©2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

A Nonlinear M-Estimation Approach to Robust Asynchronous Multiuser Detection in Non-Gaussian Noise

Duc Son Pham, *Member, IEEE*, Abdelhak M. Zoubir, *Senior Member, IEEE*, Ramon F. Brcic, *Member, IEEE*, and Yee Hong Leung, *Member, IEEE*

***Abstract*—A nonlinear M-estimation approach is proposed to solve the multiuser detection problem in asynchronous code-division multiple-access (CDMA) systems where the ambient noise is impulsive and the delays are not known. We treat the unknown delays as nuisance parameters and the transmitted symbols as parameters of interest. We also analyze the asymptotic performance of the proposed estimator and propose suboptimal but computationally efficient procedures for solving the nonlinear optimization function. Simulation results show considerable improvements over the conventional approaches.**

***Index Terms*—Multiuser detection, nonlinear M-estimation, nuisance parameters, robust detection.**

I. Introduction

MULTIPLE-ACCESS techniques have been studied extensively over the last decade due to their improved system capacity over single access techniques in nonorthogonal signalling schemes. By exploiting the special structure of the system, the effects of both multiple-access interference (MAI) and ambient noise can be significantly mitigated, bringing the achievable performance bound close to that of the single access case.

In the context of up-link code-division-multiple-access (CDMA) multiuser detection, signal processing techniques have been limited mainly to the synchronous case [6], [21], [25], [28], [35]. The real issue here is that as the chip duration in later CDMA specifications is shortened [12], one can no longer ignore the asynchronous nature of the different users' received signals. The general asynchronous case is a very difficult problem and so attempts to solve subproblems have been made under certain assumptions, e.g., 1) the delays are assumed to be perfectly known [14], or at least chip synchronous [16], [34] and 2) the focus is to estimate the delays rather than signal detection, which is the ultimate goal [19], [20], [24]. Other suboptimum alternatives to the exponentially complex sequence maximum-likelihood (ML) estimator, which tackles the computational complexity issue, were presented in [8], [9], and [32].

Another limitation of many multiple-access techniques is the assumption of additive white Gaussian noise (AWGN). Practical measurements have indicated that in many cases both man-made and natural noise appear to be impulsive [1], [17], [18]. Non-Gaussian ambient noise is a major impairment to signal processing techniques that are based on a Gaussian assumption [29], [35]. Very little effort has been made in solving the problem of asynchronous multiuser detection in non-Gaussian noise. Recently, an approach based on robust statistics was formulated to tackle the non-Gaussian problem, and the method was extended to the asynchronous case, but the delays were assumed to be known *a priori* [29]. One can argue that a simple approach would be to first estimate the delays using standard techniques, for example [24], followed by robust demodulation of the received signal. However, there are two issues associated with this approach, namely, the requirement of very high signal-to-noise ratio (SNR) and a sufficiently large sample size, and the high sensitivity of the performance measures with respect to a model mis-specification. For practical sample sizes and low to moderate SNRs, as will be shown later, this approach may suffer from considerable performance loss.

We propose a nonlinear M-estimation approach to perform asynchronous CDMA multiuser detection in non-Gaussian impulsive noise. It can be shown that our approach is more general than the separable nonlinear least-squares method [15], and it allows for contamination in the noise density function. Notably, for the aforementioned multiuser detection problem, special computational procedures are proposed to solve the nonlinear M-estimation equations and very general asymptotic results are obtained.

The paper is organized as follows. In Section II, we present the signal model for asynchronous CDMA. Upon showing that the detection problem is a special nonlinear regression problem where the nuisance parameters lie in the design matrix, we formulate the optimization function under the framework of robust statistics in Section III. We then propose some modifications to the iteratively reweighted least-squares (IRLS) algorithm suggested in the robust statistics literature. In Section IV, we present some analysis of the estimators. Simulation results are given in Section V to illustrate the robustness and near–far resistance of the proposed approach. Section VI concludes the paper.

Manuscript received March 25, 2005; revised June 10, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Simon J. Godsill.

D. S. Pham was with the Communications and Signal Processing Group, Department of Electrical and Computer Engineering, Curtin University of Technology. He is now with the Institute for Multi-Sensor Processing and Content Analysis, Department of Computing, Curtin University of Technology, WA 6845, Perth, Australia (e-mail: dspham@ieee.org).

A. M. Zoubir and R. F. Brcic are with the Darmstadt University of Technology, D-64283 Darmstadt, Germany (e-mail: zoubir@ieee.org).

Y. H. Leung is with the Department of Electrical and Computer Engineering, Curtin University of Technology, WA 6845, Perth, Australia (e-mail: leungyh@ece.curtin.edu.au).

Digital Object Identifier 10.1109/TSP.2006.889471

1053-587X/$25.00 © 2007 IEEE

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

II. Signal Model

Consider a CDMA system with K users transmitting simultaneously. The received signal is the sum of the delayed versions of the transmitted signals and the channel noise

$$r(t) = \sum_{k=1}^{K}\sum_{l=0}^{\infty} A_k b_k(l) s_k(t - lT - d_k) + n(t) \tag{1}$$

where d_k presents the delay of the kth user with respect to the timing reference of the base station, $b_k(l)$ denotes the lth information bit, A_k denotes the amplitude of user k, and

$$s_k(t) = \sum_{j=0}^{M-1} c_j^k \beta(t - jT_c) \tag{2}$$

is the normalized signaling waveform, where $\{c_j^k \in \{+1, -1\}\}_{j=0}^{M-1}$ is an M-pseudorandom sequence, $T = MT_c$ is the period of the spreading sequence, T_c is the chip duration, and $\beta(t)$ is the pulse shaping filter impulse response. Without loss of generality, the following are assumed:

- $\beta(t)$ is approximately rectangular and confined within the chip duration;
- the maximum delay is less than the symbol duration, i.e., $\max_k d_k < T$.

The received signal is chip-matched filtered and sampled, which produces the following output at the jth chip and ith symbol:

$$\begin{aligned} r_j(i) &= \int_{iT+jT_c}^{iT+(j+1)T_c} r(t)\beta(t - iT - jT_c)dt \\ &= \sum_{k=1}^{K} q_j^k(i) + n_j(i) \end{aligned} \tag{3}$$

where $n_j(i) = \int_{iT+jT_c}^{iT+(j+1)T_c} \beta(t - iT - jT_c)n(t)\,dt, j = 0, \ldots, M-1$, are the noise samples that follow the two-term Gaussian mixture that is usually used to characterize impulsive noise [29]

$$f(x) = (1-\varepsilon)f_G(x;\nu^2) + \varepsilon f_G(x;\kappa\nu^2). \tag{4}$$

In (4), ε represents the amount of contamination, ν^2 is the variance of the Gaussian background noise, and κ represents the relative strength of the impulsive component. Note that this noise distribution is unknown. The contribution of the kth user in (3) is

$$\begin{aligned} q_j^k(i) = \int_{iT+jT_c}^{iT+(j+1)T_c} &\sum_{l=0}^{\infty} A_k b_k(l) c_j^k \\ &\times \beta(t - lT - jT_c - d_k)\beta(t - iT - jT_c)\,dt. \end{aligned} \tag{5}$$

Denote

$$p_k = \lfloor \frac{d_k}{T_c} \rfloor, \quad \eta_k = \frac{d_k}{T_c} - p_k \tag{6}$$

and the signature of the kth user $\mathbf{s}_k = (1/\sqrt{M})[c_0^k, \ldots, c_{M-1}^k]^T = [s_0^k, \ldots, s_{M-1}^k]^T$. By making use of the assumption on $\beta(t)$, it can be shown from Fig. 1 that the evaluation of (5) is, loosely speaking, proportional to the area under the signal waveform of the kth user during the jth chip-matched filtering duration. This is dependent of the relative timing between the arrival of the signal of the kth user and the base station as follows.

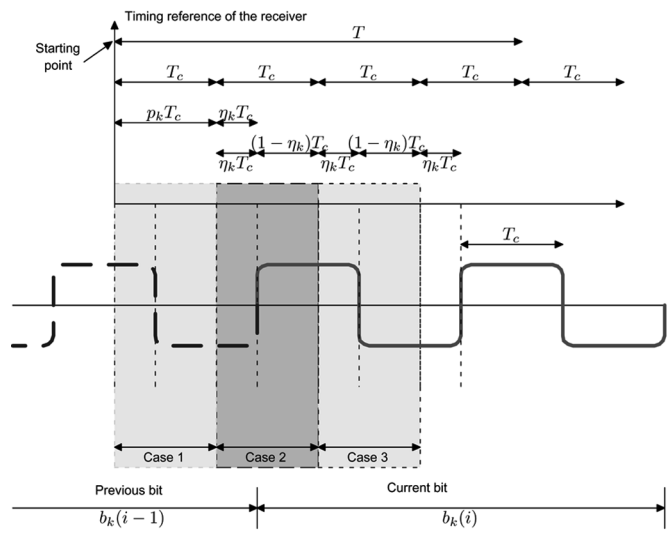

Fig. 1. Illustration of three cases in asynchronous CDMA detection.

