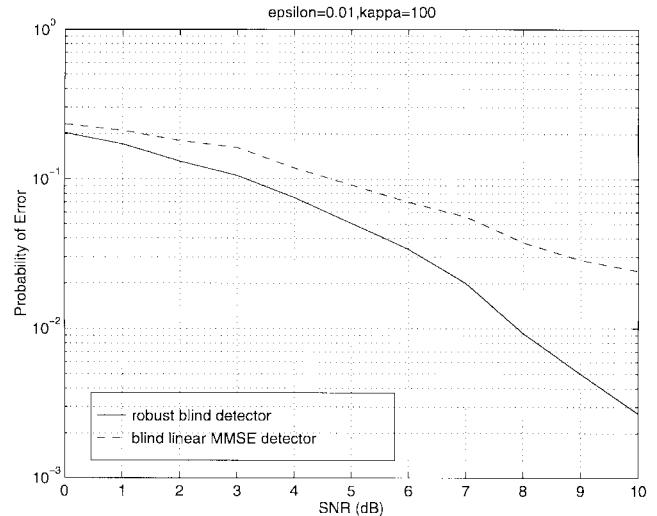

Fig. 13. Probability of error versus signal-to-noise ratio (SNR) for user 1 for the blind robust detector and the blind linear detector, based on the subspace tracking in an asynchronous CDMA channel with non-Gaussian noise. $N = 31$, $K = 6$. The powers of the interferers are 10 dB above the power of user 1.

Notice that the signal subspace has dimension $K(m+1)$. We then proceed with the robust estimation in the signal subspace as

$$\mathbf{z}^l \triangleq \psi(\mathbf{r} - \mathbf{U}_s \zeta^l) \qquad (82)$$

$$\zeta^{l+1} = \zeta^l + \frac{1}{\mu} \mathbf{U}_s^T \mathbf{z}^l. \qquad (83)$$

Finally, the kth user data bit is demodulated according to $\hat{b}_k = \operatorname{sgn}(\theta_k)$, where

$$\theta_k = \sum_{j=1}^{K(m+1)} \frac{\mathbf{u}_j^T \mathbf{h}_k}{\lambda_j} \zeta_j, \quad k = 1, \ldots, K \qquad (84)$$

with $\mathbf{h}_k \triangleq [\underline{h}_k^T \ \underline{0}_{N(m-2)}^T]^T$.

B. Simulation Examples

We consider an asynchronous CDMA system with spreading gain $N = 31$ and the number of users $K = 6$. The user spreading sequences are the same as before. The delays of the six users are randomly generated and are given by $[24.73 \ 29.46 \ 14.00 \ 1.93 \ 15.45 \ 16.78]T_c$. The performance of the robust multiuser detectors and the linear decorrelating detector in this system is plotted in Fig. 12. As before, both the exact robust multiuser detector and the approximate robust multiuser detector are considered. It is seen that, similarly to the synchronous case, the robust detectors significantly outperform the linear decorrelator. Moreover, the performance of the robust detector is insensitive to the parameters.

Fig. 13 depicts the performance of the subspace-based blind adaptive robust multiuser detector for the same asynchronous CDMA system. The performance of the subspace-based blind adaptive linear MMSE detector is also plotted. It is seen that the robust techniques offer significant performance gain in impulsive noise.

VI. CONCLUDING REMARKS

In many practical wireless channels in which multiuser detection techniques may be applied, the ambient noise is likely to have an impulsive component that gives rise to larger tail probabilities than is predicted by the Gaussian model. Impulsive noise can seriously degrade the error probability of the linear multiuser detectors for a given level of the noise power. In this paper, we have proposed a robust multiuser detection technique that is seen to significantly outperform linear multiuser detectors in non-Gaussian ambient noise. This technique is based on the M-estimation method for robust regression. The asymptotic performance of several versions of the robust detectors are analyzed and compared. Simulations show that the robust multiuser detectors offer significant performance gain over the linear multiuser detectors in non-Gaussian noise, and they incur little performance loss when the channel noise is indeed Gaussian.

We have also developed a subspace-based blind adaptive technique for implementing the robust multiuser detectors, which requires only the signature waveform of the user of interest in order to robustly demodulate that user's signal. Finally, we have shown that the robust multiuser detection method and its blind adaptive version can be readily applied to asynchronous CDMA channels.

