Nonparametric Multiuser Detection in Non-Gaussian Channels

Babak Seyfe and Ahmad R. Sharafat, *Senior Member, IEEE*

***Abstract*—Existing multiuser detection techniques in wireless systems are based on the assumption that some information on the parameters of the probability density function (pdf) of ambient noise is available. Such information may not be available in all cases, particularly for non-Gaussian and impulsive noises, or may change depending on circumstances. In this paper, we present a technique for multiuser detection that does not require any *a priori* knowledge about the noise parameters. This method is based on using pseudo norms for linear nonparametric regression. Analytical and simulation results show that the proposed method offers an improved, or at least comparable, performance over existing robust techniques in the absence of any information on the nature of noise in the environment. The increased computational complexity is marginal compared to existing parametric detectors. In addition, the proposed nonparametric detector is portable in the sense that it does not need to be tuned for different noise models without any considerable degradation of performance. We also show that in non-Gaussian noise, the performance of blind adaptive nonparametric multiuser detectors is better than that of robust multiuser detectors.**

***Index Terms*—Nonparametric multiuser detection, pseudonorm, sign detector, Wilcoxon detector.**

I. Introduction

In the last two decades, multiuser detection has been the subject of continuous studies with a view to developing methods that would enhance the performance of multiple access communication systems [1]. The performance of such methods is limited by the interference generated by other users, called "multiple access interference (MAI)" and ambient noise in the environment. The majority of documented research assume that ambient noise is additive white Gaussian noise (AWGN) [1]. However, it is well recognized that even though the Gaussian noise model is mathematically appealing, it simply does not apply in many situations [2]. For example, in underwater acoustic communication systems [3]–[5], urban and indoor wireless systems [6]–[10], low-frequency data recordings [11], and very high-speed digital subscriber line (VDSL) systems [12], ambient noise is known through experimental measurements to be decidedly non-Gaussian due to the impulsive nature of the nonmade electromagnetic interference (or other effects). The Gaussian noise is most entropic among all finite variance noise models. This implies that in non-Gaussian channels, the performance of a detector designed assuming Gaussian noise model would be inferior to that of a detector designed for non-Gaussian environments [14]. This has led researchers in recent years to address the problem of non-Gaussian noise in wireless systems [13]–[21].

It has been shown that a conventional (i.e., single user) detector designed for non-Gaussian environments but used for multiuser detection has a better performance in non-Gaussian environments compared to a multiuser detector designed for Gaussian noise [13], [21]. Considering this, techniques were developed for designing multiuser detectors assuming that the parameters of the actual non-Gaussian noise model are available [13], [15], [22]. In particular, robust methods developed by Huber [23] were used by Wang and Poor to design the minimax multiuser detector [13]. However, the performance of such detectors in particular and parametric methods in general, depends on the availability of the parameters of the non-Gaussian noise model.

Notwithstanding the preference for an accurate and actual non-Gaussian noise model, deriving the parameters of such a model is usually difficult, and is impossible in some cases. Therefore, there is a need to devise nonparametric detectors which work without accurate information on the noise model.

Such methods have been widely used in various signal processing applications with encouraging results [24]–[27]. In this paper, we propose a nonparametric multiuser detector that uses minimal assumptions about noise characteristics. We show that in such environments, our proposed detector performs better than existing robust detectors (especially in highly impulsive noise). Our proposed detector uses pseudonorms for nonparametric regression by applying a system model that is based on *regression with an intercept parameter* instead of the conventional system model that is based on *regression through the origin*. Since regression is not performed through the origin, we therefore need to estimate the value of *intercept parameter*.

This paper is organized as follows. In Section II, we describe the synchronous CDMA model studied in this paper, as well as the impulsive noise model, and the conventional and nonparametric models for multiuser detection in non-Gaussian channels. In Section III, we develop our nonparametric multiuser detector, followed by Section IV, where we propose a blind nonparametric multiuser detector using subspace methods. Section V examines simulation results, and finally in Section VI, we present our conclusions.

Manuscript received December 1, 2003; revised January 13, 2005. This work was supported in part by Tarbiat Modarres University, Tehran, Iran. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Joseph Tabrikian.

B. Seyfe was with the Department of Electrical Engineering, Tarbiat Modarres University, Tehran, Iran. He is currently with the Centre for Digital Signal Processing Research, King's College, London, U.K. (e-mail: seyfe@iit.modares.ac.ir).

A. R. Sharafat is with the Department of Electrical Engineering, Tarbiat Modarres University, Tehran, Iran (e-mail: sharafat@isc.iranet.net).

Digital Object Identifier 10.1109/TSP.2005.861102

II. System Model

A. The Conventional System Model for Multiuser Detection

We assume a baseband digital synchronous DS-CDMA system. A waveform received by a terminal with a coherent BPSK modulation is modeled as

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

where $n(t)$ is the ambient noise (assumed to be white), and $S(t)$ is the transmitted signal. The received signal is comprised of data signals of K active users and noise. It passes through a chip-rate sampler and produces an output vector with N elements during each symbol interval T. Thus

$$\mathbf{r} = \mathbf{SAb} + \mathbf{n} \tag{2}$$

where $\mathbf{r}$ is a $N \times 1$ received signal vector and $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2 \ldots, \mathbf{s}_k]$ is a matrix whose column $\mathbf{s}_k, k = 1, \ldots, K$, is the normalized signature waveform of the kth user. In (2), $A = \mathrm{diag}(A_1, A_2, \ldots, A_k)$ is the matrix of users' amplitudes, $\mathbf{b} = [b_1, b_2, \ldots, b_k]^T$ is the vector of users' symbols ($b_k = \pm 1, k = 1, \ldots, K$), and $\mathbf{n}$ is a $N \times 1$ vector of independent and identically distributed (i.i.d.) random variables. Let $\boldsymbol{\theta} = \mathbf{Ab}$, then $\mathbf{b} = \mathrm{sgn}(\boldsymbol{\theta})$ where $\mathrm{sgn}(\boldsymbol{\theta})$ is the sign of $\boldsymbol{\theta}$. Now we have

$$\mathbf{r} = \mathbf{S}\boldsymbol{\theta} + \mathbf{n}. \tag{3}$$

We obtain an estimate of $\boldsymbol{\theta}$ by minimizing the sum of squared errors, i.e., through the least-squares (LS) method [1], [13]

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{\mathrm{LS}} &= \arg\min_{\boldsymbol{\theta}} \Sigma_{j=1}^{N} \left(\mathbf{r}_j - \Sigma_{k=1}^{k} s_j^k \theta_k\right)^2 \\ &= \arg\min_{\boldsymbol{\theta}} \|\mathbf{r} - \mathbf{S}\boldsymbol{\theta}\|^2\end{aligned} \tag{4}$$

where s_j^k is the jth element of the kth user's signature waveform. The Euclidean norm $\|\cdot\|$ is a measure of dispersion used to minimize the power of noise by minimizing the sum of squared residuals in (4). This measure is suitable for Gaussian noise, but the performance of a detector using this measure deteriorates in non-Gaussian noise models [13]. In the robust multiuser detector, the Huber's proposal in [23] is employed to minimize the sum of a less rapidly increasing function $(\rho_H(\cdot))$ of residuals in non-Gaussian environments [13]

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in R^K} \Sigma_{j=1}^{N} \rho_H\left(r_j - \Sigma_{k=1}^{K} s_j^k \theta_k\right) \tag{5}$$

where $\rho_H(\cdot)$ is defined as

$$\rho_H(x) = \begin{cases} \frac{x^2}{2\nu^2}, & |x| \leq \xi\nu^2 \\ \frac{\xi^2\nu^2}{2} - \xi|x|, & |x| > \xi\nu^2 \end{cases} \tag{6}$$

in which the index H in $\rho_H(\cdot)$ refers to the Huber penalty function, and values of v and ξ, depend on the distribution of additive noise. It has been shown in [29] that nonparametric methods can be based on *regression with an intercept parameter*. Therefore, we now modify the system model in such a way that the result would be useful for nonparametric detection.

B. The Modified System Model for Nonparametric Multiuser Detection

We modify the system model in (3) to obtain a system model based on *regression with an intercept parameter*. This enables us to use nonparametric regression to estimate the data vector $\boldsymbol{\theta}$ by estimating the intercept parameter. In (3), we center the matrix $\mathbf{S}$ to obtain $\mathbf{S}_c$ by

$$\mathbf{S}_c = \mathbf{S} - \mathbf{1}_N[\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_K] \tag{7}$$

where the scalar $\bar{s}_k (k = 1, 2, \ldots, K)$ is the arithmetic mean of the signature vector elements of the kth user and $\mathbf{1}_N$ is a $N \times 1$ vector whose elements are all equal to 1, respectively. We rewrite (3) as

$$\mathbf{r} = \mathbf{S}_c\boldsymbol{\theta} + \psi\mathbf{1}_N + \mathbf{n} \tag{8}$$

where $\psi = [\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_K]\boldsymbol{\theta}$ is a scalar. In (8), ψ is not an independent intercept parameter needed in nonparametric regression, because it is a function of $\boldsymbol{\theta}$ [29], In what follows, we obtain an estimate of $\boldsymbol{\theta}$ that is independent of ψ (i.e., location free).

Definition 1: Let $D(\cdot)$ be the dispersion function that satisfies the following two properties for every $\mathbf{r} \in R^N$ and $\gamma \in R$:

1-$D(\mathbf{r} - \gamma\mathbf{1}_N) = D(\mathbf{r})$
2-$D(-\mathbf{r}) = D(\mathbf{r})$

then $D(\cdot)$ is called an even and location-free dispersion function.