- Case 1: When $j < p_k$, only the signal waveform of the previous symbol $b_k(i-1)$ is present.

$$\begin{aligned} q_j^k(i) = \eta_k A_k b_k(i-1) s_{M-p_k+j}^k \\ + (1-\eta_k) A_k b_k(i-1) s_{M-p_k+j+1}^k. \end{aligned} \tag{7}$$

- Case 2: When $p_k \le j < p_k + 1$, the transition from $b_k(i-1)$ to $b_k(i)$ is occurring.

$$q_j^k(i) = \eta_k A_k b_k(i-1) s_M^k + (1-\eta_k) A_k b_k(i) s_1^k. \tag{8}$$

- Case 3: When $p_k + 1 \le j$, only the signal waveform of the current symbol $b_k(i)$ is present.

$$q_j^k(i) = \eta_k A_k b_k(i) s_{j-p_k}^k + (1-\eta_k) A_k b_k(i) s_{j-p_k}^k. \tag{9}$$

Note that extensions to root-raised cosine pulses can be readily made with more mathematical complication provided the integral in (5) can be evaluated as a function of d_k. Denote $\mathbf{q}^k(i) = [q_0^k(i), \ldots, q_{M-1}^k(i)]^T$ the contribution of user k during the chip-matched filtering of the bit i as seen by the base station. Equations (7), (8), and (9) suggest the following representation:

$$\mathbf{q}^k(i) = \mathbf{h}_p^k A_k b_k(i-1) + \mathbf{h}_c^k A_k b_k(i) \tag{10}$$

where $\mathbf{h}_p^k$ denotes the coefficient vector due to *previous* information bit, and $\mathbf{h}_c^k$ denotes the coefficient vector due to *current* information bit and they are defined as follows [29]:

$$\begin{bmatrix} \mathbf{h}_c^k \\ \mathbf{h}_p^k \end{bmatrix} = \eta_k \begin{bmatrix} \mathbf{0}_{p_k+1} \\ \mathbf{s}_k \\ \mathbf{0}_{M-p_k-1} \end{bmatrix} + (1-\eta_k) \begin{bmatrix} \mathbf{0}_{p_k} \\ \mathbf{s}_k \\ \mathbf{0}_{M-p_k} \end{bmatrix} \tag{11}$$

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

where $\mathbf{0}_m$ denotes an m-size zero vector. From (3), we define the output vector for the ith bit as $\mathbf{r}(i) = [r_0(i), \ldots, r_{M-1}(i)]^T$. It then follows that

$$\mathbf{r}(i) = [\mathbf{H}_p \quad \mathbf{H}_c] \begin{bmatrix} \mathbf{A}\mathbf{b}(i-1) \\ \mathbf{A}\mathbf{b}(i) \end{bmatrix} + \mathbf{n}(i) \tag{12}$$

where $\mathbf{H}_p = [\mathbf{h}_p^1, \mathbf{h}_p^2, \ldots, \mathbf{h}_p^K], \mathbf{H}_c = [\mathbf{h}_c^1, \mathbf{h}_c^2, \ldots, \mathbf{h}_c^K]$, and $\mathbf{A} = \mathrm{diag}(A_1, \ldots, A_k)$, which is unknown. To maximize the information about the transmitted bits, it is appropriate to collect samples over m consecutive bits. Denote $\mathbf{y} = [\mathbf{r}(i)^T, \mathbf{r}(i+1)^T, \ldots, \mathbf{r}(i+m-1)^T]^T, \boldsymbol{\theta} = \boldsymbol{\mathcal{A}}\mathbf{b}_m(i)$, $\boldsymbol{\mathcal{A}} = \mathbf{I}_{m+1} \otimes \mathbf{A}$ is a $[K(m+1) \times K(m+1)]$ diagonal matrix, $\otimes$ denotes the Kronecker product, $\mathbf{b}(i) = [b_1(i), \ldots, b_K(i)]^T, \mathbf{b}_m(i) = [\mathbf{b}(i-1)^T, \mathbf{b}(i)^T, \ldots, \mathbf{b}(i+m-1)^T]^T, \mathbf{x} = [\mathbf{n}(i) \ldots \mathbf{n}(i+m-1)]^T$, and

$$\mathbf{H}(\mathbf{d}) = \begin{bmatrix} \mathbf{H}_p & \mathbf{H}_c & \ldots & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_p & \mathbf{H}_c & \ldots & \mathbf{0} \\ \vdots & \ldots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ldots & \ldots & \mathbf{H}_p & \mathbf{H}_c \end{bmatrix} \tag{13}$$

where $\mathbf{d} = [d_1, d_2, \ldots, d_K]^T$ is the vector of delays. The value of m must satisfy $m \geq \lfloor (K/M - K) \rfloor + 1$ so that $\mathbf{H}(\mathbf{d})$ will have full column rank [29]. In this paper, we select m to be the smallest integer that satisfies this criterion as to minimise the size of $\mathbf{H}(\mathbf{d})$. Hence, we obtain the signal model

$$\mathbf{y} = \mathbf{H}(\mathbf{d})\boldsymbol{\theta} + \mathbf{x}. \tag{14}$$

Given an estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, the estimate of the ith transmitted bits is

$$\hat{\mathrm{b}}(i) = \mathrm{sgn}(\hat{\boldsymbol{\theta}}[K+1:2K]). \tag{15}$$

For notational convenience, introduce $N = mM$, $P = (m+1)K$ and rewrite (14) as

$$y_n = \mathbf{h}_n^T(\mathbf{d})\boldsymbol{\theta} + x_n, \quad n = 1, \ldots, N \tag{16}$$

where $\mathbf{d} \in \mathbb{R}^K$, $\boldsymbol{\theta} \in \mathbb{R}^P, \mathbf{h}_n^T(\mathbf{d})$ is the nth row of $\mathbf{H}(\mathbf{d})$, and $x_i, i = 1, \ldots, N$ also follow the Gaussian mixture (4).

III. Multiuser Detection Techniques

A. Exhaustive Search

Since the noise density is unknown in the model (16), the maximum-likelihood (ML) solution is not available. A common approach to estimating $\boldsymbol{\theta}$ is to carry out an exhaustive search all over the parameter space of the quadratic cost function

$$(\hat{\boldsymbol{\theta}}, \hat{\mathbf{d}}) = \arg \min_{\substack{\mathbf{d} \in \mathbb{R}^K \\ \theta_p \in \{+\mathcal{A}_p, -\mathcal{A}_p\}}} \sum_{n=1}^{N} (y_n - \mathbf{h}_n^T(\mathbf{d})\boldsymbol{\theta})^2 \tag{17}$$

where θ_p and $\mathcal{A}_p$ denote the pth elements of $\boldsymbol{\theta}$ and $\boldsymbol{\mathcal{A}}$, respectively. There are several drawbacks to this approach that make it impractical: the prohibitive computational complexity of exponential order, lack of robustness in impulsive noise, and the strict assumption of perfect knowledge of the users' received signal amplitudes [24].

B. Simple Methods

Another common approach is to separate delay estimation from the signal detection problem. Typically, the preliminary estimates of the delays, denoted by $\hat{\mathbf{d}}^{(0)}$ are then obtained from, for example, MUSIC based on the sample covariance matrix [24]. Estimates of the transmitted bits are obtained by replacing the unknown design matrix $\mathbf{H}(\mathbf{d})$ with its estimate $\mathbf{H}(\hat{\mathbf{d}}^{(0)})$ and performing either linear detection, such as least-squares (LS) or minimum mean-square error (MMSE) techniques, or robust detection, which is described next. Even though this simplifies the estimation procedure, the performance of the estimator of $\boldsymbol{\theta}$ depends largely on the accuracy of the initial estimate $\hat{\boldsymbol{\theta}}^{(0)}$. Standard delay estimation techniques often assume sufficiently high SNR and a very large sample size, over which the sample covariance matrix is measured. In conditions such as low to moderate SNRs, time-varying delays, and especially in impulsive noise, the estimates from these techniques are poor. The critical issue then is the sensitivity of the final estimate of $\boldsymbol{\theta}$ with respect to the error of the design matrix that makes performance measures such as mean-squared error (MSE) or bit error rate (BER) worse. It is sensible then to consider joint estimation and detection under a robust framework to account for non-Gaussian noise. This motivates the approach described next.

C. Nonlinear M-Estimation

In robust statistics, instead of using the quadratic function as in (17), a less rapidly increasing penalty function $\rho(x)$ is used to down-weight outliers. With no nuisance parameters, i.e., when $\mathbf{d}$ is known, $\mathbf{h}_n(\mathbf{d}) = \mathbf{h}_n$, and the linear M-estimates are found by minimizing the following cost function [7]:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^P} \sum_{n=1}^{N} \rho\left(y_n - \mathbf{h}_n^T\boldsymbol{\theta}\right) \tag{18}$$

or, equivalently, by solving for $\boldsymbol{\theta}$ the equation

$$\sum_{n=1}^{N} \mathbf{h}_n \psi(y_n - \mathbf{h}_n^T\boldsymbol{\theta}) = \mathbf{0} \tag{19}$$

where $\psi(x) = \partial\rho(x)/\partial x$ is the location score function, which is generally odd symmetric, bounded and linear near the origin. The choice of $\psi(x)$ depends on the nature of the robust solution (see [4] for a comprehensive list). In this paper, we use Huber's minimax penalty function due to its simplicity and robustness. It is described by

$$\rho(x) = \begin{cases} \frac{x^2}{2\nu^2}, & \text{for } |x| \leq k_{\mathrm{H}}\nu^2 \\ k_{\mathrm{H}}|x| - \frac{k_{\mathrm{H}}^2\nu^2}{2}, & \text{for } |x| > k_{\mathrm{H}}\nu^2 \end{cases} \tag{20}$$

and its score function is given by

$$\psi(x) = \begin{cases} \frac{x}{\nu^2}, & \text{for } |x| \leq k_{\mathrm{H}}\nu^2 \\ k_{\mathrm{H}}\mathrm{sign}(x), & \text{for } |x| > k_{\mathrm{H}}\nu^2 \end{cases} \tag{21}$$

where k_{H} is dependent on ε and ν via

$$\frac{\Upsilon(k_{\mathrm{H}}\nu)}{k_{\mathrm{H}}\nu} - \Phi(k_{\mathrm{H}}\nu) = \frac{\varepsilon}{2(1-\varepsilon)} \tag{22}$$

in which $\Upsilon(x) = (1/\sqrt{2\pi})e^{-x^2/2}, \Phi(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-t^2/2}dt$ [7], [29], ε and ν are defined in (4). As ε is not known, it is possible in practice to use an approximate

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

version of the minimax as shown in [29, p. 298] for the value of k_{H} or nonparametric techniques [10, Ch. 5] to obtain ancillary estimates of the minimax parameters. Hereinafter, we assume that the minimax detector is fully specified.

