Robust Multiuser Detection for Multicarrier CDMA Systems

Rensheng Wang, *Member, IEEE*, Hongbin Li, *Member, IEEE*, and Tao Li

***Abstract*—Multiuser detection (MUD) for code-division multiple-access (CDMA) systems usually relies on some *a priori* channel estimates, which are obtained either blindly or by using training sequences, and the covariance matrix of the received signal, usually replaced by the sample covariance matrix. However, such prior estimates are often affected by errors that are typically ignored in subsequent detection. In this paper, we present robust channel estimation and MUD techniques for multicarrier (MC) CDMA by explicitly taking into account such estimation errors. The proposed techniques are obtained by optimizing the worst case performance over two bounded uncertainty sets pertaining to the two types of estimation errors. We show that although the estimation errors associated with the prior channel estimate and the sample covariance matrix are generally not bounded, it is beneficial to optimize the worst case performance over properly chosen *bounded* uncertainty sets determined by a parameter called *bounding probability*. At a slightly higher computational complexity, our proposed robust detectors are shown to yield improved performance over the standard detectors that ignore the prior estimation errors.**

***Index Terms*—Multicarrier code-division multiple access (MC-CDMA), multiuser detection (MUD), robust channel estimation, robust MUD.**

I. Introduction

MULTICARRIER (MC) code-division multiple-access (CDMA), a popular multiple-access scheme for broadband transmission [1], [2], has received much research interest recently. So far, numerous multiuser detection (MUD) schemes for MC-CDMA have been proposed (see [3]–[6] and references therein). While these MUD schemes may be derived from different principles, they usually rely on some prior estimate of the channel, obtained by either a blind or a training-sequence assisted channel estimation algorithm. However, channel estimation are usually affected by errors and most of existing MUD schemes are known to be sensitive to such errors. Moreover, for effective interference suppression, many MUD schemes also require an estimate of the covariance matrix of the received signal, which is typically the sample covariance matrix. The sample covariance matrix converges slowly, resulting in a poor estimate of the true covariance matrix when the number of samples of the received signal is relatively low [7], [8].

In this paper, we develop robust multiuser detection schemes that explicitly account for both channel and covariance matrix estimation errors. Our schemes build on recent developments in robust adaptive beamforming (RAB) [9]–[13], which can successfully deal with uncertainty in the antenna steering vector. Although RAB in the form of diagonal loading has been known for more than two decades in the array processing community (see [9] and [14]), it was unclear how to choose the diagonal loading factor precisely and relate it to the amount of uncertainty of the steering vector. The problem was studied in several recent studies that involve optimizing the worst case performance over a *bounded* ambiguity set determined by the amount of uncertainty on the steering vector [15]–[19]. It was shown the diagonal loading can be computed exactly by a interior point method [18] or a Newton search over a one-dimensional (1-D) bounded set [15], [16], among other approaches.

While the uncertainty on the steering vector in array processing is usually bounded, this is not the case in our problem, where the prior channel estimation errors and the sample covariance matrix estimation errors are unbounded. Even so, it turns out beneficial to optimize the worst case performance over a properly chosen bounded uncertainty set. By using the Chebyshev inequality [20], we show that, in either case involving channel estimation error or covariance matrix estimation error, a bounded uncertainty set can be determined by a parameter referred to as the *Chebyshev bounding probability*. By choosing the Chebyshev bounding probability sufficiently large (e.g., ≥ 0.9), we neglect the small-probability event that the estimation error exceeds the chosen bounded set. This is because in that case the prior estimation is so poor and our robust schemes are not expected to help (in fact, few methods will succeed when the initial estimates are very poor). Hence, our strategy is to try to achieve robustness against small-to-medium prior estimation errors. This makes our work distinct from earlier RAB studies.

Driven by the advances in RAB, robust detection and estimation for wireless communications is receiving more attention. It is interesting to note that a similar approach [21], which was brought to our attention by one reviewer, was considered for robust linear detection in multiuser multiple-input multiple-output (MIMO) systems. The Chebyshev inequality was employed there to simplify a probability constrained optimization-based receiver, which requires nonlinear programming, to one

Manuscript received January 16, 2005; revised July 11, 2005. This work was supported in part by the U.S. National Science Foundation under Grant CCF-0514938, in part by the U.S. Army Research Office under Grant DAAD19-03-1-0184, in part by the Secure Infrastructure Technology Laboratory (SInTeL) project under Grant N00014-05-1–0632, and in part by the National Natural Science Foundation of China under Grant 60502011. This paper was presented in part at the 38th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, November 7–10, 2004.

R. Wang and H. Li are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: rwang1@stevens.edu; hli@stevens.edu).

T. Li is with the Department of Computer Science, Sichuan University, Chengdu, 610065, China (e-mail: litao@scu.edu.cn).

Digital Object Identifier 10.1109/JSAC.2005.862418

that requires only second-order programming, which is computationally much simpler. Although the problems addressed are different, the idea behind [21] and our work is to ensure the probability that the design constraint is violated is small. In [22], robust blind MUD was considered for synchronous CDMA systems to deal with user signature mismatch. Different from our approach, it was assumed that the mismatch is bounded with a known upper bound. Other recent relevant studies in wireless communications can be found in the references of [21] and [22].