In this case, we use (7) and (8) to write

$$D(\mathbf{r} - \mathbf{S}\boldsymbol{\theta}) = D(\mathbf{r} - \psi\mathbf{1}_N - \mathbf{S}_c\boldsymbol{\theta}) = D(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta}). \tag{9}$$

Hence, we can use $\mathbf{S}$ and $\mathbf{S}_c$ interchangeably in $D(\cdot)$. We minimize $D(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta})$ as a function of (8) to obtain an estimate of $\boldsymbol{\theta}$. Thus, our modified model, which is a simple regression model, is

$$\mathbf{r} = \mathbf{S}_c\theta + \mathbf{n}. \tag{10}$$

Note that since we use the even and location free dispersion function, the model (10) is also based on regression through the origin. In [29], it is shown that in general, *even and location-free* dispersion functions cannot be used in the regression through the origin. Then, we transform (10) into a modified system model for nonparametric regression, which needs the intercept parameter, and write

$$\begin{aligned}\mathbf{r} &= \gamma\mathbf{1}_N + \mathbf{S}_c\boldsymbol{\theta}_1 + \mathbf{n} \\ &= [\mathbf{1}_N\mathbf{S}_c]\begin{bmatrix}\gamma \\ \boldsymbol{\theta}_1\end{bmatrix} + \mathbf{n}\end{aligned} \tag{11}$$

where the true γ is 0. Let $\mathbf{S}_{c1} = [1_N \ \ \mathbf{S}_c]$ and assume Ω and Ω_1 denote the column spaces of $\mathbf{S}_c$ and $\mathbf{S}_{c1}$ respectively. Let $\widehat{\mathbf{q}}_1 = \widehat{\gamma}\ \mathbf{1}_N + \mathbf{S}_c\ \widehat{\boldsymbol{\theta}}_1$, denote a fitted value based on the fit model (11). Note that $\widehat{\mathbf{q}}_1$, lies in the space of Ω_1. Since the space Ω does not include $\mathbf{1}_N$ then $\Omega \subset \Omega_1$. To obtain a fitted value that lies in the column space of $\mathbf{S}_c$, i.e., in the desired space Ω, we project $\widehat{\mathbf{q}}_1$ on to the space Ω. Then $\widehat{\mathbf{q}} = \mathbf{H}_\Omega\ \widehat{\mathbf{q}}_1$ is the projection of this fitted value on to the desired space Ω, where $H_\Omega = \mathbf{S}_c(\mathbf{S}_c^T\mathbf{S}_c)^{-1}\mathbf{S}_c^T$. Since each user's signature waveform is independent of other users' signature waveforms, $\mathbf{S}$ has a full rank K, and so has $\mathbf{S}_c$. Thus the inverse of $\mathbf{S}_c^T\mathbf{S}_c$ exists and then $\mathbf{H}_\Omega$ exists. Our estimate of $\boldsymbol{\theta}$ is denoted by $\widehat{\boldsymbol{\theta}}$, which is a solution to $\mathbf{S}_c\,\widehat{\boldsymbol{\theta}}=\mathbf{H}_\Omega\,\widehat{\mathbf{q}}_1$. Now, $\mathbf{S}_c\,\widehat{\boldsymbol{\theta}}$ lies in the column space of $\mathbf{S}_c$, We thus have

$$\widehat{\boldsymbol{\theta}}=\widehat{\boldsymbol{\theta}}_1+\widehat{\gamma}\,(\mathbf{S}_c^T\mathbf{S}_c)^{-1}\mathbf{S}_c^T\mathbf{1}_N, \tag{12}$$

Since $\mathbf{S}_c$ is available, we need an estimate of $\boldsymbol{\theta}_1$ and γ to obtain an estimate of $\boldsymbol{\theta}$. In order to compute $\boldsymbol{\theta}$ as a function of $\boldsymbol{\theta}_1$, and γ we modify (10) and write

$$\mathbf{r}=\mathbf{S}_c\boldsymbol{\theta}_1+\gamma\mathbf{1}_N+\mathbf{n}. \tag{13}$$

Note that (8) and (13) are different models of the same observation, and as such are not identical. Also note that while in this specific model and thanks to (7), the columns of $\mathbf{S}_c^T$ and the subspace spanned by $\mathbf{1}_N$ are orthogonal, for other models for regression through the origin [29] which do not center the matrix $\mathbf{S}$, the columns of modified $\mathbf{S}$ are not orthogonal to $\mathbf{1}_N$ and all such models can be used with (11) and (12). Thus, in general since $\boldsymbol{\theta}_1$ is independent of γ while $\boldsymbol{\theta}$ is not independent of ψ, we proceed with (13) and estimate $\boldsymbol{\theta}_1$. We use an even and location free dispersion function to estimate $\boldsymbol{\theta}_1$ and the intercept parameter γ in (13). We obtain an estimate of $\boldsymbol{\theta}_1$ by minimizing an even and location free dispersion function of $\boldsymbol{\theta}_1$ residuals. This function is based on *rank pseudonorm* that is discussed in Section III.

III. Nonparametric Multiuser Detection

In this section, we propose and analyze a nonparametric multiuser detector in CDMA channels with non-Gaussian ambient noise. We start by presenting some basic notions of nonparametric linear regression and *pseudonorm* from [29] and [30] that are needed for estimating $\boldsymbol{\theta}_1$ in (13).

A. Definition and Characteristics of Pseudonorm

We use Definition 1 to introduce a distance measure (in terms of norms) that is invariant to a *uniform shift in location*. A pseudonorm, as defined below, has this property.

Definition 2: An operator $\|\cdot\|$, is a pseudonorm if for $\forall\mathbf{u},\mathbf{v}\varepsilon R^N$ and $\forall\alpha\,\varepsilon R$, the following four conditions are satisfied:

$$\|\mathbf{u}+\mathbf{v}\|_*\le\|\mathbf{u}\|_*+\|\mathbf{v}\|_* \tag{14-1}$$
$$\|\alpha\mathbf{u}\|_*=|\alpha|\|\mathbf{u}\|_* \tag{14-2}$$
$$\|\mathbf{u}\|_*\ge 0 \tag{14-3}$$
$$\|\mathbf{u}\|_*=0 \text{ if only if } u_1=\ldots=u_N. \tag{14-4}$$

A regular norm satisfies the first three properties, but the fourth one forces the norm of any vector with equal elements to zero. Consider the N-dimensional vectors $\mathbf{z}\varepsilon R^N,\mathbf{c}\varepsilon R_N,\mathbf{1}_N$, and the scalar $\gamma\varepsilon R$. The following inequalities establish the invariance of pseudonorms to any *uniform shift in location*:

$$\begin{aligned}&\|\mathbf{z}-\gamma\mathbf{1}_N-\mathbf{c}\|_*\\&\quad\le\|\mathbf{z}-\mathbf{c}\|_*+\|\mathbf{1}_N\|_*=\|\mathbf{z}-\mathbf{c}\|_*\\&\quad=\|\mathbf{z}-\gamma\mathbf{1}_N-\mathbf{c}+\gamma\mathbf{1}_N\|_*\le\|\mathbf{z}-\gamma\mathbf{1}_N-\mathbf{c}\|_*.\end{aligned} \tag{15}$$

Hence, $\|z-\gamma\mathbf{1}_N-\mathbf{c}\|_*=\|z-c\|_*$.

Now we define a *rank pseudonorm* as

$$\|\mathbf{u}\|_*=\Sigma_{i=1}^N\alpha(R(u_i))u_i, \tag{16}$$

where $R(u_i)$ denotes the rank of u_i among $u_1,u_2,\ldots,u_N$, $a(R(u_i))$ are scores such that $a(1)\le a(2)\le\ldots\le a(N)$, and $\sum_i a(i)=0$ [29].

Theorem 1: Suppose $a(1)\le a(2)\le\ldots\le a(N),\Sigma_i a(i)=0$, and $a(i)=-a(N+1-i)$. Then the function $\|\mathbf{u}\|_*=\Sigma_{i=1}^N a(R(u_i))u_i$ is a pseudonorm.

Proof: See Appendix A

We repeat the following as presented by Hettmansperger and McKean [29] for easy reference here. If a set of scores satisfy conditions of Theorem 1, then an estimation that uses pseudonorms (called rank estimation or R-estimation) can be obtained. We obtain such estimation by using a general rank score of the form

$$a_\varphi(i)=\varphi\left(i/_{N+1}\right) \tag{17}$$

where $\varphi(x)$ is a bounded, nondecreasing, and differentiable function defined on the interval (0, 1) that satisfies

$$\int_0^1\varphi(x)dx=0 \text{ and } \int_0^1\varphi^2(x)dx=1. \tag{18}$$

Normalization of scores in (18) is for convenience. Wilcoxon scores [29], [30] are generated in this way by the linear function $\varphi_w(x)=\sqrt{1}2(x-1/2)$ and sign scores are generated by $\varphi_s(x)=\mathrm{sgn}(2x-1)$. For scores generated by $\varphi(x)$, we denote the corresponding pseudonorm by

$$\|u\|_\varphi=\Sigma_{i=1}^N a_\varphi(R(u_i))u_i. \tag{19}$$

B. Rank Estimation Using Pseudonorm

Now we define an even and location-free dispersion function $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$ based on pseudonorms that produces a rank estimate of $\boldsymbol{\theta}_1$. We denote the R-estimate of $\boldsymbol{\theta}_1$ by $\widehat{\boldsymbol{\theta}}_{1_\varphi}$ which is

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{1_\varphi}&=\arg\min_{\theta_1}D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)\\&=\arg\min_{\theta_1}\|\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1\|_\varphi\end{aligned} \tag{20}$$

where $\|\cdot\|_\varphi$ is a pseudonorm based on (18) and (19). Since $D_\varphi(\cdot)$ is expressed in terms of norm, it is a continuous and convex function of $\boldsymbol{\theta}_1$ as stated in [30].