To allow for nuisance parameters, we reformulate the nonlinear optimization problem as follows:

$$\begin{aligned}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{d}}) &= \arg \min_{\substack{\boldsymbol{\theta}\in\mathbb{R}^P \\ \mathbf{d}\in\mathbb{R}^K}} \sum_{n=1}^{N} \rho(y_n - \mathbf{h}_n^T(\mathbf{d})\boldsymbol{\theta}) \\ &= \arg \min_{\substack{\boldsymbol{\theta}\in\mathbb{R}^P \\ \mathbf{d}\in\mathbb{R}^K}} \mathcal{C}(\boldsymbol{\theta}, \mathbf{d}) \end{aligned} \tag{23}$$

where $\mathcal{C}(\boldsymbol{\theta}, \mathbf{d})$ will be referred to as the *minimax cost function*. We have also relaxed $\boldsymbol{\theta}$ in (23) for computational convenience in finding a suboptimal but robust solution. However, as the transmitted bits are estimated as in (15), this relaxation does not have any misleading effect on the final results. Continuing, we make the following assumptions.

Assumption 1: The delays $d_k, k = 1, \ldots, K$, must not fall on multiples of the chip durations, i.e., $0 < \eta_k < 1 \quad \forall k$.

Assumption 2: The delays $d_k, k = 1, \ldots, K$, are known up to p_k, i.e., the integer part, through some preliminary estimation procedure.

Assumption 1 avoids the technical difficulty and should not affect the results derived later. The delays can be as close as possible to any multiples of the chip durations. With Assumption 1, the first and second derivatives of the cost function can be evaluated, which will be useful for the asymptotic analysis presented subsequently. This assumption is reasonable, if we assume that the delays are uniformly distributed in $[0, T]$. With Assumption 2, we implicitly assume that an initial estimate $\hat{\mathbf{d}}^0$ is available from the acquisition phase as in the simple method. The difference here is that we use this initial estimate for joint parameter estimation and signal detection under a nonlinear M-estimation framework. Assumption 2 also means that we only need to update the fractional parts of $\hat{\mathbf{d}}$. This assumption is reasonable given the availability of standard delay estimation techniques, such as the MUSIC method presented in [24]. The motivations for joint parameter estimation and signal detection here are the scenarios of low to moderate SNRs, possibly fast time-varying delays, and non-Gaussian noise, which can significantly affect performance measures such as BER or MSE. These will be evident subsequently in the simulations results.

Denote $\mathcal{S} = \{\boldsymbol{\varphi} \in \mathbb{R}^K : \lfloor d_k \rfloor < \varphi_k < \lfloor d_k \rfloor + T_c,\ k = 1, 2, \ldots, K\}$ a subset of the K-dimension Euclidean space over which the vector of the delays is assumed to exist. With the above assumptions, we find the estimates from the equivalent optimization problem

$$\left(\hat{\boldsymbol{\theta}}, \hat{\mathbf{d}}\right) = \arg \min_{\substack{\boldsymbol{\theta}\in\mathbb{R}^P \\ \mathbf{d}\in\mathcal{S}}} \sum_{n=1}^{N} \rho\left(y_n - \mathbf{h}_n^T(\mathbf{d})\boldsymbol{\theta}\right). \tag{24}$$

When the penalty function $\rho(x)$ is not quadratic, which is the case here, (24) can be solved numerically. A number of nonlinear programming techniques are found in, for example, [13]. For fast convergence, we propose to use the IRLS algorithm. Briefly speaking, this algorithm sequences through a number of iterations at which the weights of a weighted least squares problem are updated according to the current values of the residuals. (For a historical perspective of the IRLS algorithm, please see [3].) Many variations of the IRLS algorithm exist with different updating schemes for the weights and the unknown parameters. In the signal processing context, the IRLS algorithm has been successfully applied in ℓ_p norm optimization applications [2], [11]. As the optimization (24) is not a ℓ_p norm problem, we will use the scheme frequently employed in robust statistics [7, Ch. 7], [23]. Under this scheme, the weights depend on the residuals and the score function (21) in such a way that the weighted least-squares cost function is the "best" approximation to the minimax cost function (23). Though (24) is, strictly speaking, a constrained optimization problem, it is of common practice to try the unconstrained solution first and only modify the search procedure if some constraints are violated. Thus, we shall describe the unconstrained IRLS algorithm then discuss the necessary modification when the constraint violation is found. However, we first need the following assumption before going into details of the IRLS algorithm.

Assumption 3: The solution $(\hat{\boldsymbol{\theta}},\ \hat{\mathbf{d}})$ to (24) exists and is the globally optimum solution of (23).

1) Unconstrained IRLS Procedure: Denote $\hat{\boldsymbol{\theta}}^{(i-1)}$ and $\hat{\mathbf{d}}^{(i-1)}$ the estimates at iteration $i-1$. The parameter updates of the IRLS algorithm at iteration i are generally given by

$$\hat{\boldsymbol{\theta}}^{(i)} = \hat{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\tau}^{(i)} \tag{25}$$

$$\hat{\mathbf{d}}^{(i)} = \hat{\mathbf{d}}^{(i-1)} + \boldsymbol{\delta}^{(i)} \tag{26}$$

where the search directions are found from minimizing an approximate local cost function about the point $(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)})$

$$\left(\boldsymbol{\tau}^{(i)}, \boldsymbol{\delta}^{(i)}\right) = \arg \min_{\substack{\boldsymbol{\tau}\in\mathbb{R}^P \\ \boldsymbol{\delta}\in\mathcal{D}}} \tilde{\mathcal{C}}^{(i-1)}\left(\hat{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\tau}, \hat{\mathbf{d}}^{(i-1)} + \boldsymbol{\delta}\right) \tag{27}$$

where the set $\mathcal{D} = \{\boldsymbol{\xi} \in \mathbb{R}^K : 0 < \xi_k < T_c, k = 1, 2, \ldots, K\}$, and the approximate local cost function is a quadratic function that has the same gradient as the minimax cost function at the point $(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)})$

$$\begin{aligned}&\tilde{\mathcal{C}}^{(i-1)}\left(\hat{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\tau}, \hat{\mathbf{d}}^{(i-1)} + \boldsymbol{\delta}\right) \\ &\quad = \sum_{n=1}^{N} \frac{\psi\left(z_n^{(i-1)}\right)}{z_n^{(i-1)}} \left(y_n - \mathbf{h}_n^T(\hat{\mathbf{d}}^{(i-1)} + \boldsymbol{\delta})\left(\hat{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\tau}\right)\right)^2\end{aligned} \tag{28}$$

in which the residuals are

$$z_n^{(i-1)} = y_n - \mathbf{h}_n^T\left(\hat{\mathbf{d}}^{(i-1)}\right)\hat{\boldsymbol{\theta}}^{(i-1)},\ n = 1, \ldots, N. \tag{29}$$

For notational convenience, we denote the weights at the ith iteration by

$$w(z_n^{(i)}) = \frac{\psi(z_n^{(i-1)})}{z_n^{(i-1)}}, \quad n = 1, \ldots, N. \tag{30}$$

Definition 1: Define the nth local model by a first-order approximation

$$\mathbf{h}_n\left(\hat{\mathbf{d}}^{(i-1)} + \boldsymbol{\delta}\right) \approx \mathbf{h}_n\left(\hat{\mathbf{d}}^{(i-1)}\right) + \mathbf{G}_n\left(\hat{\mathbf{d}}^{(i-1)}\right)\boldsymbol{\delta}, \quad n = 1, 2, \ldots, N \tag{31}$$

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

where

$$\mathbf{G}_n(\hat{\mathbf{d}}^{(i-1)}) = \frac{\partial \mathbf{h}_n(\hat{\mathbf{d}}^{(i-1)})}{\partial \mathbf{d}^T}. \tag{32}$$