Our robust channel estimation and MUD schemes are first developed based on batch processing. We also extend them for adaptive implementations. We show that in the latter case, the uncertainty set pertaining to the estimation error contributed by the sample covariance matrix is no longer fixed and should decrease as more and more symbols are received. As such, time-varying covariance estimation errors are employed in our adaptive implementations. To allow variable uncertainty set in adaptive implementation is another unique feature of our work.

To exemplify the usefulness of our robust design, we consider a class of standard linear MUD receivers, including the minimum variance (MV) detector (e.g., [23]–[25]) assuming knowledge of only one user's spreading code and training,[1] and the zero-forcing MV (ZF-MV) detector [26], [27] with spreading codes and training of multiple users. In the latter case, we consider two scenarios with or without residual unmodeled interference [e.g., intercell interference (ICI)]. We discuss robust versions of these detectors obtained from our robust design. We compare the performance of these MUD detectors driven by training-based channel estimates and our robust channel estimates. Numerical results show that our proposed robust schemes yield significant performance improvements over the standard detectors that ignore the prior channel and sample covariance matrix estimation errors.

The rest of the paper is organized as follows. In Section II, we introduce the MC-CDMA data model and formulate the problem of interest. In Section III, we first review a class of linear MUD receivers. Then, we examine estimation errors in the initial channel estimate and the sample covariance matrix. In Section IV, we develop our robust channel and MUD schemes. Adaptive implementations are discussed in Section V, followed by numerical results in Section VI. Finally, we draw conclusions in Section VII.

A. Notation

Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the conjugate, transpose, and conjugate transpose, respectively; $\|\cdot\|$ denotes the Frobenius norm; $\mathbf{I}_L$ denotes the $L \times L$ identity matrix; $\mathrm{diag}\{\cdot\}$ denotes a diagonal matrix; and finally, $\mathrm{tr}\{\cdot\}$ denotes the trace a matrix.

II. Data Model and Problem Formulation

Consider a baseband MC-CDMA (uplink) system with P subcarriers and K active users. For the kth user, each transmit symbol is modulated by a $P\times 1$ preassigned spreading code $\mathbf{c}_k$ (frequency domain spreading). After a P-point inverse discrete Fourier transform (IDFT) and parallel-to-serial (P/S) conversion, a cyclic prefix (CP) is inserted between successive symbols to avoid inter-symbol interference. Then, the signal is sent out, passing through an $L\times 1$ fading channel $\mathbf{g}_k$, which lumps together the transmit/receiver filters and the physical channel. At the receiver, the received signal is first serial-to-parallel (S/P) converted, followed by CP removal and discrete Fourier transform (DFT). We assume that the propagation delay spread is within the duration of the CP so that the output of the DFT processor during the mth symbol interval can be expressed as (e.g., [28])

$$\mathbf{y}(m) = \sum_{k=1}^{K} d_k(m)\sqrt{P_k}\mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{g}_k + \mathbf{e}(m) \tag{1}$$

where P_k and $d_k(m)$ denote user k's transmission power and data symbol, respectively, $\mathbf{C}_k = \mathrm{diag}\{\mathbf{c}_k\}$ with $\mathbf{c}_k$ denoting user k's spreading code, $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{P\times L}$ denotes the DFT matrix with the klth element given by $(1/\sqrt{P})e^{-j2\pi(k-1)(l-1)/P}$, and $\mathbf{e}(m)$ is the disturbance containing the additive channel noise and possibly unmodeled interference (e.g., ICI). Since the transmission power P_k and channel impulse response $\mathbf{g}_k$ cannot be separated from one another in channel estimation, they are treated as a combined channel vector $\mathbf{h}_k \triangleq \sqrt{P_k}\mathbf{g}_k$. Then, (1) can be rewritten as

$$\mathbf{y}(m) = \sum_{k=1}^{K} d_k(m)\mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{h}_k + \mathbf{e}(m).$$

We consider MUD schemes that require the channel state information (CSI) $\mathbf{h}_k$ for multiuser interference (MUI) suppression. The true CSI can never be known exactly and has to be replaced by some channel estimate $\hat{\mathbf{h}}_k$ which is inevitably subject to estimation error. The problem of interest is to develop robust linear MUD schemes by taking into account imperfect channel estimate $\hat{\mathbf{h}}_k$, along with the other type of imperfection occurred in the estimation of the covariance matrix of $\mathbf{y}(m)$, which is often needed in many MUD schemes.

III. Preliminaries

To facilitate our discussion of the proposed robust MUD scheme in Section IV, we first briefly review a class of linear MUD receivers that utilize the ZF constraint or MV criterion or a combination of both, as well as training-based channel estimation with the least-squares (LS) criterion. Then, we analyze the estimation error of the LS based channel estimator and that of the sample covariance matrix. These analyses can help us determine the boundaries of the uncertainty sets in our robust MUD receiver design in Section IV.