Jaeckel in [31] points out that if $\mathbf{S}_c$ is a full rank matrix, then $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$ attains its minimum. Since $\mathbf{S}_c$ has full column rank K, then $\widehat{\boldsymbol{\theta}}_{1_\varphi}$ in (20) may be taken to be a value that minimizes $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$. In order to minimize $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$, we need to compute its gradient [30] as

$$\nabla D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)=-\mathbf{g}_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1) \tag{21}$$

where

$$\mathbf{g}_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)=\mathbf{S}_c^T\mathbf{a}_\varphi(R(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)), \tag{22}$$

$$\begin{aligned}\mathbf{a}_\varphi^T(R(\mathbf{r}-\mathbf{S}_c\theta_1))=\big[&\alpha_\varphi\left(R\left(r_1-\mathbf{s}_{c_1}^T\boldsymbol{\theta}_1\right)\right),\ldots,\\&\alpha_\varphi\left(R\left(r_N-\mathbf{s}_{c_N}^T\boldsymbol{\theta}_1\right)\right)\big]\end{aligned} \tag{23}$$

and $\mathbf{S}_{c_i}.(i = 1, 2, \ldots, N)$ is the ith column of $\mathbf{S}_c^T$. Thus $\widehat{\boldsymbol{\theta}}_{1\varphi}$ (the R-estimate of $\boldsymbol{\theta}_1$) is the solution to the following equations (also called R-normal equations):

$$\mathbf{g}_\varphi(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta}_1) = \mathbf{S}_c^T \mathbf{a}_\varphi(R(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta}_1)) = \mathbf{0}_K. \tag{24}$$

If noise has a symmetric distribution (i.e., $f(x) = f(-x)$) the R-estimate would be unbiased for all sample sizes of N [29].

We assume that the noise density function f is absolutely continuous, and the Fisher Information as denoted by $I(f)$ satisfies the following:

$$0 < I(f) < \infty \tag{25-1}$$

where $I(f) = \int_{-\infty}^{+\infty} [f'^2(x)/f(x)]\,dx$. These assumptions imply that f is uniformly bounded and uniformly continuous. We rewrite $I(f)$ as $I(f) = \int_0^1 \varphi_f^2(x)dx$. and from [29] we know that for any given noise model, if the score is $\varphi_f(x) = f'(F^{-1}(x))/f(F^{-1}(x))$, then an estimate of $\boldsymbol{\theta}_1$, with the minimum error is obtained by the rank detector that uses this score, where $F(x)$ is the cumulative distribution function (cdf) of noise [29]. We further assume

$$\lim_{N\to\infty} N^{-1}\mathbf{S}_c^T\mathbf{S}_c = \mathbf{\Sigma} \tag{25-2}$$

where $\mathbf{\Sigma}$ is a positive definite matrix. We also assume that the Noether's condition, as stated below is imposed

$$\lim_{N\to\infty} \max_{1\le i\le N} \frac{S_{c_i}^{k^2}}{\sum_{j=1}^N S_{c_j}^{k^2}} \to 0 \text{ for all } k = 1, 2, \ldots, K \tag{25-3}$$

where S_{cj}^k is they jth element of the kth column of $\mathbf{S}_c$. In a CDMA system, the norm of users' signature vectors are linearly independent from one another, normalized and bounded, which implies that (25-2) and (25-3) are satisfied.

Now we will show that $D_\varphi(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta}_1)$ is approximately quadratic. Suppose

$$\begin{aligned} Q(\mathbf{r} - \mathbf{S}_c\,\widehat{\boldsymbol{\theta}}_1) = (2\tau_\varphi)^{-1}(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1)^T\mathbf{S}_c^T\mathbf{S}_c(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1)& \\ +(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1)^T\mathbf{g}_\varphi(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta}_1) + D_\varphi(\mathbf{r} - \mathbf{S}_c\boldsymbol{\theta}_1)& \end{aligned} \tag{26}$$

where $\widehat{\boldsymbol{\theta}}_1$ is an estimate of $\boldsymbol{\theta}_1$, and

$$\tau_\varphi^{-1} = \int_0^1 \varphi(u)\varphi_f(u)du \tag{27}$$

is a scale parameter. Note that $Q(\cdot)$ depends on τ_φ and $\boldsymbol{\theta}_1$, so it cannot be used to estimate $\boldsymbol{\theta}_1$. As we will show, $Q(\cdot)$ is quite useful for establishing the asymptotic properties of R-estimates. It will also lead us to a Gauss-Newton type algorithm for obtaining the R-estimate. Theorems (3.5.2)–(3.5.6) in [29] show that $Q(\cdot)$ provides a local asymptotic approximation to $D_\varphi(\cdot)$, and if $\widetilde{\boldsymbol{\theta}}_1$ minimizes the quadratic function $Q(\cdot)$, then under conditions (18), (25-1), (25-2), and (25-3) we have $\sqrt{N}(\widehat{\boldsymbol{\theta}}_{1_\varphi} - \widetilde{\boldsymbol{\theta}}_1) \xrightarrow{\text{Pr}} 0$, where θ_{1_φ} is the R-estimate of $\boldsymbol{\theta}_{1_\varphi}$ and $\xrightarrow{\text{Pr}}$ means the limit in probability.

Furthermore, in [29] it is shown that under conditions (18), (25-1), (25-2), and (25-3), we have $\sqrt{N}(\widehat{\boldsymbol{\theta}}_{1_\varphi} - \boldsymbol{\theta}_1) \xrightarrow{\text{Dist.}} \mathrm{N}_k(\mathbf{0_K}, \tau_\varphi^2\mathbf{\Sigma^{-1}})$ where $\widehat{\boldsymbol{\theta}}_{1_\varphi}$ is the R-estimate of $\boldsymbol{\theta}_1$, and $\mathrm{N}_K(\cdot,\cdot)$ is the K dimensional Gaussian distribution. Now, if the variance of noise σ^2 is finite, it is further shown in [29] that the Least-Squares (LS) estimate of $\boldsymbol{\theta}_1$ (denoted by $\widehat{\boldsymbol{\theta}}_{1_{LS}}$ satisfies $\sqrt{N}(\widehat{\boldsymbol{\theta}}_{1_{\text{LS}}} - \boldsymbol{\theta}_1) \xrightarrow{\text{Dist.}} N_K(\mathbf{0}_K, \sigma^2\mathbf{\Sigma}^{-1})$. Hence, the asymptotic efficiency of the R-estimate relative to the Least-Square (LS) estimate is $(\sigma^2)/(\tau_\varphi^2)$. If the pdf of noise is $f(x)$, the *asymptotic relative efficiency* (ARE) of the Wilcoxon detector relative to the LS-estimator [32] is

$$\mathrm{ARE}_{W,\mathrm{LS}} = 12\sigma^2 \left(\int_{-\infty}^{+\infty} f^2(x)\,dx\right)^2. \tag{28}$$

For example, if the noise is standard Gaussian, the asymptotic relative efficiency is 0.955. For longer-tailed noise distributions this efficiency is higher. In other words, the efficiency of Wilcoxon detector is better than that of the LS detector in impulsive noise. Lehmann [32] has shown that $\mathrm{ARE}_{W,\mathrm{LS}} \ge 0.864$ for all pdfs of noise. Also Huber [23] has shown that $\sqrt{N}(\widehat{\boldsymbol{\theta}}_{1_{\text{minimax}}} - \boldsymbol{\theta}_1) \xrightarrow{\text{Dist.}} \mathrm{N}_K(\mathbf{0}_K, \beta^2\sum^{-1})$, where $\widehat{\boldsymbol{\theta}}_{1_{\text{minimax}}}$ is the robust minimax estimate of $\boldsymbol{\theta}_1$ and $\beta^2 = (E[\phi^2])/(\{E[\phi']\}^2)$, in which $\phi(\cdot)$ is

$$\phi(x) = \begin{cases} \frac{x}{\sigma^2}, & |x| \le \xi\sigma^2 \\ \xi\mathrm{sgn}(x), & |x| > \xi\sigma^2 \end{cases} \tag{29}$$

where σ^2 is the noise variance and ζ is a parameter that depends on the noise model. The ARE of the nonparametric detector to the robust detector is $(\beta^2)/(\tau_\varphi^2)$.

C. Implementation of the Rank-Based Estimator

Now we consider four issues that are involved in the synthesis of our proposed nonparametric multiuser detector. We begin by estimating the scale parameter τ_φ, which we will use later to construct the rank based estimator in (35).