Note that due to the definition of $\mathbf{h}_p^k$, $\mathbf{h}_c^k$ as in (11), $\mathbf{h}_n(\mathbf{d})$ is linear with respect to $\boldsymbol{\eta}$ where $\boldsymbol{\eta}$ is the vector of the fractional parts of $\mathbf{d}$ as defined in (6). Hence, $\mathbf{G}_n(\hat{\mathbf{d}})$ is fixed for $\hat{\mathbf{d}} \in \mathcal{S}$. With the linearization above, we can rewrite the approximate local cost function as

$$\tilde{\mathcal{C}}^{(i-1)}(\boldsymbol{\tau},\boldsymbol{\delta}) = \sum_{n=1}^{N} w\left(z_n^{(i-1)}\right)\left(y_n - \left(\mathbf{h}_n^T\left(\hat{\mathbf{d}}^{(i-1)}\right) + \left(\mathbf{G}_n\left(\hat{\mathbf{d}}^{(i-1)}\right)\boldsymbol{\delta}\right)^T\right)\left(\hat{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\tau}\right)\right)^2. \tag{33}$$

To simplify the notation, denote

$$\mathbf{h}_n^{(i-1)} = \mathbf{h}_n\left(\hat{\mathbf{d}}^{(i-1)}\right) \tag{34}$$

$$\mathbf{G}_n^{(i-1)} = \mathbf{G}_n\left(\hat{\mathbf{d}}^{(i-1)}\right) \tag{35}$$

$$\mathbf{H}^{(i-1)} = \mathbf{H}\left(\hat{\mathbf{d}}^{(i-1)}\right) \tag{36}$$

$$\mathbf{z}^{(i-1)} = \left[z_1^{(i-1)}, z_2^{(i-1)}, \ldots, z_N^{(i-1)}\right]^T \tag{37}$$

$$\mathbf{W}^{(i-1)} = \mathrm{diag}\left(w\left(z_1^{(i-1)}\right), w\left(z_2^{(i-1)}\right), \ldots, w\left(z_N^{(i-1)}\right)\right). \tag{38}$$

The approximate local cost function can be then written as

$$\tilde{\mathcal{C}}^{(i-1)}(\boldsymbol{\tau},\boldsymbol{\delta}) \approx \sum_{n=1}^{N} w\left(z_n^{(i-1)}\right)\left(y_n - \mathbf{h}_n^{T(i-1)}\hat{\boldsymbol{\theta}}^{(i-1)} - \mathbf{h}_n^{T(i-1)}\boldsymbol{\tau} - \hat{\boldsymbol{\theta}}^{T(i-1)}\mathbf{G}_n^{(i-1)}\boldsymbol{\delta}\right)^2 \tag{39}$$

$$\approx \sum_{n=1}^{N} w\left(z_n^{(i-1)}\right)\left(z_n^{(i-1)} - \mathbf{h}_n^{T(i-1)}\boldsymbol{\tau} - \hat{\boldsymbol{\theta}}^{T(i-1)}\mathbf{G}_n^{(i-1)}\boldsymbol{\delta}\right)^2 \tag{40}$$

where we have ignored the term $\boldsymbol{\tau}^T\mathbf{G}_n^{(i-1)}\boldsymbol{\delta}$ as $\boldsymbol{\tau}$ and $\boldsymbol{\delta}$ are assumed to be small. Now, introduce

$$\mathbf{P}^{(i-1)} = \begin{bmatrix} \hat{\boldsymbol{\theta}}^{T(i-1)}\mathbf{G}_1^{(i-1)} \\ \hat{\boldsymbol{\theta}}^{T(i-1)}\mathbf{G}_2^{(i-1)} \\ \vdots \\ \hat{\boldsymbol{\theta}}^{T(i-1)}\mathbf{G}_N^{(i-1)} \end{bmatrix} \tag{41}$$

and to further simplify the notation, we shall drop the superscript $^{(i-1)}$ where applicable. The approximate cost function at iteration i can be expressed in a more compact form

$$\tilde{\mathcal{C}}(\boldsymbol{\tau},\boldsymbol{\delta}) \approx (\mathbf{z} - \mathbf{H}\boldsymbol{\tau} - \mathbf{P}\boldsymbol{\delta})^T\mathbf{W}(\mathbf{z} - \mathbf{H}\boldsymbol{\tau} - \mathbf{P}\boldsymbol{\delta}). \tag{42}$$

It is then straightforward to show that if we assume $\mathbf{D} = [\mathbf{H}\ \mathbf{P}]$ is of full column rank, then the solution of (42) is found at

$$\begin{bmatrix} \boldsymbol{\tau} \\ \boldsymbol{\delta} \end{bmatrix} = \left(\mathbf{D}^T\mathbf{W}\mathbf{D}\right)^{-1}\mathbf{D}^T\mathbf{W}\mathbf{z} \tag{43}$$

or, equivalently, by using the matrix inversion lemma

$$\boldsymbol{\delta} = (\mathbf{P}^T\mathbf{W}\mathbf{P} - \mathbf{P}^T\mathbf{W}\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{P})^{-1} \times (\mathbf{P}^T\mathbf{W} - \mathbf{P}^T\mathbf{W}\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W})\mathbf{z}, \tag{44}$$

$$\boldsymbol{\tau} = (\mathbf{H}^T\mathbf{W}\mathbf{H} - \mathbf{H}^T\mathbf{W}\mathbf{P}(\mathbf{P}^T\mathbf{W}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{W}\mathbf{H})^{-1} (\mathbf{H}^T\mathbf{W} - \mathbf{H}^T\mathbf{W}\mathbf{P}(\mathbf{P}^T\mathbf{W}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{W})\mathbf{z}. \tag{45}$$

To further simplify (44) and (45), we start from the assumption that the diagonal elements of $\mathbf{W}$ are nonnegative because of the monotonic nature of the score function so that it is possible to introduce

$$\mathbf{V} = \mathbf{W}^{1/2}. \tag{46}$$

Furthermore, denote

$$\tilde{\mathbf{P}} = \mathbf{V}\mathbf{P} \tag{47}$$

$$\tilde{\mathbf{H}} = \mathbf{V}\mathbf{H} \tag{48}$$

so that

$$\mathbf{P}^T\mathbf{W}\mathbf{P} - \mathbf{P}^T\mathbf{W}\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{P} = \tilde{\mathbf{P}}\mathbf{P}_{\tilde{\mathbf{H}}}^{\perp}\mathbf{P} \tag{49}$$

where $\mathbf{P}_{\tilde{\mathbf{H}}}^{\perp} = \mathbf{I} - \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}$ denotes the orthogonal projection onto the subspace of $\tilde{\mathbf{H}}$. Using a similar argument for the other terms, it is easy to show that (44) and (45) are equivalent to

$$\boldsymbol{\delta} = \left(\tilde{\mathbf{P}}^T\mathbf{P}_{\tilde{\mathbf{H}}}^{\perp}\tilde{\mathbf{P}}\right)^{-1}\tilde{\mathbf{P}}\mathbf{P}_{\tilde{\mathbf{H}}}^{\perp}\mathbf{V}\mathbf{z} \tag{50}$$

$$\boldsymbol{\tau} = \left(\tilde{\mathbf{H}}^T\mathbf{P}_{\tilde{\mathbf{P}}}^{\perp}\tilde{\mathbf{H}}\right)^{-1}\tilde{\mathbf{H}}\mathbf{P}_{\tilde{\mathbf{P}}}^{\perp}\mathbf{V}\mathbf{z}. \tag{51}$$

We have the following result regarding the convergence behaviour of the extended IRLS algorithm.

Proposition 1: With the extended IRLS update scheme as described above and the fact that $\psi(x)/x$ is monotonically decreasing for $x > 0$, the minimax cost function decreases at a rate that

$$\mathcal{C}(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)}) - \mathcal{C}(\hat{\boldsymbol{\theta}}^{(i)}, \hat{\mathbf{d}}^{(i)}) \geq \frac{1}{2}\mathbf{z}^{(i-1)}\mathbf{W}^{(i-1)}\mathbf{D}^{(i-1)}\left(\mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}\mathbf{D}^{(i-1)}\right)^{-1} \times \mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}\mathbf{z}^{(i-1)}. \tag{52}$$

The proof of this result can be found in Appendix A.

2) Practical Considerations: In the preceding derivations, we have made relaxation when solving partially constrained optimization problem (24) because of Assumptions 1, 2, and 3. Even though it may happen in practice that during the convergence to the global solution, the searching path may fall outside the constrained set $\mathcal{S}$. In practice, one should always try an unconstrained IRLS procedure. However, if one or more constraints are violated, necessary modifications are needed to remedy this problem and ensure that the global cost function is always decreased. In what follows, we follow standard nonlinear programming techniques [13] to derive a procedure for doing so.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

Definition 2: Define the *active* set $\mathcal{Q}_a(\hat{\mathbf{d}})$ and its complementary (*inactive*) set $\mathcal{Q}_i(\hat{\mathbf{d}})$ as follows:

$$\mathcal{Q}_a(\hat{\mathbf{d}}) = \{k : \hat{d}_k = d_k + \epsilon \text{ or } \hat{d}_k = d_k + T_c - \epsilon\} \quad (53)$$
$$\mathcal{Q}_i(\hat{\mathbf{d}}) = \{k : d_k + \epsilon < \hat{d}_k < d_k + T_c - \epsilon\} \quad (54)$$

where ϵ is arbitrarily small, i.e., $0 < \epsilon/T_c \ll 1$.