A. Linear Multiuser Detection

Let $\mathbf{s}_k = \mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{h}_k \in \mathbb{C}^{P\times 1}$, for $k = 1, \ldots, K$, $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_K]$, and $\mathbf{d}(m) \triangleq [d_1(m), \ldots, d_K(m)]^T$. Then $\mathbf{y}(m)$ can be rewritten in a compact form

$$\mathbf{y}(m) = \mathbf{S}\mathbf{d}(m) + \mathbf{e}(m).$$

[1]While the MV principle can be employed for blind channel estimation, we assume that the channel estimation is initially obtained by training (and refined by our robust techniques).

The ZF detector, which assumes the signature vectors $\mathbf{s}_1, \ldots, \mathbf{s}_K$ are known at the receiver, is given by [29]

$$\mathbf{W}_{\text{ZF}} = \mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}$$

which detects all K users simultaneously. The ZF detector ignores any possible residual interference in $\mathbf{e}(m)$ (e.g., ICI or unmodeled interference). To account for such residual interference, we can use ZF-MV criterion (e.g., [26], [30])

$$\mathbf{W}_{\text{ZF-MV}} = \arg\min_{\mathbf{W}} \operatorname{tr}\{\mathbf{W}^H\mathbf{R}_y\mathbf{W}\}, \text{ s.t. } \mathbf{W}^H\mathbf{S} = \mathbf{I}_K \quad (2)$$
$$= \mathbf{R}_y^{-1}\mathbf{S}(\mathbf{S}^H\mathbf{R}_y^{-1}\mathbf{S})^{-1} \quad (3)$$

where $\mathbf{R}_y \triangleq E\{\mathbf{y}(m)\mathbf{y}^H(m)\}$ denotes the covariance matrix of $\mathbf{y}(m)$ and in obtaining the second equality, we used a standard constrained quadratic optimization result (e.g., [31, p. 283]). The ZF constraint in (2) ensures that each row of $\mathbf{W}^H$ passes one user with unit gain but completely suppress (with zero gain) the other $K-1$ interfering users. Meanwhile, the MV criterion in (2) further cancels the residual/unmodeled interference in $\mathbf{e}(m)$.

The ZF-MV receiver reduces to the MV or minimum output energy (MOE) detector [23]–[25] if $K = 1$, i.e., only a single user, say user 1, is detected

$$\mathbf{w}_{\text{MV}} = \mathbf{R}_y^{-1}\mathbf{s}_1(\mathbf{s}_1^H\mathbf{R}_y^{-1}\mathbf{s}_1)^{-1}. \quad (4)$$

We note that if $\mathbf{s}_1$ and $\mathbf{R}_y$ are known exactly, $\mathbf{w}_{\text{MV}}$ is the same as the optimum linear minimum mean-squared error (MMSE) detector [23], [24]; meanwhile, a similar ZF-MV approach was used in [27] for deriving group blind MUD schemes.

The minimum output power of the ZF-MV detector is

$$V \triangleq \operatorname{tr}\{(\mathbf{S}^H\mathbf{R}_y^{-1}\mathbf{S})^{-1}\} \quad (5)$$

which can be verified by substituting (3) into the cost function in (2). The minimum output (5) will be used in our robust receiver design in Section IV.

B. Channel Estimation

While the above MUD detectors are based on different principles and prior knowledge, they all rely on the CSI. Suppose M training symbols are available. We stack $\mathbf{y}(m)$ as $\mathbf{y} \triangleq [\mathbf{y}^T(1), \ldots, \mathbf{y}^T(M)]^T$. Then

$$\mathbf{y} = \begin{bmatrix} d_1(1)\mathbf{C}_1\boldsymbol{\mathcal{F}} & \cdots & d_K(1)\mathbf{C}_K\boldsymbol{\mathcal{F}} \\ \vdots & \ddots & \vdots \\ d_1(M)\mathbf{C}_1\boldsymbol{\mathcal{F}} & \cdots & d_K(M)\mathbf{C}_K\boldsymbol{\mathcal{F}} \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_K \end{bmatrix} + \begin{bmatrix} \mathbf{e}(1) \\ \vdots \\ \mathbf{e}(M) \end{bmatrix}$$
$$\triangleq \mathbf{A}\mathbf{h} + \mathbf{e}$$

where we assume that we have the training information of all K users and, therefore, $\mathbf{A}$ is known. The LS estimate of $\mathbf{h}$ is given by

$$\hat{\mathbf{h}} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{y}. \quad (6)$$

Note that, if there is no residual interference in $\mathbf{e}(m)$, the LS channel estimate in (6) is equivalent to the optimum maximum-likelihood (ML) estimate [32].