1) Estimates of the Scale Parameter τ_φ: In order to estimate $\boldsymbol{\theta}_1$ an estimate of τ_φ is needed. As we explain below, we use (33)–(35) to estimate $\boldsymbol{\theta}_1$ and update τ_φ in an iterative manner. We also use the score function φ defined in (17) to form our estimator in the following manner. Consider the standard score function

$$\varphi^*(u) = \frac{\varphi(u) - \varphi(0)}{\varphi(1) - \varphi(0)}. \tag{30}$$

Recall from (27) that $\tau_\varphi = 1/\eta$, where

$$\eta = \int_0^1 \varphi(u)\varphi_f(u)du. \tag{31}$$

We define $\widehat{H}_N(\cdot)$ as

$$\begin{aligned} \widehat{H}_N(y) = \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N \Bigg(\varphi^*\left(\frac{j}{N}\right)& \\ -\varphi^*\left(\frac{j-1}{N}\right)\Bigg)\, \mathrm{IN}(|\,\widehat{e}_{(i)} - \widehat{e}_{(j)}\,| \le y)& \end{aligned} \tag{32}$$

where $\mathrm{IN}(\Omega)$ is an indicator function such that if Ω is true then $\mathrm{IN} = 1$ otherwise $\mathrm{IN} = 0$; and $\widehat{\mathbf{e}} = [\widehat{e}_1, \widehat{e}_2 \ldots \widehat{e}_N]^T$ is the estimated residual vector (i.e., $\widehat{\mathbf{e}} = \mathbf{r} - \mathbf{S}_c\,\widehat{\boldsymbol{\theta}}_1$). Then $\widehat{e}_{(i)}, i =$

$1, \ldots N$ is the ith least element of $\widehat{\mathbf{e}}$. Let $\widehat{t}_{N,\delta}$ denote the δ quintile of $\widehat{H}_N$ (i.e., $\widehat{t}_{N,\delta}=\widehat{H}_N^{-1}(\delta)$). Our estimate of η as in [29] is

$$\widehat{\eta}_{N,\delta}=\widehat{\tau}_\varphi^{-1}=\frac{\big(\varphi(1)-\varphi(0)\big)\widehat{H}_N\left(\frac{t_{N,\delta}}{\sqrt{N}}\right)}{\frac{2t_{N,\delta}}{\sqrt{N}}}. \quad (33)$$

It is quite complicated to use $\widehat{H}_N^{-1}(\delta)$ and compute $\widehat{t}_{N,\delta}$. In [29] it is noted that we have to estimate $h_N(0)$ (denoted by $\widehat{h}_N(0)$), which is the derivative of $\widehat{H}_N(y)$ at $y=0$. For a small (near zero) value of $\widehat{t}_{n,\delta}/\sqrt{N}$, we note that $(\widehat{H}_N((\widehat{t}_{N,\delta})/(\sqrt{N}))/((\widehat{t}_{N,\delta})/(\sqrt{N}))$ would approximately be equal to $\widehat{h}_N(0)$. In our simulations, we set $\widehat{t}_{n,\delta}=0.1$ and $N=31$, which correspond to a small value $\widehat{t}_{N,\delta}/\sqrt{N}=0.018$. The consistency of the above estimation is shown in [33].

From the above, we conclude that $\widehat{\tau}_\varphi=1/\widehat{\eta}_{N,\delta}$ is a consistent estimate of τ_φ. In all cases it has been found [29] that the following simple degree of freedom correction is useful, which we have used in our simulations:

$$\widehat{\tau}\,\varphi=\sqrt{\frac{N}{N-k-1}}\,\widehat{\eta}^{-1}. \quad (34)$$

2) Iterative Algorithms for Obtaining the R-Estimate of $\boldsymbol{\theta}_1$: As stated earlier, the dispersion function $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$ is a continuous convex function of $\boldsymbol{\theta}_1$. Thus, gradient-type algorithms, such as the steepest descent, can be used to minimize $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$, but they are often slow. The following algorithm [29] is a Newton-type algorithm, which is based on the asymptotic quadraticity of $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$, and needs an initial estimate of $\boldsymbol{\theta}_1$ (denoted by $\widehat{\boldsymbol{\theta}}_1^{(0)}$). Let $\widehat{\mathbf{e}}^{(0)}=\mathbf{r}-\mathbf{S}_c\,\widehat{\boldsymbol{\theta}}_1^{(0)}$ denote the initial estimate of residuals and let $\widehat{\boldsymbol{\tau}}_\varphi^{(0)}$ denote the initial estimate of $\boldsymbol{\tau}_\varphi$ from (33). From (26) (i.e., the quadratic approximation of $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$), the estimate of $\boldsymbol{\theta}_1$ that minimizes $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$ is

$$\widehat{\boldsymbol{\theta}}_1^{(l+1)}=\widehat{\boldsymbol{\theta}}_1^{(l)}+\widehat{\tau}_\varphi^{(l)}\left(\mathbf{S}_c^T\mathbf{S}_c\right)^{-1}\mathbf{g}_\varphi\left(\mathbf{r}-\mathbf{S}_c\,\widehat{\boldsymbol{\theta}}_1^{(l)}\right) \quad (35)$$

where $l=0,1,2,\ldots$ is the iteration number. In practice, we want to know if $D_\varphi\big(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1^{(l+1)}\big)$ is less than $D_\varphi\big(r-S_c\boldsymbol{\theta}_1^{(l)}\big)$ before proceeding further. In Appendix B, we prove the convergence of this algorithm and show that when N is large, for any $l\geq 1$ we have

$$\widehat{\boldsymbol{\theta}}_1^{(l)}\overset{\text{as}\,l\to+\infty}{\longrightarrow}\widehat{\boldsymbol{\theta}}_{1_\varphi} \quad (36)$$

where $\widehat{\boldsymbol{\theta}}_{1_\varphi}$ denotes a value that minimizes $D_\varphi(\mathbf{r}-\mathbf{S}_c\boldsymbol{\theta}_1)$. In comparison with robust multiuser detection, nonparametric methods need to perform additional computations for $\mathbf{g}_\varphi\left(\mathbf{r}-\mathbf{S}_c\,\widehat{\boldsymbol{\theta}}_1^{(l)}\right)$ in (35), and to sort and rank the elements of residuals.

3) Estimation of the Intercept Parameter γ: In estimating the value of intercept parameter, the median of residuals has an important role. If $\widehat{\eta}_s$ is the median of residuals

$$\widehat{\gamma}_S=\mathrm{Median}(\widehat{\mathbf{e}}), \quad (37)$$

where $\widehat{\mathbf{e}}=\mathbf{r}-\mathbf{S}_c\widehat{\boldsymbol{\theta}}_1$ is the residual vector, then as shown in [29], we have $\sqrt{N}(\widehat{\gamma}_s-\gamma)\overset{\text{Dist.}}{\longrightarrow}\mathrm{N}_1(0,\tau_s^2)$, in which $\mathrm{N}(\cdot,\cdot)$ is the one dimensional Gaussian distribution, $\tau_s=[2f(m)]^{-1}$ and m is the median of noise distribution, i.e., $m=F^{-1}(1/2)$. This implies that $\widehat{\gamma}_s$ is an unbiased estimate of γ.

4) Estimation of Data Vector $\boldsymbol{\theta}$: Now, by using (12), (13), (35), and (37) we get

$$\widehat{\boldsymbol{\theta}}=\widehat{\boldsymbol{\theta}}_{1_\varphi}+\widehat{\gamma}_S\left(\mathbf{S}_c^T\mathbf{S}_c\right)^{-1}\mathbf{S}_c^T\mathbf{1}_N. \quad (38)$$

In order to compute $\boldsymbol{\theta}_1$ in (24), we can use the matrix of signature waveforms $\mathbf{S}$ instead of $\mathbf{S}_c$. This is due to the "location free" feature of the dispersion function $D_\varphi(\cdot)$.

IV. Blind Nonparametric Multiuser Detection

In order to detect data vectors, we need signature waveforms of all users. This information is available at the base station, but not at a user terminal. Thus, we need a method, called *blind multiuser detection*, which does not require any information on other users at a given user's terminal. In recent years, the *subspace approach*, also known as subspace tracking, has been used to develop simple blind algorithms [34]. We will use such algorithms to construct our blind nonparametric multiuser detector.

For subspace estimation, we use the *Projection Approximation Subspace Tracking with deflation* (PASTd) algorithm [35]–[37]. Several methods have been proposed in [38]–[42] to improve the performance of this algorithm. Here, we briefly present a few notations on this algorithm, and refer the interested reader to [35]–[37] for a more detailed description and analysis.

For convenience and without loss of generality, we assume that the signature waveforms $\{\mathbf{s}_k\}_{k=1}^K$ of all K users are linearly independent from each other. The autocorrelation matrix of the received signal $\mathbf{r}$ is

$$\mathbf{C}=E\{\mathbf{r}^T\mathbf{r}\}=\sum_{k=1}^K A_k^2\mathbf{s}_k\mathbf{s}_k^T+\sigma^2\mathbf{I}_N=\mathbf{S}\mathbf{A}^2\mathbf{S}^T+\sigma^2\mathbf{I}_N \quad (39)$$

where $\mathbf{A}^2=\mathrm{diag}(A_1^2,A_2^2,\ldots,A_K^2),\sigma^2$ is the variance of noise, and $\mathbf{I}_N$ is the identity matrix. By performing an eigen decomposition of matrix $\mathbf{C}$, we get

$$\mathbf{C}=\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T=[\mathbf{U}_S\mathbf{U}_n]\begin{bmatrix}\boldsymbol{\Lambda}_S & \\ & \boldsymbol{\Lambda}_n\end{bmatrix}\begin{bmatrix}\mathbf{U}_S^T\\ \mathbf{U}_n^T\end{bmatrix} \quad (40)$$

where $\mathbf{U}=[\mathbf{U}_s\mathbf{U}_n],\boldsymbol{\Lambda}=\mathrm{diag}[\boldsymbol{\Lambda}_S\ \ \boldsymbol{\Lambda}_n]$. In (40), $\Lambda_S=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_K)$ contains K largest eigenvalues of $\mathbf{C}$ in descending order, $\mathbf{U}_s=[\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_K]$ contains the corresponding orthonormal eigenvectors, $\Lambda_n=\sigma^2\mathbf{I}_{N-K}$, and $\mathbf{U}_n=[\mathbf{u}_{K+1},\mathbf{u}_{K+2},\ldots,\mathbf{u}_N]$ contains the $(N-K)$ orthonormal eigenvectors that correspond to the eigenvalue σ^2. It is easy to see that $\mathrm{range}(\mathbf{S})=\mathrm{range}(\mathbf{U}_s)$. The range space of $\mathbf{U}_s$ is called the *signal subspace* and its orthogonal complement, the *noise subspace*, is spanned by $\mathbf{U}_n$.