The active set refers to the delay estimates that are deemed to be on the boundary whilst the inactive set refers to the delay estimates that are well within the constraints, i.e., $\mathbf{d} \in \mathcal{S}$, which can be *practically* rewritten as follows:

$$\mathbf{c}(\boldsymbol{\theta}, \mathbf{d}) = \begin{bmatrix} \mathbf{I}_{K\times K} \\ -\mathbf{I}_{K\times K} \end{bmatrix} \mathbf{d} - \begin{bmatrix} \mathbf{d}^u \\ -\mathbf{d}^l \end{bmatrix} \leq \mathbf{0}_{2K} \quad (55)$$

where $\mathbf{d}^u = [\lfloor d_1 \rfloor + T_c - \epsilon, \ldots, \lfloor d_K \rfloor + T_c - \epsilon]^T$ and $\mathbf{d}^l = [\lfloor d_1 \rfloor + \epsilon, \ldots, \lfloor d_K \rfloor + \epsilon]^T$. It is noted that if in (55) the equality occurs for the kth constraint $c_k(\boldsymbol{\theta}, \mathbf{d})$, then it will not happen for the $(k+K)$th constraint $c_{k+K}(\boldsymbol{\theta}, \mathbf{d})$ and *vice versa*, where $k \leq K$. Denote $\mathcal{V}(\hat{\mathbf{d}})$ the subset of row indexes of constraints $\mathbf{c}$ where the equalities occur. Suppose at iteration $i-1$ the active set is nonempty. The solution at iteration $i-1$ should satisfy the Kuhn–Tucker conditions [13, p. 328]

$$\nabla \tilde{C}^{(i-1)}(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)}) + \sum_{k \in \mathcal{V}(\hat{\mathbf{d}}^{(i-1)})} \sigma_k \nabla c_k \left(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)}\right) = \mathbf{0} \quad (56)$$

where $c_k(\boldsymbol{\theta}, \mathbf{d})$ is the kth row of $\mathbf{c}(\boldsymbol{\theta}, \mathbf{d})$ in (55) and ∇ is the gradient with respect to $(\boldsymbol{\tau}, \boldsymbol{\delta})$ and evaluated at $\boldsymbol{\tau} = \mathbf{0}, \boldsymbol{\delta} = \mathbf{0}$, i.e.,

$$\nabla \tilde{C}^{(i-1)} \left(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)}\right) = -\mathbf{D}^{T(i-1)} \mathbf{W}^{(i-1)} \mathbf{z}^{(i-1)}. \quad (57)$$

If any Lagrangian multiplier $\sigma_k < 0$ in (56), we can relax the constraint k from the active set $\mathcal{Q}_a(\hat{\mathbf{d}}^{(i-1)})$. Now introduce $\bar{\mathbf{P}}^{(i-1)}$ consisting of the columns of $\mathbf{P}^{(i-1)}$ that correspond to the inactive set, i.e., $\bar{\mathbf{P}}^{(i-1)}$ is obtained similarly as in (41) but the derivatives of $\mathbf{G}_n$ as in (32) are only with respect to elements of $\mathbf{d}$ that correspond to the inactive set. This implies that we will only compute subdirections $\bar{\boldsymbol{\delta}}^{(i)}$ as in (44) with $\mathbf{P}^{(i-1)}$ being replaced with $\bar{\mathbf{P}}^{(i-1)}$. Define the selection matrix $\mathbf{U}^{(i)}$ consisting of the columns of the identity matrix of size $K \times K$ whose column indexes are found in the inactive set $\mathcal{Q}(\hat{\mathbf{d}}^{(i-1)})$. The search directions for $\hat{\mathbf{d}}$ are

$$\boldsymbol{\delta}^{(i)} = \mathbf{U}^{(i)} \bar{\boldsymbol{\delta}}^{(i)}. \quad (58)$$

Furthermore, instead of using (25) and (26), we introduce the scale parameter $\lambda \in [0, 1]$ for the search directions and propose the updates as follows:

$$\hat{\boldsymbol{\theta}}^{(i)} = \hat{\boldsymbol{\theta}}^{(i-1)} + \lambda^{(i)} \boldsymbol{\tau}^{(i)} \quad (59)$$
$$\hat{\mathbf{d}}^{(i)} = \hat{\mathbf{d}}^{(i-1)} + \lambda^{(i)} \boldsymbol{\delta}^{(i)}. \quad (60)$$

To find the optimal value for the scale parameter, denoted by $\lambda_{\min}^{(i)}$, we follow the approach in [22, p. 384]. First, define

$$\lambda_1^{(i)} = \sup_\lambda \left\{ \left(\hat{\mathbf{d}}^{(i-1)} + \frac{\lambda}{\mu} \boldsymbol{\delta}^{(i)} \right) \in \mathcal{S} \right\}. \quad (61)$$

and

$$g^{(i)}(\lambda) = \mathcal{C} \left(\hat{\boldsymbol{\theta}}^{(i-1)} + \frac{\lambda}{\mu} \boldsymbol{\tau}^{(i)}, \hat{\mathbf{d}}^{(i-1)} + \frac{\lambda}{\mu} \boldsymbol{\delta}^{(i)} \right) \quad (62)$$

where $\mathcal{C}(\boldsymbol{\theta}, \mathbf{d})$ is the minimax cost function as defined in (23). By using a quadratic regression over this minimax cost function, the value of $\lambda_{\min}^{(i)}$ is found from

$$\lambda_{\min}^{(i)} = \arg\min_\lambda \left(g^{(i)}\left(\lambda_1^{(i)}\right) - g^{(i)}\left(\lambda_0^{(i)}\right) - \frac{\partial g^{(i)}\left(\lambda_0^{(i)}\right)}{\partial \lambda} \right) \lambda^2 + \frac{\partial g^{(i)}\left(\lambda_0^{(i)}\right)}{\partial \lambda} \lambda + g^{(i)}\left(\lambda_0^{(i)}\right) \quad (63)$$

where $\lambda_0^{(i)} = 0$. Now, it can be seen that with this approach, the convergence property of either Lemma 1 can still be achieved with $\mathbf{P}^{(i-1)}$ being replaced with $\bar{\mathbf{P}}^{(i-1)}$ and $\mathbf{z}^{(i-1)}$ being replaced with $\lambda_{\min}^{(i)} \mathbf{z}^{(i-1)}$. It is found in practice that with a reasonable level of numerical accuracy, the proposed algorithm converges within a few iterations.

IV. Asymptotic Analysis

In this section, we study the asymptotic behaviour of the nonlinear M-estimates for asynchronous CDMA multiuser detection in non-Gaussian noise. By "asymptotic" we mean when the code length becomes large, i.e., $M \to \infty$ or, equivalently, $N \to \infty$. We shall only focus on consistency of the estimate as a full treatment is beyond the scope of this paper.

Recall that for the minimax solution used in this work, $\rho(x)$ is convex, and $\psi(x)$ is odd symmetric and bounded. Strictly speaking, the derivative of $\psi(x)$ does not exist at the corner points. However, this can be relaxed in practice by introducing a smooth version as shown in [29]. Hence, without loss of generality, we assume that $\psi(x)$ is differentiable over the real line.

Proposition 2: For asynchronous CDMA multiuser detection as described by Equation (16), there exists no strongly consistent estimate under the nonlinear M-estimation approach.

Proof: Note that this result also applies to maximum likelihood and least squares estimates of (16). We will prove this result by assuming the converse is true, that the nonlinear M-estimates is strongly consistent. Then, as a direct consequence of the results in [27], [31], [33], we would require that the minimum eigenvalue of $\mathbf{D}^T\mathbf{D}$ approaches infinity

$$\lim_{N\to\infty} \gamma_{\min}(\mathbf{D}^T\mathbf{D}) \to \infty \quad (64)$$

or, more conveniently, as a special case

$$\lim_{N\to\infty} \gamma_{\min}(\mathbf{H}^T\mathbf{H}) \to \infty \quad (65)$$

where both $\mathbf{H}$ and $\mathbf{D}$ *are evaluated at the true parameters*, and $\gamma_{\min}$ denotes the minimum eigenvalue. Note that following the arguments in [31] this form is less strict than the form used in [15]. Hence, it suffices to show that (65) is not met by practical systems. It follows from a standard result on the norm of a square matrix (see, e.g., [5, p. 316]) that if a square matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is positive definite, then

$$\gamma_{\min}(\mathbf{A}) \leq \frac{1}{m} \|\mathbf{A}\|^2 \quad (66)$$

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

where $||\cdot||$ denotes the Frobenius norm. Using this result, it remains to show that the norm of $\mathbf{H}^T\mathbf{H}$, or, equivalently, the norm of $\mathbf{H}$, is bounded.

For CDMA systems, the code matrix is designed such that $\mathbf{s}_k = (1/\sqrt{M})[c_1^k, c_2^k, \ldots, c_M^k]^T$ to satisfy the energy requirement $\mathbf{s}_k^T\mathbf{s}_k = 1$, where $\{c_j^k \in \{+1, -1\}\}_{j=1}^M$. The code cross-correlation $\mathbf{s}_k^T\mathbf{s}_l \in (-1, 1)$ and $\lim_{M\to\infty} \mathbf{s}_k^T\mathbf{s}_l \to \varrho_{kl}, \forall k \neq l$. It follows that the norm of $\mathbf{S}^T\mathbf{S}$ and hence $\gamma_{\min}(\mathbf{S}^T\mathbf{S})$ is bounded. From the definition of $\mathbf{H}$ in (13), the definitions of $\mathbf{H}_p$ and $\mathbf{H}_c$ after (12), and the definitions of $\mathbf{h}_c^k$ and $\mathbf{h}_p^k$ in (11), it follows that the norm of $\mathbf{H}$ is also bounded. ■

Although strong consistency cannot be asserted for the problem of interest, weak consistency can be asserted under special circumstances. Following [26], [30], and [31], we require the following assumptions.