The channel estimation error of (6) is

$$\Delta\mathbf{h} = \hat{\mathbf{h}} - \mathbf{h} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{e}. \quad (7)$$

Suppose $\mathbf{e}$ contains white Gaussian noise with $E\{\mathbf{e}\mathbf{e}^H\} = \sigma^2\mathbf{I}$. While this assumption is not satisfied in the presence of residual interference, this is a simple and suitable assumption if the statistical property of the residual interference is unknown. Following the assumption, $\Delta\mathbf{h}$ is zero-mean Gaussian (conditioned on the training symbols) with covariance matrix

$$\operatorname{cov}\{\Delta\mathbf{h}\} = \sigma^2(\mathbf{A}^H\mathbf{A})^{-1}. \quad (8)$$

Let $\boldsymbol{\mathcal{A}} = (\mathbf{A}^H\mathbf{A})$ and $\boldsymbol{\mathcal{A}}_{ij}$ the (ij)th $L \times L$ submatrix of $\boldsymbol{\mathcal{A}}$: $\boldsymbol{\mathcal{A}}_{ij} = \boldsymbol{\mathcal{F}}^H\mathbf{C}_i\mathbf{C}_j\boldsymbol{\mathcal{F}}\sum_{m=1}^M d_i^*(m)d_j(m)$, $i,j = 1,\ldots,K$. For independent and identically distributed (i.i.d.) training symbols, the *sample cross-correlation* $r_{ij} \triangleq (1/M)\sum_{m=1}^M d_i^*(m)d_j(m) = O(M^{-1/2})$, $i \neq j$, which means that r_{ij} approaches zero (which is the statistical cross-correlation) at rate $M^{-1/2}$ as M increases. Hence, for large M, $\boldsymbol{\mathcal{A}} = M(\mathbf{I}_{KL} + O(M^{-1/2}))$ and, in turn, $\boldsymbol{\mathcal{A}}^{-1} = (1/M)\mathbf{I}_{KL} + O(M^{-3/2})$ [33, p. 58] (i.e., inverse of a perturbed matrix). Substituting $\boldsymbol{\mathcal{A}}^{-1}$ (i.e., $(\mathbf{A}^H\mathbf{A})^{-1}$) into (8), for large M, we have

$$\operatorname{cov}\{\Delta\mathbf{h}\} \approx \frac{\sigma^2}{M}\mathbf{I}_{KL}. \quad (8a)$$

Recall that $\mathbf{h} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T$. We can further infer that

$$\operatorname{cov}\{\Delta\mathbf{h}_k\} = \frac{\sigma^2}{M}\mathbf{I}_L. \quad (9)$$

Let us denote $\beta \triangleq \|\Delta\mathbf{h}_k\|^2$. Clearly, β is χ^2 distributed with $2L$ degrees of freedom [34], whose mean μ_β and variance σ_β^2 are given by

$$\mu_\beta = E\{\beta\} \approx \frac{L}{M}\sigma^2, \quad \sigma_\beta^2 = \operatorname{var}\{\beta\} \approx \frac{L}{M^2}\sigma^4. \quad (10)$$

The above calculation is used to choose the size of the uncertainty set. Simulation results in Section VI show that the performance of the proposed approach is insensitive to the choice. Hence, an estimate of $\operatorname{cov}\{\Delta\mathbf{h}\}$ suffices. Note approximation is not needed for orthogonal training. Although for nonorthogonal training and small M, the above approximation may be inaccurate, it still offers a guideline regarding how to set the ambiguity set. Nonorthogonal training and relatively small M are used in our computer simulation. Our results show significant performance improvement can still be obtained with the proposed approach in such cases.

C. Covariance Matrix Estimation

Note that $\mathbf{R}_y$ is used in the ZF-MV (3) and MV (4) detectors for interference suppression. In reality, the true covariance matrix $\mathbf{R}_y$ is not available and usually substituted by the sample covariance matrix

$$\hat{\mathbf{R}}_y = \frac{1}{M}\sum_{m=1}^{M}\mathbf{y}(m)\mathbf{y}(m)^H. \tag{11}$$

While $\hat{\mathbf{R}}_y$ is the ML estimate of $\mathbf{R}_y$, the above linear multiuser detectors suffer from small M, due to the well-known convergence problem. We next characterize the estimation errors in $\hat{\mathbf{R}}_y$. In particular, we consider estimation error $\gamma \triangleq \|\hat{\mathbf{R}}_y - \mathbf{R}_y\|^2$. Let $\hat{R}_y(i,j)$ and $R_y(i,j)$ be the (i,j)th element of $\hat{\mathbf{R}}_y$ and $\mathbf{R}_y$, respectively. Then

$$\gamma = \sum_{i=1}^{P}\sum_{j=1}^{P}\left[\hat{R}_y(i,j) - R_y(i,j)\right]\left[\hat{R}_y(i,j) - R_y(i,j)\right]^*$$

is a sum of P^2 random variables. As shown in Appendix I, the mean and variance of γ are given by

$$\begin{aligned}\mu_\gamma &= E\{\gamma\} = \frac{1}{M}\mathrm{tr}^2\{\mathbf{R}_y\}\\ \sigma_\gamma^2 &= \mathrm{var}\{\gamma\} = \frac{2}{M^2}\|\mathbf{R}_y\|^4.\end{aligned} \tag{12}$$

We will use the above statistics of γ for robust MUD design in Section IV.