Now we use the PASTd algorithm [35] for blind nonparametric multiuser detection. Using (39) and (40) it has been shown in [13] that there is a vector $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots \omega_K]^T$ such that

$$\sum_{k=1}^{K} \theta_k \mathbf{s}_k = \sum_{j=1}^{K} \omega_j \mathbf{u}_j, \quad \theta_\mathrm{k} \in R, \quad \omega_\mathrm{j} \in R \tag{41}$$

and from [13] we obtain

$$\theta_k = \left[\sum_{i=1}^{K} \frac{(\mathbf{u}_i^T \mathbf{s}_k)^2}{\lambda_i - \sigma^2}\right]^{-1} \sum_{j=1}^{K} \frac{\mathbf{u}_j^T \mathbf{s}_k}{\lambda_j - \sigma^2} \omega_j, \quad k = 1, 2, \ldots, K. \tag{42}$$

Hence, the system model is

$$\begin{aligned}\mathbf{r} &= \mathbf{S}\boldsymbol{\theta} + \mathbf{n} \\ &= \mathbf{U}_S \boldsymbol{\omega} + \mathbf{n}\end{aligned} \tag{43}$$

where $\omega = [\omega_1, \omega_2, \ldots, \omega_K]^T$. We use a procedure similar to the one that we used in (7)–(13) to write

$$\mathbf{r} = \mathbf{U}_{S_c} \boldsymbol{\omega}_\mathbf{1} + \gamma_U \mathbf{1}_N + \mathbf{n}, \tag{44}$$

where $\mathbf{U}_{S_c} = \mathbf{U}_S - \mathbf{1}_N[\bar{u}_1 \bar{u}_2, \ldots \bar{u}_K]$ is the centered $\mathbf{U}_S, \gamma_U$ is a scalar, $\boldsymbol{\omega}_1$ is a vector and the scalar $\bar{u}_k, (\mathrm{k} = 1, 2, \ldots, K)$ is the arithmetic mean of corresponding eigenvector elements of the kth user. We employ nonparametric methods to estimate γ_U and $\boldsymbol{\omega}_1$ in (45) as follows:

$$\widehat{\boldsymbol{\omega}}_1^{(l+1)} = \widehat{\boldsymbol{\omega}}_1^{(l)} + \boldsymbol{\tau}_\varphi^{(l)} \left(\mathbf{U}_{S_c}^T \mathbf{U}_{S_c}\right)^{-1} \mathbf{g}_\varphi \left(\mathbf{r} - \mathbf{U}_{S_c} \widehat{\boldsymbol{\omega}}_1^{(l)}\right) \tag{45}$$

where $l = 0, 1, 2, \ldots$ is the iteration number and $-\mathbf{g}_\varphi(\cdot)$ is the gradient of dispersion function of residuals;. We can use a procedure similar to the one in the Appendix B to show that if $l \to +\infty$ then $\widehat{\boldsymbol{\omega}}_1^{(l)} \to \widehat{\boldsymbol{\omega}}_{1_\varphi}$, where $\widehat{\boldsymbol{\omega}}_{1_\varphi}$, is the R-estimate of ω_1. Also, for an estimate of γ_U, we have

$$\widehat{\gamma}_U = \mathrm{Median}(\mathbf{r} - \widehat{\mathbf{U}}_{S_c} \widehat{\boldsymbol{\omega}}_{1_\varphi}). \tag{46}$$

Now, we use the same notion as in (38) and write

$$\widehat{\boldsymbol{\omega}} = \widehat{\boldsymbol{\omega}}_1 + \widehat{\gamma}_U \left(\mathbf{U}_{S_c}^T \mathbf{U}_{S_c}\right)^{-1} \mathbf{U}_{S_c}^T \mathbf{1}_N. \tag{47}$$

We use the PASTd Algorithm [34], [35] to estimate the *signal subspace coordinates* $\mathbf{U}_S$, as denoted by $\widehat{\mathbf{U}}_S = [\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2, \ldots, \widehat{\mathbf{u}}_K]$. If we ignore σ^2 in (42), the sensitivity of our estimator to errors in estimating the parameters will be reduced [13]. Also, since $[\Sigma_{i=1}^k (\mathbf{u}_i^T \mathbf{S}_k)^2 / (\lambda_i - \sigma^2)]^{-1}$ in (42) is positive, it does not affect the sign of $\widehat{\boldsymbol{\theta}}_k$, and so we can write

$$\widehat{\theta}_k = \sum_{i=1}^{k} \frac{\widehat{\mathbf{u}}_i^T \mathbf{s}_k}{\widehat{\lambda}_i} \widehat{\omega}_i, \quad k = 1, \ldots, K. \tag{48}$$

Equation (48) provides a blind estimate of the kth user's data.

In Section V, we explain the behavior of our nonparametric detectors using simulation results.

V. Simulation Results

In this section, we compare the performances of nonparametric multiuser detectors and robust (minimax) multiuser detector in the presence of non-Gaussian noise. We consider a synchronous system in which the spreading sequence for each user is a shifted m-sequence. For robust multiuser detection, we note that the "approximate" minimax decorrelating detector converges [13], and so we simulate its performance. In all simulations we assume that the signature vectors' lengths are $N = 31$ and there are 6 users in a cell, i.e., $\boldsymbol{K} = \mathbf{6}$. All users have the same power, i.e., perfect power control.

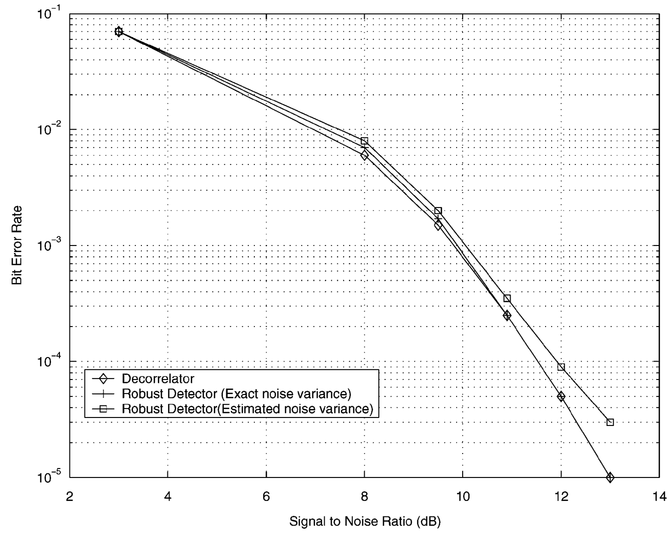

Fig. 1. Probability of error versus SNR for the robust multiuser detector in Gaussian noise for user 1 for the exact value of noise variance and for the estimated variance.

In Fig. 1 we show the bit error rate (BER) of the robust multiuser detector versus the signal-to-noise ratio (SNR). The SNR is the ratio of signal power to the noise variance. If the variance of noise is unknown, we need to estimate it. As can be seen in Fig. 1, this would result in degradation in the performance of robust detector particularly for large values of SNR.

For the robust detector, Wang and Poor in [13] used $\xi = 1.5/\sigma$ in (29), where σ^2 is the noise variance. For Gaussian or nearly Gaussian noise, this is accurate. But, our simulations show that for some non-Gaussian noise, this is not the case. For the Laplacian noise, with the following pdf:

$$f(x) = \frac{\mu}{2} e^{-\mu|x|}, \quad -\infty < x < +\infty \tag{49}$$

the value of ξ should be $1/\sigma$. Fig. 2 shows performance degradation of the robust (minimax) detector for $\xi = 1.5/\sigma$ in the Laplacian noise, which indicates that the value of ξ depends on the noise model.

Fig. 3 shows the performance of various detectors in the presence of Gaussian noise. It is shown that the performance of the nonparametric Wilcoxon detector closely follows that of the optimum detector in Gaussian noise, but has a maximum loss of about 0.4 dB compared to the decorrelator for SNR of about 8 dB.

In our simulations, we also use the Gaussian mixture noise with the following cdf:

$$F(x) = (1 - \varepsilon)\mathrm{N}_1(0, v^2) + \varepsilon \mathrm{N}_1(0, \kappa \nu^2) \tag{50}$$

where ε is the probability of impulse noise, $\mathrm{N}_1(0, \nu^2)$ is the zero mean one-dimensional Gaussian cdf with variance ν^2 that represents the background noise, and $\mathrm{N}_1(0, \boldsymbol{\kappa}\nu^2)$ is the impulsive part of noise. This is an approximation to the Middleton (Class

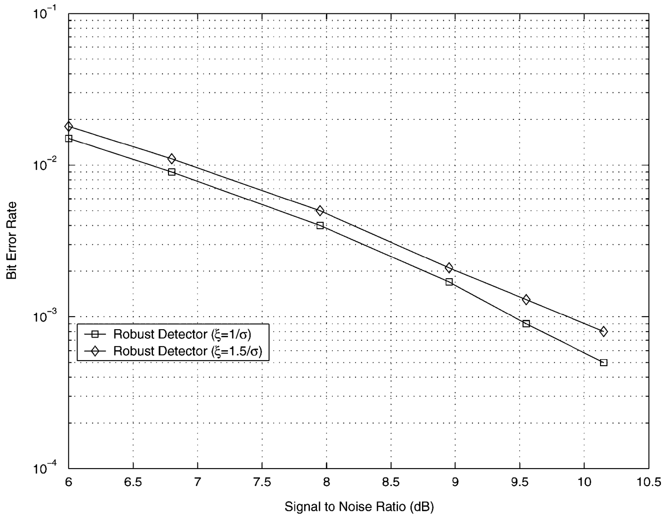

Fig. 2. Probability of error versus SNR for the robust multiuser detector in Laplacian noise for user 1 and for two values of ξ in (33).