Assumption 4: For notational convenience, hereinafter we denote the true parameters by $\boldsymbol{\omega}_0 = (\boldsymbol{\theta}_0, \mathbf{d}_0)$ and the estimates by $\hat{\boldsymbol{\omega}}_N = (\hat{\boldsymbol{\theta}}_N, \hat{\mathbf{d}}_N)$. There exists a compact set $\mathcal{W} \subset \mathbb{R}^{P+Q}$ and N_0 such that $\forall N > N_0$ and $\boldsymbol{\omega} \in \mathcal{W}$, the eigenvalues of $\mathbf{D}^T\mathbf{D}$ are positive.

The next assumption concerns the noise distribution and the choice of the penalty function.

Assumption 5: The penalty function $\rho(x)$ is symmetric, convex, and continuous on the real line. The zero mean non-Gaussian noise has finite variance σ^2. Denote by $F_N(\mathbf{y}, \boldsymbol{\omega})$ the multivariate cumulative distribution function For any $\boldsymbol{\omega} = (\boldsymbol{\theta}, \mathbf{d}) \in \mathcal{W}, \int_{-\infty}^{+\infty} \mathcal{C}_N(\boldsymbol{\omega}) dF_N(\mathbf{y}, \boldsymbol{\omega}) = \int_{-\infty}^{+\infty} \sum_{n=1}^N \rho(y_n - \mathbf{h}_n^T(\mathbf{d})\boldsymbol{\theta}) dF_N(\mathbf{y}, \boldsymbol{\omega})$ is finite.

Then, we have the following result.

Proposition 3: Suppose that $\forall \Delta > 0$ and $\boldsymbol{\omega} \in \mathcal{W}$

$$\lim_{N\to\infty} \inf_{||\boldsymbol{\omega}-\boldsymbol{\omega}_0||\geq\Delta} \{\mathcal{C}_N(\boldsymbol{\omega}) - \mathcal{C}_N(\boldsymbol{\omega}_0)\} > 0 \tag{67}$$

in probability, then under Assumptions 4 and 5, $\hat{\omega}_N \xrightarrow{P} \boldsymbol{\omega}_0$, meaning $\hat{\boldsymbol{\theta}}_N \xrightarrow{P} \boldsymbol{\theta}_0$, where $\xrightarrow{P}$ denotes convergence in probability and $||\cdot||$ denotes the Euclidean norm.

Proof: The proof of this result follows the same approach in [31]. Details are given here for completeness. First of all, recall that by definition it is $\hat{\boldsymbol{\omega}}_N$ that minimizes the cost function

$$\mathcal{C}_N(\boldsymbol{\omega}) = \sum_{n=1}^N \rho\left(y_n - \mathbf{h}_n^T(\mathbf{d})\boldsymbol{\theta}\right) \tag{68}$$

not $\boldsymbol{\omega}_0$. This means

$$\mathcal{C}_N(\hat{\boldsymbol{\omega}}_N) - \mathcal{C}_N(\boldsymbol{\omega}_0) < 0. \tag{69}$$

The assumption in the proposition implies that $\forall \Delta > 0$

$$\lim_{N\to\infty} \Pr\left\{\inf_{||\boldsymbol{\omega}-\boldsymbol{\omega}_0||\geq\Delta} (\mathcal{C}_N(\boldsymbol{\omega}) - \mathcal{C}_N(\boldsymbol{\omega}_0)) < 0\right\} = 0. \tag{70}$$

Now suppose the converse is true, that

$$\lim_{N\to\infty} \Pr\{||\hat{\boldsymbol{\omega}}_N - \boldsymbol{\omega}_0|| < \Delta\} = 1 \tag{71}$$

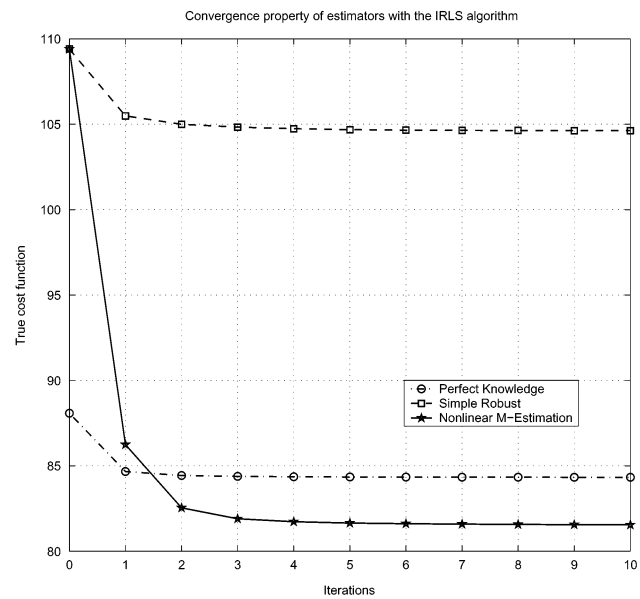

Fig. 2. Typical average cost function in non-Gaussian noise.

is not correct for some $\Delta > 0$. Equivalently

$$\lim_{N\to\infty} \Pr\{||\hat{\boldsymbol{\omega}}_N - \boldsymbol{\omega}_0|| \geq \Delta\} > 0. \tag{72}$$

Consequently, by the definition of $\hat{\boldsymbol{\omega}}_N$, that is to minimize $\mathcal{C}_N(\boldsymbol{\omega})$, it follows that

$$\lim_{N\to\infty} \Pr\{\inf_{||\boldsymbol{\omega}-\boldsymbol{\omega}_0||\geq\Delta} (\mathcal{C}_N(\boldsymbol{\omega}) - \mathcal{C}_N(\boldsymbol{\omega}_0)) < 0\} > 0 \tag{73}$$

which contradicts (70). Hence, the proof follows. ■

V. Simulation Results

In this section, we present some simulation results to assess the proposed approach in solving the asynchronous CDMA multiuser detection problem in non-Gaussian noise. The noise parameter settings for non-Gaussian noise described in (4) are $\varepsilon = 0.1, \kappa = 100$ and the total noise variance is $\sigma^2 = (1-\varepsilon)\nu^2 + \kappa\nu^2$. We consider a CDMA system with $K = 4$ users and a code length $M = 31$. The delays are randomly generated $\mathbf{d}^T = [4.89T_c\ 23.24T_c\ 3.64T_c\ 24.23T_c]$. User 1 is assumed to be the user of interest. We will compare the proposed approach using the extended IRLS algorithm (Section III-C) with the following:

- the method with perfect knowledge of the delays followed by robust detection (Section III-A);
- the simple method (Section III-B) followed by linear LS detection;
- the simple method (Section III-B) followed by robust detection.

In the first simulation, we compare the proposed numerical procedures for solving the nonlinear M-estimation problem. In the first simulation, we measure the value of the global cost function at each iteration to assess convergence using the IRLS algorithm. A typical example for non-Gaussian noise is shown in Fig. 2. As can be seen, 99% convergence can be reached within ten iterations. It is noted that since the true parameters

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

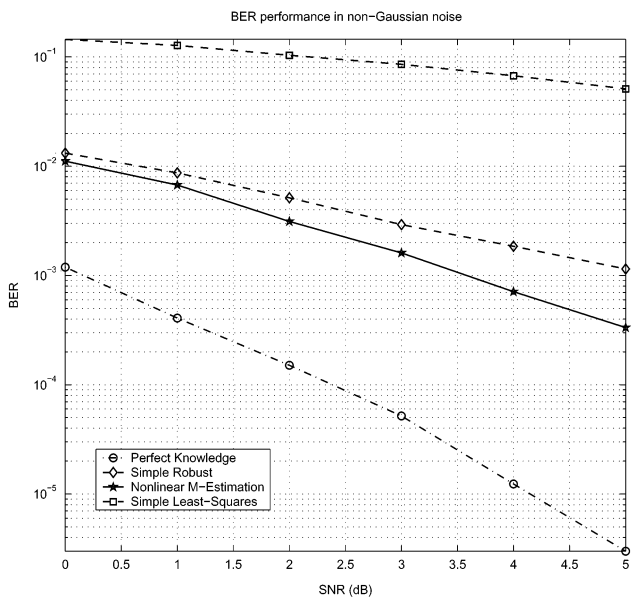

Fig. 3. BER performance in non-Gaussian noise, $\varepsilon = 0.1, \kappa = 100$.

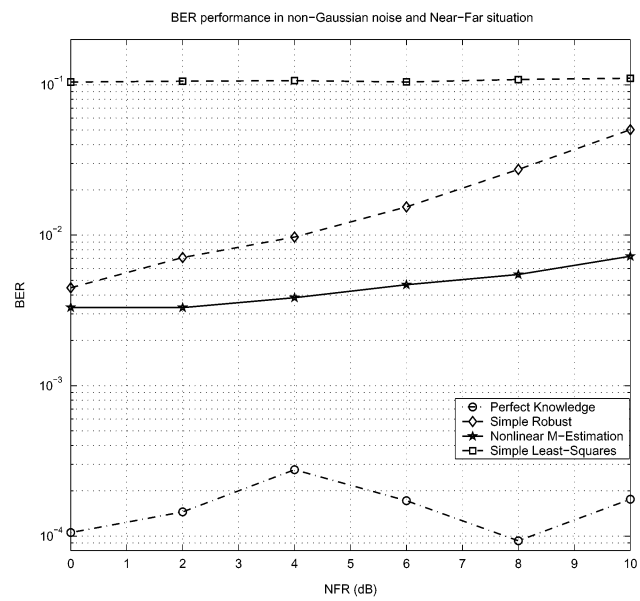

Fig. 4. BER performance in non-Gaussian noise, $\varepsilon = 0.1, \kappa = 100$, and near–far situation.

are not the solution of the finite sample optimization problem, the curve for the method with perfect knowledge is above the proposed nonlinear M-estimation curves.