IV. Robust Multiuser Detection

In this section, we will develop a class of robust MUD schemes by taking into account the estimation errors in $\hat{\mathbf{h}}$ and $\hat{\mathbf{R}}_y$. Our schemes optimize the worst case performance over bounded sets of the above estimation errors. First, we discuss how to bound the above estimation errors.

A. Bounding Channel Estimation Error

Although $\beta = \|\Delta\mathbf{h}_k\|^2$ is χ^2 distributed and, therefore, not bounded, by the Chebyshev inequality, the unbounded channel estimation error is bounded in probability

$$P_\beta(|\beta - \mu_\beta| > \delta_\beta) \le \frac{\sigma_\beta^2}{\delta_\beta^2} \tag{13}$$

where μ_β and σ_β^2 denote the mean and variance, respectively, of β, which are given in (10), and δ_β is any positive number. The right-hand side of (13) can be made very small by choosing a large δ_β. We will ignore the *unbounded channel estimation error*, which is a small-probability event for sufficiently large δ_β, and consider a bounded set

$$P_\beta(\beta \le \mu_\beta + \delta_\beta) \ge 1 - \frac{\sigma_\beta^2}{\delta_\beta^2}. \tag{14}$$

Let $\epsilon_k \triangleq \mu_\beta + \delta_\beta$ denote a chosen boundary of β. Then

$$P_\beta(\|\Delta\mathbf{h}_k\|^2 \le \epsilon_k) \ge 1 - \frac{\sigma_\beta^2}{\delta_\beta^2} \tag{15}$$

where $P_\beta(\|\Delta\mathbf{h}_k\|^2 \le \epsilon_k)$ is henceforth referred to as the *Chebyshev bounding probability*. Although we can choose a sufficiently large ϵ_k to make the Chebyshev bounding probability P_β approach 1, it is not worth seeking robustness when the channel estimation error is too large (see Section I). Hence, our strategy is to improve detection robustness against small to moderate channel estimation errors. For a chosen Chebyshev bounding probability P_β (say, $P_\beta = 0.9$), we can determine the corresponding boundary ϵ_k by setting the two sides of (15) equal. In particular, we have

$$\epsilon_k = \mu_\beta + \sqrt{\frac{\sigma_\beta^2}{(1-P_\beta)}}. \tag{16}$$

In light of (10), we can rewrite ϵ_k as

$$\epsilon_k = \frac{1}{M}\left(L + \sqrt{\frac{L}{(1-P_\beta)}}\right)\sigma^2. \tag{17}$$

B. Bounding Covariance Matrix Estimation Error

Likewise, we can bound covariance matrix estimation error $\gamma = \|\hat{\mathbf{R}}_y - \mathbf{R}_y\|^2$ by the Chebyshev inequality

$$P_\gamma(|\gamma - \mu_\gamma| > \delta_\gamma) \le \frac{\sigma_\gamma^2}{\delta_\gamma^2} \tag{18}$$

where μ_γ and σ_γ^2 denote the mean and variance, respectively, of γ, which are given by (12), and δ_γ is a positive number. Again, we can ignore the *unbounded small-probability* event (for sufficiently large δ_γ) and consider a bounded set

$$P_\gamma(\gamma \le \mu_\gamma + \delta_\gamma) \ge 1 - \frac{\sigma_\gamma^2}{\delta_\gamma^2}. \tag{19}$$

It will be more convenient to work with the estimation error $\sqrt{\gamma}$ instead of γ. To this end, let $\eta = \sqrt{\mu_\gamma + \delta_\gamma}$ denote the boundary for γ. Then

$$P_\gamma(\|\hat{\mathbf{R}}_y - \mathbf{R}_y\| \le \eta) \ge 1 - \frac{\sigma_\gamma^2}{\delta_\gamma^2}. \tag{20}$$

For a given Chebyshev bounding probability P_γ, we set the two sides of (20) equal in order to determine the boundary η, which is given by

$$\eta = \sqrt{\mu_\gamma + \sqrt{\frac{\sigma_\gamma^2}{(1-P_\gamma)}}}. \tag{21}$$

In light of (12), we have

$$\eta = \frac{1}{\sqrt{M}}\sqrt{\mathrm{tr}^2\{\mathbf{R}_y\} + \frac{\sqrt{2}\,\|\mathbf{R}_y\|^2}{\sqrt{1-P_\gamma}}}. \tag{22}$$

C. Robust Channel Estimation and MUD

The ZF-MV detector in (3) is sensitive to signal mismatch due to errors in $\hat{\mathbf{S}}$ (equivalently, $\hat{\mathbf{h}}$) and errors in $\hat{\mathbf{R}}_y$. Next, we develop a class of robust ZF-MF detectors by accounting for such imperfect prior estimation. We consider channel estimation by *maximizing* the following multichannel output power [cf. (5)]:

$$V(\mathbf{h}) = \mathrm{tr}\{[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})]^{-1}\} \tag{23}$$

subject to certain constraints (to be specified). The idea is after interference suppression [enforced by the ZF-MV criterion in (2)], we should maximize the output power to avoid signal cancellation (e.g., [24]).