The robust multiuser detection methods developed in this paper can be extended in a number of ways. For example, the robust regression techniques can be applied to robustify some nonlinear multiuser detectors, such as the decision-feedback multiuser detectors [7] and the multistage multiuser detectors [24]. A previous study [23] indicates that these nonlinear detectors suffer similar performance loss in the presence of impulsive noise, just as do linear detectors. Another possible extension is to incorporate the channel dispersion due to multipath fading effects. A general framework for blind joint equalization and multiuser detection in dispersive

channels with Gaussian ambient noise is developed in [26]. By integrating the robust methods proposed in this paper into such a framework, nonlinear signal processing techniques can be developed for jointly combating multiple-access interference, intersymbol interference, and impulsive ambient channel noise.

APPENDIX A
DERIVATION OF (59)

We follow the technique used in [13] by defining the function

$$
\begin{aligned}
d(\underline{\tau}) = {} & \frac{1}{\mu}[C(\underline{\theta}^l) - C(\underline{\theta}^l + \underline{\tau})] \\
& + \frac{1}{2}\left[\underline{\tau}^T(\underline{S}^T\underline{S})\underline{\tau} - \frac{2}{\mu}\underline{\tau}^T\underline{S}^T\underline{z}(\underline{\theta}^l)\right], \quad \underline{\tau} \in \mathcal{R}^K. \quad (85)
\end{aligned}
$$

Notice that

$$
d(\underline{0}) = 0 \quad (86)
$$

$$
\frac{\partial}{\partial \underline{\tau}}d(\underline{0}) = \frac{1}{\mu}\sum_{j=1}^{N}\psi\left(r_j - \underline{\xi}_j^T(\underline{\theta}^l + \underline{\tau})\right)\underline{\xi}_j
$$

$$
+ \underline{S}^T\underline{S}\underline{\tau} - \frac{1}{\mu}\underline{S}^T\underline{z}(\underline{\theta}^l)|_{\underline{\tau}=\underline{0}} \quad (87)
$$

$$
= \frac{1}{\mu}\underline{S}^T\underline{z}(\underline{\theta}^l + \underline{\tau}) + \underline{S}^T\underline{S}\underline{\tau}
$$

$$
- \frac{1}{\mu}\underline{S}^T\underline{z}(\underline{\theta}^l)|_{\underline{\tau}=\underline{0}} = \underline{0} \quad (88)
$$

$$
\frac{\partial^2}{\partial^2 \underline{\tau}}d(\underline{\tau}) = -\frac{1}{\mu}\sum_{j=1}^{N}\psi'\left(r_j - \underline{\xi}_j^T(\underline{\theta}^l + \underline{\tau})\right)\underline{\xi}_j\underline{\xi}_j^T + \underline{S}^T\underline{S}
$$

$$
\succeq -\sum_{j=1}^{N}\underline{\xi}_j\underline{\xi}_j^T + \underline{S}^T\underline{S} = \underline{O} \quad (89)
$$

where in (89), $\underline{O}$ denotes a $K \times K$ zero matrix, and $\underline{A} \succeq \underline{B}$ denotes that the matrix $\underline{A}-\underline{B}$ is positive semidefinite. Equation (89) follows from the assumption that $\psi'(x) \leq \mu$. It then follows from (87)–(89) that $d(\underline{\tau}) \geq 0$ for any $\underline{\tau} \in \mathcal{R}^K$. Now, letting $\underline{\tau} = \underline{\theta}^{l+1} - \underline{\theta}^l = \frac{1}{\mu}(\underline{S}^T\underline{S})^{-1}\underline{S}^T\underline{z}(\underline{\theta}^l)$, we obtain

$$
\begin{aligned}
0 \leq d(\underline{\tau}) = {} & \frac{1}{\mu}[C(\underline{\theta}^l) - C(\underline{\theta}^{l+1})] \\
& - \frac{1}{2\mu^2}\underline{z}(\underline{\theta}^l)^T\underline{S}(\underline{S}^T\underline{S})^{-1}\underline{S}^T\underline{z}(\underline{\theta}^l) \\
= {} & \frac{1}{\mu}[C(\underline{\theta}^l) - C(\underline{\theta}^{l+1})] - \frac{1}{2}\underline{\tau}^T(\underline{S}^T\underline{S})\underline{\tau}. \quad (90)
\end{aligned}
$$

Assume that the penalty function $\rho(x)$ is convex and bounded from below; then, the cost function $C(\underline{\theta})$ is convex and has a unique minimum $C(\underline{\theta}^*)$. In this case, $\underline{\theta}^*$ is the unique solution to (19) such that $\underline{z}(\underline{\theta}^*) = \underline{0}$. Since the sequence $C(\underline{\theta}^l)$ is decreasing and bounded from below, it converges. Therefore, from (90), we have

$$
\begin{aligned}
& \underline{z}(\underline{\theta}^l)^T[\underline{S}\,\underline{R}^{-1}\underline{S}^T]\underline{z}(\underline{\theta}^l) \\
& \quad \leq 2\mu[C(\underline{\theta}^l) - C(\underline{\theta}^{l+1})] \to 0, \quad \text{as } l \to \infty. \quad (91)
\end{aligned}
$$

Since the matrix $[\underline{S}\underline{R}^{-1}\underline{S}^T]$ has full rank, (91) implies that $\underline{z}(\underline{\theta}^l) \to \underline{0}$. Since $\underline{z}(\underline{\theta})$ is a continuous function of $\underline{\theta}$ and has a unique minimum point $\underline{\theta}^*$, we thus have $\underline{\theta}^l \to \underline{\theta}^*$ as $l \to \infty$.