In the second simulation, we assess the BER performance for non-Gaussian noise. All users have the same power, and the SNR ranges from 0 to 5 dB. Results are given in Fig. 3. Again, the two numerical procedures yield approximately the same BER performance. Compared with the simple method followed by robust detection, a performance gain of 8 dB can be achieved at SNR = 5 dB, and this increases with the SNR. The simple method followed by linear LS detection obviously suffers from a large performance loss.

In the third simulation, we look at the BER performance when the other users have stronger power than user 1. This is known as the *near–far* problem in CDMA systems. We keep the SNR = 2 dB and vary the near–far ratio (NFR), that is the power ratio of any other user compared to user 1, from 0 to 10 dB. The simulation is conducted using non-Gaussian noise. The result is given in Fig. 4. It is observed that the simple method followed by robust detection is more sensitive to the near–far problem, while the proposed method is more near–far resistant. For the simple method followed by linear LS detection, non-Gaussianity is the dominant problem; hence, it exhibits less sensitivity to the near–far problem within the simulation settings.

Finally, we investigate the case where the noise is Gaussian. All users have the same power. The result is given in Fig. 5. It is observed that the proposed method and the simple methods achieve similar BER performance.

In all simulations, we have not included the maximum likelihood estimate, i.e., when $\rho(x)$ is replaced with $-\log f(x)$. The reason is that the nonmonotonic nature of the score function $-f'(x)/f(x)$ leads to numerical accuracy problems. However, as noted from the remarks in previous section, in terms of asymptotic MSE, this can be directly inferred by replacing $\psi(x)$ with $-f'(x)/f(x)$. In practice, where some prior knowledge of the delays might be available and if a strictly constrained optimization problem can be set up, it is possible to get closer to the perfect knowledge curve.

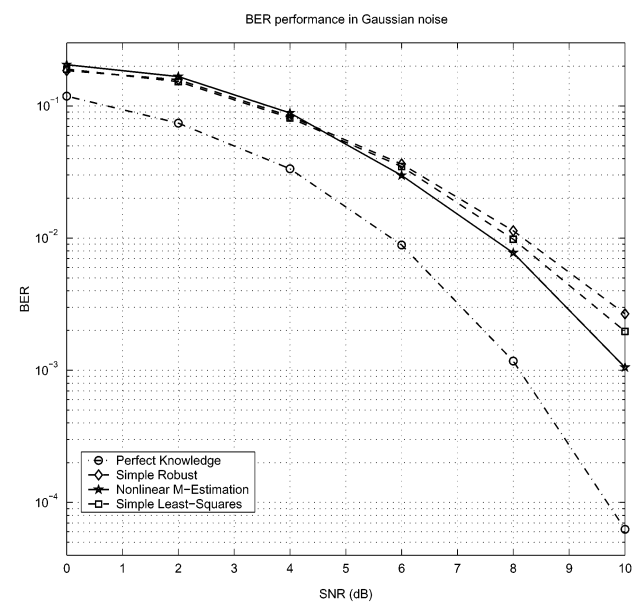

Fig. 5. BER performance in Gaussian noise.

VI. Conclusion

A novel nonlinear M-estimation approach has been proposed to solve the asynchronous CDMA multiuser detection problem in non-Gaussian impulsive noise. Even though we only considered the simple case to simplify the derivation, extensions to more realistic situations, including fading and multipath channels are possible. The proposed approach is developed on the reasonable assumption that some preliminary estimates of the delays (nuisance parameters) are available through the initial acquisition phase using standard delay estimation techniques.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

This approach is an extension to linear M-estimation and is suitable for detection in non-Gaussian noise. In contrast to many previous works, we have paid more attention to computational procedures for solving the nonlinear optimization problem. Extensions to the IRLS and MR algorithms for linear M-estimation have been presented and shown to achieve convergence. Simulations demonstrate considerable performance gain can be achieved over conventional methods. We have also conducted a brief asymptotic study and established weak consistency and the asymptotic covariance matrix of the estimator. The proposed nonlinear M-estimator can be applied to practical applications other than multiuser detection, for example array signal processing, where a similar nonlinear M-estimation problem can be formulated.

Appendix A
Proof of Proposition 1

Proof: Introduce $\mathbf{\Delta}^{(i)} = \left[\begin{smallmatrix}\boldsymbol{\tau}^{(i)}\\ \boldsymbol{\delta}^{(i)}\end{smallmatrix}\right]$. We rewrite the optimization problem (33) as

$$\hat{\mathbf{\Delta}}^{(i)} = \arg\min_{\mathbf{\Delta}} \left(\mathbf{z}^{(i-1)} - \mathbf{D}^{(i-1)}\mathbf{\Delta}\right)^T \times \mathbf{W}^{(i-1)}(\mathbf{z}^{(i-1)} - \mathbf{D}^{(i-1)}\mathbf{\Delta}) \tag{74}$$

$$= \arg\min_{\mathbf{\Delta}} \mathbf{u}^{(i-1)}(\mathbf{\Delta})^T\mathbf{W}^{(i-1)}\mathbf{u}^{(i-1)}(\mathbf{\Delta}). \tag{75}$$

Following the approach in [7], we study the function

$$\mathcal{I}(\mathbf{\Delta}) = \frac{1}{2}\mathbf{u}^{T(i-1)}(\mathbf{\Delta})\mathbf{W}^{(i-1)}\mathbf{u}^{(i-1)}(\mathbf{\Delta}) - \sum_{n=1}^{N}\rho\left(u_n^{(i-1)}(\mathbf{\Delta})\right) + \sum_{n=1}^{N}\rho\left(z_n^{(i-1)}\right) - \frac{1}{2}\mathbf{z}^{(i-1)}\mathbf{W}^{(i-1)}\mathbf{z}^{(i-1)}. \tag{76}$$

It can be verified that

$$\mathcal{I}(\mathbf{0}) = 0. \tag{77}$$

Denote by $\mathbf{d}_n^{T(i-1)}$ the nth row of $\mathbf{D}^{(i-1)}$. We have

$$\frac{\partial\mathcal{I}(\mathbf{\Delta})}{\partial\mathbf{\Delta}} = -\mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}(\mathbf{z}^{(i-1)} - \mathbf{D}^{(i-1)}\mathbf{\Delta}) + \sum_{n=1}^{N}\mathbf{d}_n^{(i-1)}\psi\left(z_n^{(i-1)} - \mathbf{d}_n^{T(i-1)}\mathbf{\Delta}\right). \tag{78}$$

Due to the definition of $\mathbf{W}^{(i-1)}$ as in (38), it can easily be verified that

$$\left.\frac{\partial\mathcal{I}(\mathbf{\Delta})}{\partial\mathbf{\Delta}}\right|_{\mathbf{\Delta}=\mathbf{0}} = \mathbf{0}. \tag{79}$$

Continuing, we have the second derivatives

$$\frac{\partial^2\mathcal{I}(\mathbf{\Delta})}{\partial\mathbf{\Delta}\partial\mathbf{\Delta}^T} = \mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}\mathbf{D}^{(i-1)} - \sum_{n=1}^{N}\mathbf{d}_n^{(i-1)}\mathbf{d}_n^{T(i-1)}\psi'\left(z_n^{(i-1)} - \mathbf{d}_n^{T(i-1)}\mathbf{\Delta}\right) \tag{80}$$

$$= \mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}\mathbf{D}^{(i-1)} - \mathbf{D}^{T(i-1)}\mathbf{V}^{(i-1)}\mathbf{D}^{(i-1)} \tag{81}$$

$$= \mathbf{D}^{T(i-1)}\left(\mathbf{W}^{(i-1)} - \mathbf{V}^{(i-1)}\right)\mathbf{D}^{(i-1)} \tag{82}$$

where $\mathbf{V}^{(i-1)} = \mathrm{diag}(\psi'(z_1^{(i-1)} - \mathbf{d}_1^{T(i-1)}\mathbf{\Delta}), \ldots \psi'(z_N^{(i-1)} - \mathbf{d}_N^{T(i-1)}\mathbf{\Delta}))$. Since $\psi(x)$ is odd symmetric, and $\psi(x)/x$ is monotonically decreasing for $x > 0$, it is straightforward to show that $\psi(x)/x - \psi'(x) \geq 0, \forall x \neq 0$. Hence

$$\left.\frac{\partial^2\mathcal{I}(\mathbf{\Delta})}{\partial\mathbf{\Delta}\partial\mathbf{\Delta}^T}\right|_{\mathbf{\Delta}=\mathbf{0}} = \mathbf{D}^{T(i-1)}\left[\mathrm{diag}\left(\frac{\psi\left(z_n^{(i-1)}\right)}{z_n^{(i-1)}} - \psi'\left(z_n^{(i-1)}\right)\right)\right]\mathbf{D}^{(i-1)} \tag{83}$$

is positive definite. Therefore, $\mathcal{I}(\mathbf{\Delta}) \geq \mathbf{0}$. Now substituting

$$\mathbf{\Delta} = \mathbf{\Delta}^{(i)} = \left(\mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}\mathbf{D}^{(i-1)}\right)^{-1} \times\mathbf{D}^{T(i-1)}\mathbf{W}^{(i-1)}\mathbf{z}^{(i-1)} \tag{84}$$

as the solution for the search directions at iteration i, we have

$$\mathcal{I}(\mathbf{\Delta}^{(i-1)}) \geq \mathbf{0}. \tag{85}$$

Finally, noting that the second term of the left-hand side of (76) is $\mathcal{C}(\hat{\boldsymbol{\theta}}^{(i)}, \hat{\mathbf{d}}^{(i)})$, while the third term is $\mathcal{C}(\hat{\boldsymbol{\theta}}^{(i-1)}, \hat{\mathbf{d}}^{(i-1)})$, and after some straight forward manipulations, we obtain the inequality in Proposition 1. ■

References

[1] 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, no. 7, pp. 991–1001, 1993.