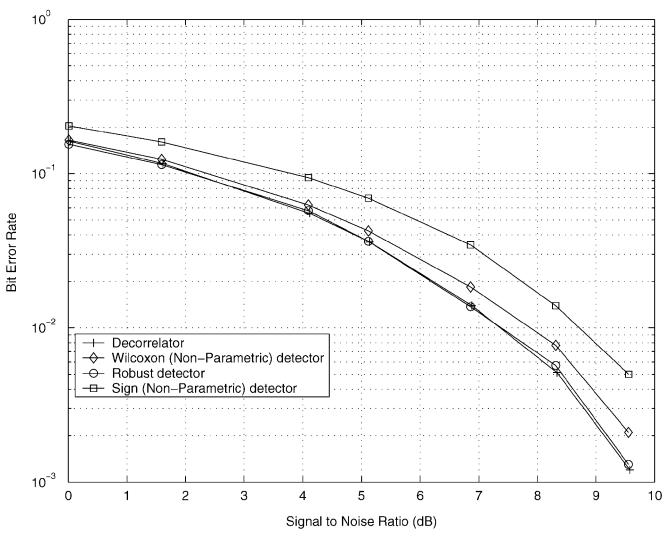

Fig. 3. Probability of error versus SNR for user 1 in the robust (minimax) detector, the Wilcoxon detector, the sign detector, and the decorrelator detector in a synchronous CDMA channel with Gaussian noise.

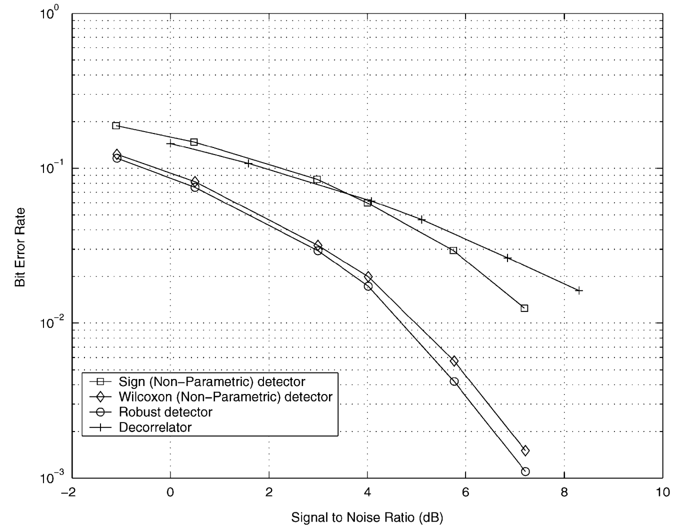

Fig. 4. Probability of error versus SNR for user 1 in the robust (minimax) detector, the Wilcoxon detector, the sign detector, and the decorrelator detector in a synchronous CDMA channel with impulsive noise where $\varepsilon = 0.01$ and $\boldsymbol{\kappa} = 100$.

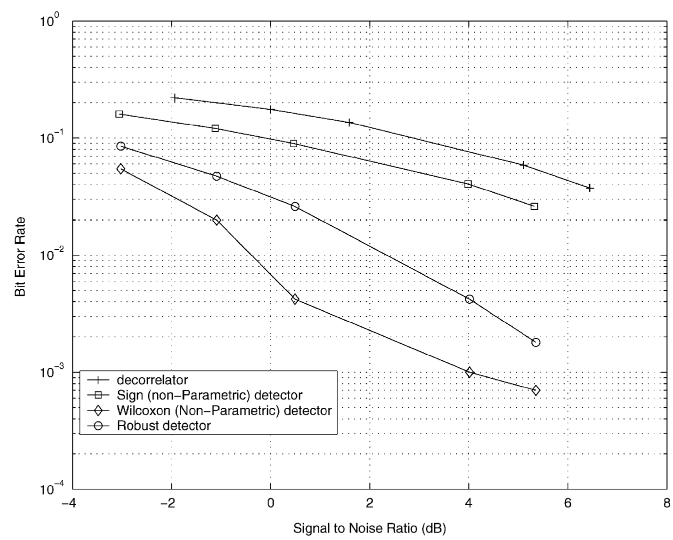

Fig. 5. Probability of error versus SNR for user 1 in the robust (minimax) detector, the Wilcoxon detector, the sign detector, and the decorrelator detector in a synchronous CDMA channel with impulsive noise where $\varepsilon = 0.1$ and $\kappa = 100$.

A) model for impulsive noise [8]–[10]. Fig. 4 shows the performances of various detectors in impulsive noise with $\varepsilon = 0.01$ and $\boldsymbol{\kappa} = 100$.

It shows that the Wilcoxon (nonparametric) detector has a good performance and closely follows the robust (minimax) detector. Note that nonparametric Wilcoxon and sign detectors do not use any *a priori* information on noise, but nevertheless the performance of Wilcoxon detector is very close to that of the robust detector in such impulsive noise environments.

Fig. 5 shows the performances of various detectors in heavy tailed (i.e., highly impulsive) noise with $\varepsilon = 0.1$ and $\boldsymbol{\kappa} = 100$. It shows that the Wilcoxon detector has more than 3 dB gain compared to the robust (minimax) detector when the signal-to-noise ratio is more than 0 dB. The nonparametric Wilcoxon detector has a good performance compared to parametric detectors without any information on impulsive noise. Poor and Tanda [19] have shown that the robust (minimax) detector has the best performance compared to all other suboptimum detectors known to them. However, we have shown here that in highly impulsive noise environment, our proposed nonparametric detector has a noticeable improvement compared to the robust (minimax) detector, and as such excels the performance of all known suboptimum detectors in such environments.

As stated earlier, signature vectors of interfering users in the downlink are unknown, which means that we have to use a blind multiuser detector. Fig. 6 shows the performance of the blind version of the same detectors in Gaussian noise using the subspace approach in [34]. It verifies that the performances of blind Wilcoxon detector in a Gaussian environment is marginally close to that of the blind robust detector.

Fig. 7 shows the performance of blind multiuser detectors in highly impulsive noise (i.e., $\boldsymbol{\varepsilon} = 0.1$ and $\boldsymbol{\kappa} = 100$). It shows

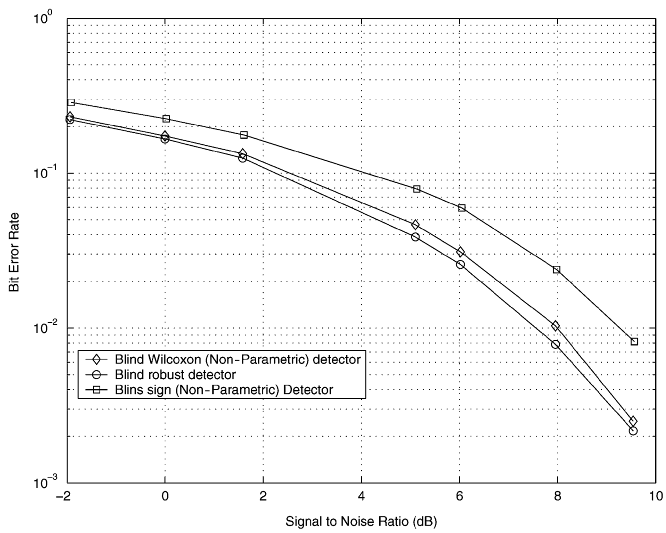

Fig. 6. Probability of error versus signal-to-noise ratio (SNR) for user 1 in the blind robust (minimax), the blind Wilcoxon and the blind sign detectors in a synchronous CDMA channel with Gaussian noise using subspace tracking.

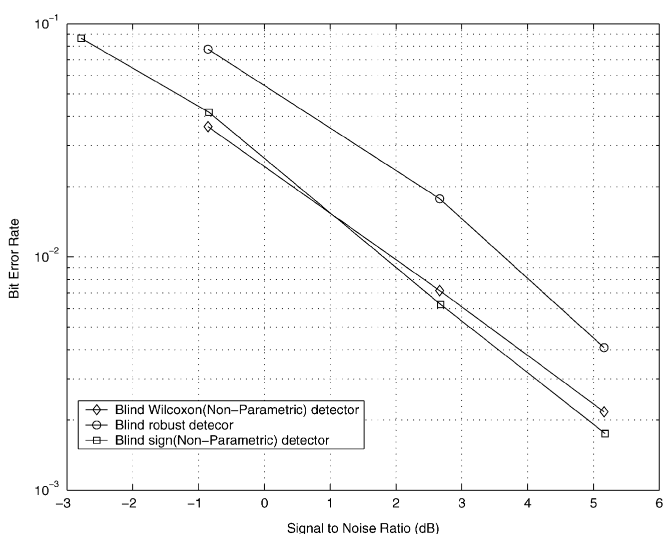

Fig. 7. Probability of error versus SNR for user 1 in the blind robust (minimax), the blind Wilcoxon, and the blind sign detectors in a synchronous CDMA channel with impulsive noise where $\varepsilon = 0.1$ and $\kappa = 100$.

good performances of blind nonparametric detectors compared to the blind robust detector, where the blind Wilcoxon and the blind sign detectors have about 2 dB gain over the blind robust detector.