APPENDIX B
DERIVATION OF (68)

Denote $\underline{\zeta} \triangleq [\zeta_1 \ \cdots \ \zeta_K]^T$. Then, (67) can be written in matrix form as

$$
\underline{S}\,\underline{\theta} = \underline{U}_s\underline{\zeta}. \quad (92)
$$

Denote $\underline{\Delta}_0 \triangleq \underline{\Delta}_s - \sigma^2\underline{I}_K = \text{diag}(\lambda_1 - \sigma^2, \ldots, \lambda_K - \sigma^2)$. Then, from (65) and (66), we obtain

$$
\underline{S}\,\underline{A}\,\underline{S}^T = \underline{U}_s\underline{\Delta}_0\underline{U}_s^T. \quad (93)
$$

Using (92) and (93), we obtain

$$
\underline{S}^T(\underline{S}\,\underline{A}\,\underline{S}^T)^\dagger\underline{S}\,\underline{\theta} = \underline{S}^T(\underline{U}_s\underline{\Delta}_0^{-1}\underline{U}_s^T)\underline{U}_s\underline{\zeta} \quad (94)
$$

$$
\Rightarrow (\underline{S}^T\underline{S}^{T^\dagger})\underline{A}^{-1}(\underline{S}^\dagger\underline{S})\underline{\theta} = \underline{S}^T\underline{U}_s\underline{\Delta}_0^{-1}(\underline{U}_s^T\underline{U}_s)\underline{\zeta} \quad (95)
$$

$$
\Rightarrow \underline{\theta} = \underline{A}\underline{S}^T\underline{U}_s\underline{\Delta}_0^{-1}\underline{\zeta} \quad (96)
$$

where in (94), $\dagger$ denotes the Moore–Penrose generalized matrix inverse [12]; in (95), we used the fact that $(\underline{S}\,\underline{A}\,\underline{S}^T)^\dagger = \underline{S}^{T^\dagger}\underline{A}^{-1}\underline{S}^\dagger$, which can be easily verified by using the definition of the Moore–Penrose generalized matrix inverse [12]; in (95), we used the facts that $(\underline{S}^T\underline{S}^{T^\dagger}) = (\underline{S}^\dagger\underline{S}) = (\underline{U}_s^T\underline{U}_s) = \underline{I}_K$. Equation (96) is the matrix form of (68). Finally, we notice that

$$
\begin{aligned}
\underline{A}^{-1} &= \underline{S}^T(\underline{S}\,\underline{A}\,\underline{S}^T)^\dagger\underline{S} \\
&= \underline{S}^T(\underline{U}_s\underline{\Delta}_0^{-1}\underline{U}_s^T)\underline{S}. \quad (97)
\end{aligned}
$$

It follows from (97) that the kth diagonal element A_k^{-2} of the diagonal matrix $\underline{A}^{-1}$ satisfies $A_k^{-2} = \sum_{j=1}^{K}\frac{(\underline{u}_j^T\underline{s}_k)^2}{\lambda_j - \sigma^2}$.

REFERENCES

[1] B. Aazhang and H. V. Poor, "Peformance of DS/SSMA communications in impulsive channels—Part I: Linear correlation receivers," *IEEE Trans. Commun.*, vol. COMM-35, pp. 1179–1187, Nov. 1987.
[2] ——, "Peformance of DS/SSMA communications in impulsive channels—Part II: Hard-limiting correlation receivers," *IEEE Trans. Commun.*, vol. 36, pp. 88–96, Jan. 1988.
[3] ——, "An analysis of nonlinear direct-sequence correlators," *IEEE Trans. Commun.*, vol. 37, pp. 723–731, July 1989.
[4] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 991–1001, Sept. 1993.
[5] T. K. Blankenship, D. M. Krizman, and T. S. Rappaport, "Measurements and simulation of radio frequency impulsive noise in hospitals and clinics," in *Proc. 1997 IEEE Veh. Technol. Conf. (VTC'97)*, 1997, pp. 1942–1946.
[6] P. L. Brockett, M. Hinich, and G. R. Wilson, "Nonlinear and non-Gaussian ocean noise," *J. Acoust. Soc. Amer.*, vol. 82, pp. 1286–1399, 1987.
[7] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for synchronous code-division multiple access channel," *IEEE Trans. Commun.*, vol. 43, pp. 421–434, Feb./Mar./Apr. 1995.