It turns out computationally involved to use (23), which is highly nonlinear in $\mathbf{h}$ (due to the outer matrix inversion). Instead, we utilize an upper bound of $V(\mathbf{h})$ obtained in a similar manner as in [30] by using the Schwartz inequality

$$\begin{aligned} K^2 &= \mathrm{tr}^2(\mathbf{I}_K) \\ &\quad \mathrm{tr}^2\left\{\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right]^{-1/2}\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right]^{1/2}\right\} \\ &\le \mathrm{tr}\left\{\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right]^{-1}\right\}\mathrm{tr}\left\{\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right]\right\} \end{aligned} \tag{24}$$

which yields

$$\mathrm{tr}\left\{\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right]^{-1}\right\} \ge \frac{K^2}{\mathrm{tr}\left\{\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right\}}. \tag{25}$$

Following a similar proof in [30], we can show that the above bound is asymptotically tight for high signal-to-noise ratio (SNR). Therefore, rather than maximizing $\mathrm{tr}\{\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right]^{-1}\}$, we can minimize $\mathrm{tr}\left\{\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right\}$. Our channel estimate is obtained by considering the following constrained optimization:

$$\begin{aligned} \{\tilde{\mathbf{R}}_y,\, \tilde{\mathbf{h}}\} &= \arg\min_{\mathbf{h},\mathbf{R}_y} \mathrm{tr}\left\{\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right\} \\ \text{s.t. } \|\hat{\mathbf{h}} - \mathbf{h}\|^2 &\le \epsilon \text{ and } \|\hat{\mathbf{R}}_y - \mathbf{R}_y\| \le \eta \end{aligned} \tag{26}$$

where $\epsilon = \sum_{k=1}^K \epsilon_k$. In plain English, we seek to optimize the cost function over two spherical constrained sets centered on the initial estimates with radius determined by their statistical uncertainty. Our updated estimates optimize the best worst case performance over the two uncertainty sets.

The cost function can be simplified as follows:

$$\begin{aligned} V_1(\mathbf{h}) &= \mathrm{tr}\left[\mathbf{S}^H(\mathbf{h})\mathbf{R}_y^{-1}\mathbf{S}(\mathbf{h})\right] \\ &= \sum_{k=1}^K \mathbf{h}_k^H\boldsymbol{\mathcal{F}}^H\mathbf{C}_k\mathbf{R}_y^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{h}_k. \end{aligned} \tag{27}$$

Note that $\mathbf{h}_k^H\boldsymbol{\mathcal{F}}^H\mathbf{C}_k\mathbf{R}_y^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{h}_k \ge 0$, for $k = 1, \ldots, K$, and they are independent of each other. Moreover, for a fixed $\mathbf{h}_k$, we have

$$\begin{aligned} \tilde{\mathbf{R}}_y &= \arg\min_{\|\hat{\mathbf{R}}_y - \mathbf{R}_y\| \le \eta} \mathbf{h}_k^H\boldsymbol{\mathcal{F}}^H\mathbf{C}_k\mathbf{R}_y^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{h}_k \\ &= \hat{\mathbf{R}}_y + \eta\mathbf{I} \end{aligned} \tag{28}$$

which is also independent of different users. Therefore, by substituting $\tilde{\mathbf{R}}_y$ into (26), the problem in (26) reduces to K separate optimization problems

$$\begin{aligned} \tilde{\mathbf{h}}_k &= \arg\min_{\mathbf{h}_k} \mathbf{h}_k^H\boldsymbol{\mathcal{F}}^H\mathbf{C}_k(\hat{\mathbf{R}}_y + \eta\mathbf{I})^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}\mathbf{h}_k \\ &\qquad \text{s.t. } \|\hat{\mathbf{h}}_k - \mathbf{h}_k\|^2 \le \epsilon_k, \quad k = 1, \ldots, K. \end{aligned} \tag{29}$$

Let $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\mathcal{F}}^H\mathbf{C}_k(\hat{\mathbf{R}}_y + \eta\mathbf{I})^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}$. Since the solution of (29) will evidently occur on the boundary of the uncertainty set (i.e., the worst case), then

$$\tilde{\mathbf{h}}_k = \arg\min_{\mathbf{h}_k} \mathbf{h}_k^H\boldsymbol{\Phi}_k\mathbf{h}_k, \quad \text{s.t. } \|\hat{\mathbf{h}}_k - \mathbf{h}_k\|^2 = \epsilon_k \tag{30}$$

where the inequality constraint has been replaced by a quadratic equality constraint. The problem in (30) can be solved by using the *Lagrange multiplier*, in a manner similar to [15]. Specifically, Let

$$V_2(\mathbf{h}_k, \lambda) = \mathbf{h}_k^H\boldsymbol{\Phi}_k\mathbf{h}_k + \lambda[(\hat{\mathbf{h}}_k - \mathbf{h}_k)^H(\hat{\mathbf{h}}_k - \mathbf{h}_k) - \epsilon_k]. \tag{31}$$