VI. Conclusion

In this paper, we have developed a nonparametric multiuser detector for non-Gaussian noise environments. Performances of our proposed detectors are comparable to or better than that of the robust detector in the presence of non-Gaussian or impulsive noise. This is achieved without any *a priori* information on the noise model. The same is true for cases in which noise is Gaussian. For blind cases, the performance of our proposed Wilcoxon detector, is almost identical to that of the blind robust detector in Gaussian noise, but is even better in non-Gaussian environments.

Compared to robust multiuser detection, nonparametric methods need to perform additional computations for $\mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\widehat{\boldsymbol{\theta}}_1^{(l)}\right)$ in (35), and to sort and rank the elements of the residuals. These additional computations are the marginal cost of nonparametric multiuser detectors methods that do not require any *a priori* knowledge on the ambient noise as compared to robust multiuser detectors. For a nonparametric detector with lower computation complexity, see [43].

Appendix A
Proof of Theorem 1

We repeat a partial proof of Theorem 1 from [30], and provide additional material to construct a complete proof. We note the relation between rank and order statistics, and rewrite (16) as

$$\|\mathbf{u}\|_* = \sum_{i=1}^{N} a(i) u_{(i)} \tag{A.1}$$

where $u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(N)}$ are the ordered values of $u_i, i = 1, 2, \ldots, N$. Next, suppose that $u_{(j)}$ is the last order statistic with a negative score. Since scores sum to 0, we write

$$\begin{aligned}\|\mathbf{u}\|_* &= \sum_{i=1}^{N} a(i)\left(u_{(i)} - u_{(j)}\right)\\ &= \sum_{i\leq j} a(i)\left(u_{(i)} - u_{(j)}\right) + \sum_{i>j} a(i)\left(u_{(i)} - u_{(j)}\right).\end{aligned} \tag{A.2}$$

Both terms on the right-hand side of (A.2) are nonnegative, hence, $\|\mathbf{u}\|_* \geq 0$. If $\|\mathbf{u}\|_* = 0$, then all terms in (A.2) are zero. But since all scores are not $0, a(1) < 0$ and $a(N) > 0$, we must have $u_{(1)} = u_{(2)} = \cdots = u_{(N)}$. Conversely if $\boldsymbol{u}_{(1)} = \boldsymbol{u}_{(2)} = \cdots = \boldsymbol{u}_{(N)}$, then $\|\mathbf{u}\|_* = 0$. From $a(i) = -a(N+1-i)$, it follows that $\|\alpha\mathbf{u}\|_* = |\alpha|\|\mathbf{u}\|_*$. In order to prove this, we note that if $\alpha > 0$ then $a(R(\alpha u_i)) = a(R)(u_i))$ and hence, $\|\alpha\mathbf{u}\|_* = |\alpha|\|\mathbf{u}\|_*$. But, if $\alpha < 0$ then $a(R(\alpha u_i)) = -a(R(u_i))$. Thus

$$\begin{aligned}\|\alpha\mathbf{u}\|_* &= \sum_{i=1}^{N} a(R(\alpha u_i))\alpha u_i = \sum_{i=1}^{N} -a(R(u_i))\alpha u_i\\ &= |\alpha| \sum_{i=1}^{N} a(R(u_i))u_i = |\alpha|\,\|\mathbf{u}\|_*.\end{aligned} \tag{A.3}$$

The equality is obvious for $\alpha = 0$. To complete the proof, we show that the triangular inequality holds.

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|_* &= \sum_{i=1}^{N} a(R(u_i+v_i))(u_i+v_i)\\ &= \sum_{i=1}^{N} a(R(u_i+v_i))u_i + \sum_{i=1}^{N} a(R(u_i+v_i))v_i.\end{aligned} \tag{A.4}$$

Furthermore, by summing through another index, we write

$$\sum_{i=1}^{N} a(R(u_i+v_i))u_i = \sum_{j=1}^{N} p_j u_{(j)}, \tag{A.5}$$

where $p_1, p_2, \ldots, p_N$ is a permutation of integers $1, 2, \ldots, N$. Suppose p_j is not in order, then there exist a ζ and a ξ such that $u_{(\zeta)} \le u_{(\xi)}$, but $p_\zeta > p_\xi$. Thus

$$\begin{aligned}&\left[p_\xi u_{(\zeta)} + p_\zeta u_{(\xi)}\right] - \left[p_\zeta u_{(\zeta)} + p_\xi u_{(\xi)}\right]\\ &\quad = (p_\zeta - p_\xi)(u_\xi - u_\zeta)\\ &\quad \ge 0. \end{aligned} \tag{A.6}$$

Therefore, such an interchange never decreases the sum. This leads to

$$\sum_{i=1}^{N} a(R(u_i + v_i))u_i \le \sum_{i=1}^{N} a(i)u_{(i)}. \tag{A.7}$$

A similar result also holds for

$$\sum_{i=1}^{N} a(R(u_i + v_i))v_i \le \sum_{i=1}^{N} a(i)v_{(i)}. \tag{A.8}$$

Therefore, the triangular inequality is proved and the proof of Theorem 1 is completed. ■

Appendix B
Proof of the Convergence of Algorithm (35)

Now we prove the convergence of our iterative algorithm. This algorithm needs an initial estimate which we denote by $\widehat{\boldsymbol{\theta}}_1^{(0)}$. Let $\widehat{\mathbf{e}}_1^{(0)} = \mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(0)}$ denote the initial residuals. From (32) and (33), and using $\widehat{\boldsymbol{\theta}}_1^{(0)}$, we obtain an initial estimate of τ_φ (denoted by $\widehat{\boldsymbol{\tau}}_\varphi^{(0)}$. which according [33], is a consistent estimate of $\widehat{\boldsymbol{\tau}}_\varphi$ based on the residual vector $\widehat{\mathbf{e}}^{(0)}$. From (35), we have

$$\widehat{\boldsymbol{\theta}}_1^{(1)} = \widehat{\boldsymbol{\theta}}_1^{(0)} + \widehat{\boldsymbol{\tau}}_\varphi^{(0)} \left(\mathbf{S}_c^T \mathbf{S}_c\right)^{-1} \mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(0)}\right). \tag{B.1}$$

We recall (35) and write

$$\widehat{\boldsymbol{\theta}}_1^{(l+1)} = \widehat{\boldsymbol{\theta}}_1^{(l)} + \widehat{\tau}_\varphi^{(l)} \left(\mathbf{S}_c^T \mathbf{S}_c\right)^{-1} \mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right) \tag{B.2}$$

where $\widehat{\boldsymbol{\tau}}_\varphi^{(l)}$ is the consistent estimate of τ_φ based on the residual vector $\widehat{\mathbf{e}}^{(l)} = \mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}$ [33]. From (26) and for $\widehat{\boldsymbol{\theta}}_1^{(l+1)}$ we have

$$\begin{aligned} &Q\left(\mathbf{r} - \mathbf{S}_c \widehat{\boldsymbol{\theta}}_1^{(l+1)}\right)\\ &= \left(2\, \widehat{\boldsymbol{\tau}}_\varphi^{(l)}\right)^{-1} \left(\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right)^T\\ &\quad \times \mathbf{S}_c^T \mathbf{S}_c \left(\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right) - \left(\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right)^T\\ &\quad \times \mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right) + D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right). \end{aligned} \tag{B.3}$$

From [29, pp. 412], we note that for any $\varepsilon > 0$ and $c > 0$ we have

$$\begin{aligned} \lim_{N\to\infty} P\Bigg[\sup_{\left\|\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right\| \le c} &\Bigg| D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}\right)\\ &- Q\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}\right)\Bigg| \ge \varepsilon \Bigg] = 0. \end{aligned} \tag{B.4}$$

Then, for large values of N, we have $Q(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}) \to D_\varphi(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)})$, so

$$\begin{aligned} &D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}\right)\\ &= \left(2\, \widehat{\tau}_\varphi^{(l)}\right)^{-1} \left(\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right)^T\\ &\quad \times \mathbf{S}_c^T \mathbf{S}_c \left(\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right) - \left(\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)}\right)^T\\ &\quad \times \mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right) + D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right). \end{aligned} \tag{B.5}$$

Now we show that $D_\varphi(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}) \le D_\varphi(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)})$. From (B.2), we have

$$\widehat{\boldsymbol{\theta}}_1^{(l+1)} - \widehat{\boldsymbol{\theta}}_1^{(l)} = \widehat{\tau}_\varphi^{(l)} \left(\mathbf{S}_c^T \mathbf{S}_c\right)^{-1} \mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right). \tag{B.6}$$

We use (B.5) and (B.6) to write

$$\begin{aligned} &D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}\right) - D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right)\\ &= -\left(\frac{\widehat{\boldsymbol{\tau}}_\varphi^{(l)}}{2}\right) \left[\mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right)\right]^T\\ &\quad \times \left(\mathbf{S}_c^T \mathbf{S}_c\right)^{-1} \left[\mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right)\right]. \end{aligned} \tag{B.7}$$

Note that because $\varphi(x)$ is a nondecreasing function, then from (32) and (33) we conclude that $\widehat{\boldsymbol{\tau}}_\varphi^{(l)}$ is positive. Furthermore, since $\mathbf{S}_c$ has full rank K, then $\mathbf{S}_c^T \mathbf{S}_c$ also has full rank. Consequently, the right-hand side of (B.7) cannot be positive. Therefore

$$D_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}\right) - \mathbf{D}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right) \le 0. \tag{B.8}$$

If $D_\varphi(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l+1)}) - D_\varphi(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}) = 0$. then from (B.7) we have

$$\left[\mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right)\right]^T \left(\mathbf{S}_c^T \mathbf{S}_c\right)^{-1} \left[\mathbf{g}_\varphi\left(\mathbf{r} - \mathbf{S}_c\, \widehat{\boldsymbol{\theta}}_1^{(l)}\right)\right] = 0. \tag{B.9}$$