[8] F. R. Hampel *et al.*, *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley, 1986.

[9] L. M. Garth and H. V. Poor, "Narrowband interference suppression techniques in impulsive channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 15–34, Jan. 1992.

[10] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[11] M. Honig and H. V. Poor, "Adaptive interference suppression," *Wireless Communications: Signal Processing Perspectives*, in H. V. Poor and G. W. Wornell, Eds. Upper Saddle River, NJ: Prentice-Hall, 1998, pp. 64–128.

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[13] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[14] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, 1985.

[15] Y. Li and K. J. R. Liu, "On blind MIMO channel identification using second-order statistics," in *Proc. 1996 Conf. Inform. Sci. Syst.*, Princeton, NJ, Mar. 1996.

[16] R. Lupas and S. Verdú, "Linear multi-user detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.

[17] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurement," *IEEE Trans. Commun.*, vol. COMM-21, pp. 1232–1241, 1973.

[18] ———, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. 19, pp. 106–127, 1977.

[19] ———, "Channel modeling and threshold signal processing in underwater acoustics: An analytical overview," *IEEE J. Oceanic Eng.*, vol. OE-12, pp. 4–28, 1987.

[20] D. Middleton and A. D. Spaulding, "Elements of weak signal detection in non-Gaussian noise," in *Advances in Statistical Signal Processing Vol. 2: Signal Detection*, H. V. Poor and J. B. Thomas, Eds. Greenwich, CT: JAI, 1993.

[21] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

[22] ———, "Non-Gaussian signal processing problems in multiple-access communications," in *Proc. 1996 USC/CRASP Workshop Non-Gaussian Signal Process.*, Ft. George Meade, MD, May 1996.

[23] H. V. Poor and M. Tanda, "An analysis of some multiuser detectors in impulsive noise," in *Proc. 16th GRETSI Symp. Signal Image Process.*, Grenoble, France, Sept. 1997.

[24] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, Apr. 1990.

[25] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[26] X. Wang and H. V. Poor, "Blind equalization and multiuser detection for CDMA communications in dispersive channels," *IEEE Trans. Commun.*, vol. 46, pp. 91–103, Jan. 1998.

[27] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–691, Mar. 1998.

[28] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Processing*, vol. 44, pp. 95–107, Jan. 1995.

[29] S. M. Zabin and H. V. Poor, "Efficient estimation of the class A parameters via the EM algorithm," *IEEE Trans. Inform. Theory*, vol. 37, pp. 60–72, Jan. 1991.

Xiaodong Wang (M'98) received the B.S. degree in electrical engineering and applied mathematics (with highest honors) from Shanghai Jiao Tong University, Shanghai, China, in 1992, the M.S. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 1995, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1998.

In June 1998, he joined the Department of Electrical Engineering, Texas A&M University, College Station, as an Assistant Professor. His research interests fall in the general areas of computing and communications. He has worked in the areas of digital communications, digital signal processing, parallel and distributed computing, nanoelectronics, and quantum computing. His current research interests include multiuser communications theory and advanced signal processing for wireless communications. He worked at AT&T Labs—Research, Red Bank, NJ, during the summer of 1997.
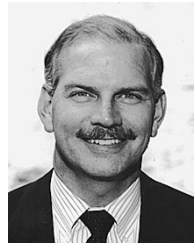
H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until he joined the Princeton faculty in 1990, he was a faculty member at the University of Illinois, Urbana-Champaign. He has also held visiting and summer appointments at several universities and research organizations in the United States, Britain, and Australia. He is currently Professor of Electrical Engineering at Princeton University. His research interests are in the area of statistical signal processing and its applications, primarily in wireless communications. His publications in this area include the the recent book, *Wireless Communications: Signal Processing Perspectives*, co-edited with G. Wornell (Englewood Cliffs, NJ: Prentice-Hall, 1998).

Dr. Poor is a Fellow of the Acoustical Society of America and of the American Association for the Advancement of Science. He has been involved in a number of IEEE activities, including serving as President of the IEEE Information Theory Society in 1990 and as a member of the IEEE Board of Directors in 1991 and 1992. In 1992, he received the Terman Award from the American Society for Engineering Education, and in 1994, he received the Distinguished Member Award from the IEEE Control Systems Society.