Then, the new channel estimate is obtained by taking the partial derivative and setting it to zero

$$\frac{\partial V_2(\mathbf{h}_k, \lambda)}{\partial \mathbf{h}_k} = \boldsymbol{\Phi}_k\tilde{\mathbf{h}}_k + \lambda(\tilde{\mathbf{h}}_k - \hat{\mathbf{h}}_k) = \mathbf{0} \tag{32}$$

which yields

$$\tilde{\mathbf{h}}_k = \lambda(\boldsymbol{\Phi}_k + \lambda\mathbf{I})^{-1}\hat{\mathbf{h}}_k = \hat{\mathbf{h}}_k - (\mathbf{I} + \lambda\boldsymbol{\Phi}_k^{-1})^{-1}\hat{\mathbf{h}}_k. \tag{33}$$

The Lagrange-multiplier λ can be calculated by setting

$$V_3(\lambda) \triangleq \|\tilde{\mathbf{h}}_k - \hat{\mathbf{h}}_k\|^2 = \|(\mathbf{I} + \lambda\boldsymbol{\Phi}_k^{-1})^{-1}\hat{\mathbf{h}}_k\|^2 = \epsilon_k. \tag{34}$$

Let the eigenvalue decomposition (EVD) of $\boldsymbol{\Phi}_k^{-1}$ be $\boldsymbol{\Phi}_k^{-1} = \mathbf{U}_k\boldsymbol{\Gamma}_k\mathbf{U}_k^H$, $\boldsymbol{\alpha} \triangleq \mathbf{U}_k^H\hat{\mathbf{h}}_k = [\alpha_1, \ldots, \alpha_L]^T$, and $\boldsymbol{\Gamma}_k = \mathrm{diag}[\rho_1, \ldots, \rho_L]$ with $\rho_1 \ge \cdots \ge \rho_L$. Then, $V_3(\lambda)$ can be rewritten as

$$V_3(\lambda) = \sum_{l=1}^L \frac{|a_l|^2}{[1 + \lambda\rho_l]^2} = \epsilon_k. \tag{35}$$

Since $V_3(\lambda)$ is monotonically decreasing, we can determine a unique solution lie in the upper and lower bounds given by (cf. [15])

$$\frac{\|\hat{\mathbf{h}}_k\| - \sqrt{\epsilon_k}}{\rho_1\sqrt{\epsilon_k}} \le \lambda \le \min\left\{\left[\epsilon_k^{-1}\sum_{l=1}^L \frac{|\alpha_l|^2}{\rho_l^2}\right]^{1/2}, \frac{\|\hat{\mathbf{h}}_k\| - \sqrt{\epsilon_k}}{\rho_L\sqrt{\epsilon_k}}\right\}. \tag{36}$$

Once λ is obtained, our robust channel estimate is given by

$$\tilde{\mathbf{h}}_k = (\lambda^{-1}\boldsymbol{\Phi}_k + \mathbf{I})^{-1}\hat{\mathbf{h}}_k. \tag{37}$$

In turn, the signature vectors for MUD are updated as follows:

$$\begin{aligned} \tilde{\mathbf{s}}_k &= \mathbf{C}_k\boldsymbol{\mathcal{F}}(\lambda^{-1}\boldsymbol{\Phi}_k + \mathbf{I})^{-1}\hat{\mathbf{h}}_k \\ \tilde{\mathbf{S}} &= [\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_K]. \end{aligned} \tag{38}$$

TABLE I
SUMMARY OF THE ROBUST AND STANDARD MUD DETECTORS AND COMPLEXITIES

Robust MUD	Standard MUD
1. Compute initial channel estimate $\hat{\mathbf{h}}_k$.	1. Compute initial channel estimate $\hat{\mathbf{h}}_k$.
2. Compute $\hat{\mathbf{R}}_y$ by (12). $\Rightarrow O(MP^2)$ flops	2. Compute $\hat{\mathbf{R}}_y$ by (12). $\Rightarrow O(MP^2)$ flops
3. Compute $[\hat{\mathbf{R}}_y + \eta\mathbf{I}]^{-1}$. $\Rightarrow O(P^3)$ flops	3. Compute $\hat{\mathbf{R}}_y^{-1}$. $\Rightarrow O(P^3)$ flops
4. Compute $\mathbf{\Phi}_k = \boldsymbol{\mathcal{F}}^H\mathbf{C}_k[\hat{\mathbf{R}}_y + \eta\mathbf{I}]^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}$. $\Rightarrow O(P^2L)$ flops	4. Compute $\mathbf{s}_k = \mathbf{C}_k\boldsymbol{\mathcal{F}}\hat{\mathbf{h}}_k,\ k = 1,\cdots,K$. $\Rightarrow O(KPL)$ flops
5. Perform the EVD of $\mathbf{\Phi}_k = \mathbf{U}_k\mathbf{\Gamma}_k^{-1}\mathbf{U}_k^H$. $\Rightarrow O(L^3)$ flops	5. Compute $\mathbf{W}_{\text{ZF-MV}}$ by (3) or $\mathbf{w}_{\text{MV}}$ by (4). $\Rightarrow O(KP^2)$ flops
6. Solve (35) for λ by a Newton search scheme.	
7. Compute $\tilde{\mathbf{h}}_k$ by (38) and form $\tilde{\mathbf{S}}$ by (39). $\Rightarrow O(KPL)$ flops	
8. Compute the robust detector $\tilde{\mathbf{W}}_{\text{robust ZF-MV}}$ by (40) or $\tilde{\mathbf{w}}_{\text{robust MV}}$ by (41). $\Rightarrow O(KP^2)$ flops	