This implies that $g_{\varphi}(\mathbf{r} - \mathbf{S}_c \overset{\frown}{\boldsymbol{\theta}}{}_1^{(l)}) = 0_k$, which is because $\mathbf{S}_c^T \mathbf{S}_c$ has full rank. Then $\overset{\frown}{\boldsymbol{\theta}}{}_1^{(l)}$ (i.e., an R-estimate of $\boldsymbol{\theta}_1$) minimizes $D_{\varphi}(\mathbf{r} - \mathbf{S}_c \boldsymbol{\theta}_1)$. Alternatively, if $D_{\varphi}(\mathbf{r} - \mathbf{S}_c \overset{\frown}{\boldsymbol{\theta}}{}_1^{(l+1)}) < D_{\varphi}(\mathbf{r} - \mathbf{S}_c \overset{\frown}{\boldsymbol{\theta}}{}_1^{(l+1)}) < 0$, then

$$\mathbf{D}_{\varphi}\left(\mathbf{r} - \mathbf{S}_c \overset{\frown}{\boldsymbol{\theta}}{}_1^{(l+1)}\right) < \mathbf{D}_{\varphi}\left(\mathbf{r} - \mathbf{S}_c \overset{\frown}{\boldsymbol{\theta}}{}_1^{(l)}\right). \tag{B.10}$$

Because $D_{\varphi}(\mathbf{r} - \mathbf{S}_c \boldsymbol{\theta}_1)$ is convex and continuous, it has a unique minimum. We conclude that $D_{\varphi}(\mathbf{r} - \mathbf{S}_c \overset{\frown}{\boldsymbol{\theta}}{}_1^{(l+1)})$ approaches this unique minimum, i.e.,

$$\overset{\frown}{\boldsymbol{\theta}}{}_1^{(l)} \overset{\text{as } l \to +\infty}{\longrightarrow} \overset{\frown}{\boldsymbol{\theta}}{}_{1_{\varphi}} = \arg \min_{\boldsymbol{\theta}_1} D_{\varphi}(\mathbf{r} - \mathbf{S}_c \boldsymbol{\theta}_1). \tag{B.11}$$

In other words, if $l \to \infty$ then $\boldsymbol{\theta}_1^{(l)}$ converges to the unique minimum of $D_{\varphi}(\mathbf{r} - \mathbf{S}_c \boldsymbol{\theta}_1)$. ■

ACKNOWLEDGMENT

The authors would like to thank Prof. S. Valaee of University of Toronto, Toronto, ON, Canada, and Prof. M. Nasiri-Kenari of Sharif University of Technology, Tehran, Iran, for their valuable suggestions.

REFERENCES

[1] S. S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[2] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*: Springer Verlag, 1988.
[3] D. Brady and J. C. Preisig, "Underwater acoustic communications," in *Wireless Communications: Signal Processing Perspectives*, H. V. Poor and G. W. Wornell, Eds. Upper Saddle River, NJ: Prentice-Hall, 1998.
[4] 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.
[5] ——, "Channel modeling and threshold signal processing in underwater acoustics: An analytical overview," *IEEE J. Ocean. Eng.*, vol. OE-12, pp. 4–28, 1987.
[6] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 11, pp. 991–1001, Sep. 1993.
[7] 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. Tech. Conf. (VTC'97)*, 1997, pp. 1942–1946.
[8] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurement," *IEEE Trans. Commun.*, vol. COM-21, pp. 1232–1241, 1973.
[9] ——, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. 19, pp. 106–127, 1977.
[10] 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. New York: JAI, 1993.
[11] J. E. Evans and A. S. Griffith, "Design of a sanguine noise processor based upon world-wide extremely low frequency recordings," *IEEE Trans. Commun.*, vol. 22, pp. 528–539, 1974.
[12] M. D. Nava and C. Del-Toso, "A short overview of the VDSL system requirements," *IEEE Commun. Mag.*, pp. 82–90, Dec. 2002.
[13] X. Wang and H. V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 289–305, Feb. 1999.
[14] S. N. Batalama, M. J. Medley, and I. N. Psaromiligkos, "Adaptive robust spread-spectrum receivers," *IEEE Trans. Comm.*, vol. 47, pp. 905–917, 1999.
[15] P. Spasojevic and X. Wang, "Improved robust multiuser detection in non-Gaussian noise," *IEEE Signal Process. Lett.*, vol. 8, pp. 83–86, Mar. 2001.
[16] P. K. Eng and D. Sarawate, "Spread-spectrum multiple-access performance of orthogonal codes: Impulsive noise," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 98–106, Jan. 1988.
[17] B. Aazhang and H. V. Poor, "Performance of DS/SSMA communications in impulsive channels—Part I: Linear correlation receivers," *IEEE Trans. Comm.*, vol. 35, pp. 1179–1187, Nov. 1987.
[18] ——, "Performance of DS/SSMA communications in impulsive channels—Part II: Hard-limiting correlation receivers," *IEEE Trans. Commun.*, vol. 36, pp. 88–96, Jan. 1988.
[19] H. V. Poor and M. Tanda, "Multiuser detection in impulsive channels," *Ann. Telecommun.*, vol. 54, pp. 392–400, July-Aug. 1999.
[20] ——, "Multiuser detection in flat fading non-Gaussian channels," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1769–1777, Nov. 2002.
[21] H. V. Poor, "Non-Gaussian signal processing problems in multiple-access communications," in *Proc. 1996 USC/CRASP Workshop on Non-Gaussian Signal Process.*, Ft. George Meade, MD, May 1996.
[22] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, 1985.
[23] P. J. Huber, *Robust Statistics*: Wiley, 1981, ch. New York.
[24] J. D. Gibson and M. L. Melsa, *Introduction to Non-Parametric Detection with Application*: Academic, 1975, ch. New York. (IEEE Press 1996).
[25] P. P. Kazakos and D. Kazakos, Eds., *Non-Parametric Methods in Communications*. New York: Marcel Decker, 1977.
[26] S. A. Kassam and J. B. Thomas, "Non-parametric detection, theory and application," in *Benchmark Papers in Electrical Engineering and Computer Science*: Academic, 1980, vol. 23, ch. New York.
[27] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer Verlag, 1994.
[28] R. Lupas and S. Verdú, "Linear multi-user detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 35, pp. 123–136, Jan. 1989.
[29] T. P. Hettmansperger and J. W. McKean, *Robust Non-Parametric Statistics*. New York: Wiley, 1998.
[30] T. P. Hettmansperger, *Statistical Inference Based on Ranks*. New York: Wiley, 1984.
[31] L. A. Jaeckel, "Estimating regression coefficients by minimizing the dispersion of residuals," *Ann. Math. Stat.*, vol. 43, pp. 1449–1458, 1972.
[32] E. L. Lehmann, *Nonparametric Statistical Methods Based on Ranks*. New York: Holden-Day, 1975.
[33] H. L. Koul, G. L. Sievers, and J. W. McKean, "An estimator of the scale parameter for the rank analysis of linear models under general score functions," *Scand. J. Stat.*, vol. 14, pp. 131–141, 1987.
[34] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 677–690, Mar. 1998.
[35] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 44, pp. 95–107, Jan. 1995.
[36] ——, "An extension of the PASTd algorithm to both rank and subspace tracking," *IEEE Signal Process. Lett.*, vol. 2, pp. 179–182, Sept. 1995.
[37] ——, "Asymptotic convergence analysis of projection approximation subspace tracking algorithms," *Signal Process.*, vol. 50, pp. 123–136, 1996.
[38] S. Attallah and K. Abed-Meraim, "Fast algorithms for subspace tracking," *IEEE Signal Process. Lett.*, vol. 8, no. 7, pp. 203–206, July 2001.
[39] Y. Wen, S. C. Chen, and K. L. Ho, "Robust subspace tracking in impulsive noise," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Finland, 2001, pp. 892–896.
[40] T. Gustafsson and C. S. MacInnes, "A class of subspace tracking algorithms based on approximation of noise-subspace," *IEEE Trans. Signal Process.*, vol. 48, no. 11, pp. 3231–3235, Nov. 2000.
[41] K. Abed-Meraim, A. Chekeif, and Y. Hua, "Fast orthonormal PAST algorithm," *IEEE Signal Process. Lett.*, vol. 7, no. 3, pp. 60–62, Mar. 2000.
[42] E. C. Real, D. W. Tufts, and J. W. Cooley, "Two algorithms for fast approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1936–1945, Jul. 1999.
[43] B. Seyfe and A. R. Sharafat, "Signed-rank non-parametric multiuser detection in non-Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1478–1486, Apr. 2005.

Babak Seyfe received the B.Sc. degree from the University of Tehran, Iran, and the M.Sc. and Ph.D. degrees both from Tarbiat Modarres University, Tehran, Iran, all in electrical engineering in 1991, 1995, and 2004, respectively.

He was with the Department of Electrical Engineering, Tarbiat Modarres University, Tehran, Iran, in 2004–2005 and a Visiting Researcher with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, in 2002. Currently, he is with the Centre for Digital Signal Processing Research, King's College, London, U.K. His research interests are detection and estimation theory, statistical signal processing, communication systems, and nonparametric and robust statistics.

Ahmad R Sharafat (S'75–M'80–SM'95) received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, and the M.Sc. and Ph.D. degrees from Stanford University, Stanford, CA, all in electrical engineering, in 1975, 1976, and 1980, respectively.

Currently, he is with the Department of Electrical Engineering, Tarbiat Modarres University, Tehran, Iran. His research interests are advanced signal processing, and communications systems, and networks.

Prof. Sharafat is a Senior Member Sigma Xi.