[2] C. S. Burrus, J. A. Barreto, and I. W. Selesnick, "Iterative reweighted least-squares design of FIR filters," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2926–2936, Nov. 1994.

[3] P. J. Green, "Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives," *J. Roy. Stat. Soc. B, Methodologic.*, vol. 46, no. 2, pp. 149–192, 1984.

[4] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley, 1986.

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

[6] A. Høst-Madsen and X. Wang, "Performance of blind and group-blind multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1849–1872, Jul. 2002.

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

[8] R. A. Iltis, "A DS-CDMA tracking mode receiver with joint channel delay estimation and MMSE detection," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1170–1179, Oct. 2001.

[9] R. A. Iltis and L. Mailaender, "An adaptive multiuser detector with joint amplitude and delay estimation," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 5, pp. 774–784, Jun. 1994.

[10] J. Jureckova and P. Sen, *Robust Statistical Procedures: Asymptotics and Interrelations*. New York: Wiley-Interscience, 1996.

[11] E. E. Kuruoglu, "Nonlinear least l_p norm filters for nonlinear autoregressive α-stable processes," *Digit. Signal Process.*, vol. 12, no. 1, pp. 119–142, Jan. 2002.

[12] J. S. Lee and L. E. Miller, *CDMA Systems Engineering Handbook*. Boston, MA: Artech House, 1998.

[13] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1984.

[14] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 496–508, Apr. 1990.

[15] K. Mahata and T. Soderström, "Large sample properties of separable nonlinear least squares estimators," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1650–1658, Jun. 2004.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.

[16] A. Mantravadi and V. V. Veeravalli, "MMSE detection in asynchronous CDMA systems: An equivalence result," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3128–3137, Dec. 2002.
[17] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, no. 3, pp. 106–127, Aug. 1977.
[18] ——, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class A and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.
[19] T. K. Moon, Z. Xie, C. K. Rushforth, and R. T. Short, "Parameter estimation in a multiuser communication system," *IEEE Trans. Commun.*, vol. 42, no. 8, pp. 2553–2560, Aug. 1994.
[20] T. Ostman, S. Parkvall, and B. Ottersten, "An improved MUSIC algorithm for estimation of time delays in asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1628–1631, Nov. 1999.
[21] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 858–871, May 1997.
[22] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[23] J. O. Street, R. J. Carroll, and D. Ruppert, "A note on computing robust regression estimates via iterative reweighted least squares," *Amer. Statistician*, vol. 42, no. 2, pp. 152–154, May 1988.
[24] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 84–93, Jan. 1996.
[25] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[26] A. Wald, "Note on the consistency of the maximum likelihood estimate," *Ann. Math. Stat.*, vol. 20, no. 4, pp. 595–601, Dec. 1949.
[27] J. Wang, "Asymptotics of least squares estimators for constrained nonlinear regression," *Ann. Stat.*, vol. 24, no. 3, pp. 1316–1326, Jun. 1996.
[28] 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.
[29] ——, "Robust multiuser detection in non-Gaussian channels," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 289–305, Feb. 1999.
[30] J. Wolfowitz, "On Wald's proof of the consistency of the maximum likelihood estimate," *Ann. Math. Stat.*, vol. 20, no. 20, pp. 601–602, Dec. 1949.
[31] C. F. Wu, "Asymptotic theory of nonlinear least squares estimation," *Ann. Stat.*, vol. 9, no. 3, pp. 501–513, May 1981.
[32] Z. Xie, C. K. Rushforth, R. T. Short, and T. K. Moon, "Joint signal detection and parameter estimation in multiuser communications," *IEEE Trans. Commun.*, vol. 41, no. 7, pp. 1208–1216, Aug. 1993.
[33] V. J. Yohai and R. A. Maronna, "Asymptotic behavior of M-estimators for the linear model," *Ann. Stat.*, vol. 7, no. 2, pp. 258–268, Mar. 1979.
[34] J. Zhang, E. Chong, and D. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1128–1144, Mar. 2001.
[35] A. Zoubir and R. Brcich, "Multiuser detection in heavy tailed noise," *Digit. Signal Process.*, vol. 12, pp. 262–273, 2002.

Duc Son Pham (S'02–M'05) received the Bachelor of Engineering degree (first-class hons.) from Monash University, Australia, in 2001 and the Ph.D. degree in electrical engineering from Curtin University of Technology, Perth, Australia, in 2005.

He is now working as a postdoctoral research fellow at Curtin University of Technology. His current research interests fall in the area of statistical signal processing with applications to communications.

Abdelhak M. Zoubir (SM'97) received the Dipl.-Ing. degree (B.Sc./B.Eng.) from Fachhochschule Niederrhein, Germany, in 1983 and the Dipl.-Ing. (M.Sc./M.Eng.) and the Dr.-Ing. (Ph.D.) degree from Ruhr University Bochum, Germany, in 1987 and 1992, all in electrical engineering.

Early placement in industry (Klöckner–Moeller & Siempelkamp AG) was then followed by Associate Lectureship in the Division for Signal Theory at Ruhr University Bochum. In June 1992, he joined Queensland University of Technology, Brisbane, Australia, where he was Lecturer, Senior Lecturer, and then Associate Professor in the School of Electrical and Electronic Systems Engineering. In March 1999, he took up the position of Professor of Telecommunications at Curtin University of Technology, where he was Head of the School of Electrical and Computer Engineering from November 2001 until February 2003. In February 2003, he took up the position of Professor in signal processing at Darmstadt University of Technology, Darmstadt, Germany. His general research interest lies in statistical methods for signal processing with applications in communications, sonar, radar, biomedical engineering, and vibration analysis. His current research interest lies in robust estimation and in bootstrap techniques for spectrum estimation and the modelling of nonstationary and non-Gaussian signals.

Dr. Zoubir was the General Co-Chairman of the Third IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) held in Darmstadt, Germany, in December 2003; the Technical Chairman of the Eleventh IEEE Workshop on Statistical Signal processing held in Singapore in August 2001; and Deputy Technical Chairman (Special Sessions/Tutorials) of the 1994 International Conference on Acoustics, Speech and Signal processing (ICASSP-94) held in Adelaide, Australia. He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 until 2005 and is currently Associate Editor of the EURASIP journals *Signal Processing* and the *Journal of Applied Signal Processing*. He is a Member of the IEEE Signal Processing Society (SPS) Technical Committee on Signal Processing Theory and Methods, Signal Processing Education, and Sensor Array and Multichannel Signal Processing. He is a Member of the Institute of Mathematical Statistics.

Ramon F. Brcic (M'03) received the B.Eng. (Hons. 1) degree in aerospace avionics from Queensland University of Technology, Brisbane, Australia, in 1998 and the Ph.D. degree in electrical engineering from Curtin University of Technology (CUT), Perth, Australia, in 2003.

In 2002, he was a Research Fellow with the School of Electrical Engineering, CUT. Since March 2003, he has been a Research Associate with the Signal Processing Group, Technische Universität Darmstadt, Darmstadt, Germany. His research interests include modeling, estimation, and detection for heavy-tailed processes, and robust signal processing.

Dr. Brcic received an IEEE Signal Processing Society Young Author Award for 2003.

Yee Hong Leung (M'80) was born in Singapore in 1955. He received the B.E. (Hons.) and Ph.D. degrees, both in electrical engineering, from the University of Western Australia, Australia, in 1977 and 1986, respectively.

He was with the School of Electrical, Electronic and Computer Engineering, the University of Western Australia, as a Tutor from 1982 to 1986 and as a Lecturer from 1988 to 1994. From 1986 to 1988, he was with the Defence Science and Technology Organisation, South Australia, as a Research Scientist. In 1994, he joined the Western Australian Telecommunications Research Institute, Curtin University of Technology, Perth, Australia, as a Senior Research Fellow. He is also associated with the Australian Telecommunications Cooperative Research Centre. He joined the Department of Electrical and Computer Engineering, Curtin University of Technology, in 2004 as a Senior Lecturer. His research interests are in adaptive and optimum signal processing, and their applications in antenna arrays and telecommunication systems.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:08 from IEEE Xplore. Restrictions apply.