The robust ZF-MV detector (multiuser spreading codes are known) is given by [cf. (3)]

$$\tilde{\mathbf{W}}_{\text{robust ZF-MV}} = (\hat{\mathbf{R}}_y + \eta\mathbf{I})^{-1}\tilde{\mathbf{S}}\left[\tilde{\mathbf{S}}^H(\hat{\mathbf{R}}_y + \eta\mathbf{I})^{-1}\tilde{\mathbf{S}}\right]^{-1}. \tag{39}$$

When only a single spreading code is available, the robust ZF-MV detector reduces to a robust MV detector [cf. (4)]

$$\tilde{\mathbf{w}}_{\text{robust MV}} = \frac{(\hat{\mathbf{R}}_y + \eta\mathbf{I})^{-1}\tilde{\mathbf{s}}_1}{\tilde{\mathbf{s}}_1^H(\hat{\mathbf{R}}_y + \eta\mathbf{I})^{-1}\tilde{\mathbf{s}}_1}. \tag{40}$$

The proposed robust MUD detectors along with the complexity (in terms of the number of flops) involved in each step is summarized in Table I. Note that the complexity of Step 1, which is identical for both the robust and standard detectors, depends on a particular channel estimator used and, therefore, is not listed. For the LS estimator discussed in Section III-B, the complexity is $O(K^2L^2P + K^2M + K^3L^3)$ (ignoring the lower order terms). Also note that the Newton search in Step 6 involves 1-D search, as shown in (35) and (36). The variable complexity involved in the search is not listed since it is negligible compared to other matrix/vector manipulations. For comparison, the standard detectors along with their complexities are also listed in Table I. It is seen that the extra steps of the robust detectors involve only slight increase in complexity since L is much smaller compared to P in practical systems. As such, the robust and standard detectors have a comparable complexity.

V. Adaptive Implementation

The above robust MUD schemes are presented in batch-processing form, based on a collection of the received signals during M symbols. In a more realistic scenario, we may have M training symbols followed by N_y data symbols which form a data packet. It is often desirable to detect the received signal in a symbol-to-symbol manner, as opposed to the above batch-processing. In this section, we discuss the symbol-by-symbol adaptive implementation of the proposed robust MUD detectors.

First, we note that the sample covariance matrix

$$\hat{\mathbf{R}}_y(m) = \frac{1}{m}\sum_{i=1}^{m}\mathbf{y}(i)\mathbf{y}^H(i), \qquad m = M,\ M+1,\ \ldots,\ M+N_y \tag{41}$$

becomes more and more accurate as m increases. This is also reflected in (22), which shows that the covariance matrix estimation error decrease in $1/\sqrt{m}$ as m increases. As such, the size of the uncertainty set in (26) with respect to covariance matrix estimation should be time varying and decrease with m

$$\eta(m) = \frac{1}{\sqrt{m}}\eta_0 \tag{42}$$

where $\eta_0 \triangleq \sqrt{\text{tr}^2\{\mathbf{R}_y\} + \sqrt{2}\|\mathbf{R}_y\|^2/\sqrt{1-P_\gamma}}$ is a constant and independent of m. With time-varying size of the uncertainty set, we have the adaptive robust ZF-MV and MV for the mth received symbol, given by

$$\tilde{\mathbf{W}}_{\text{robust ZF-MV}}(m) = [\hat{\mathbf{R}}_y(m) + \eta(m)\mathbf{I}]^{-1}\tilde{\mathbf{S}} \times\left\{\tilde{\mathbf{S}}^H[\hat{\mathbf{R}}_y(m) + \eta(m)\mathbf{I}]^{-1}\tilde{\mathbf{S}}\right\}^{-1} \tag{43}$$

and, respectively

$$\tilde{\mathbf{w}}_{\text{robust MV}}(m) = [\hat{\mathbf{R}}_y(m) + \eta(m)\mathbf{I}]^{-1}\tilde{\mathbf{s}}_1 \times\left[\tilde{\mathbf{s}}_1^H(\hat{\mathbf{R}}_y(m) + \eta(m)\mathbf{I})^{-1}\tilde{\mathbf{s}}_1\right]^{-1}. \tag{44}$$

We note that the computational complexity with (43) and (44) is high since the matrix inverse $[\hat{\mathbf{R}}_y(m) + \eta(m)\mathbf{I}]^{-1} \in \mathbb{C}^{P\times P}$ has to recomputed for every m. In the following, we discuss how to efficiently and adaptively compute it. Based on the definition of $\hat{\mathbf{R}}_y(m)$, we have

$$\hat{\mathbf{R}}_y(m) = \frac{m-1}{m}\hat{\mathbf{R}}_y(m-1) + \frac{1}{m}\mathbf{y}(m)\mathbf{y}^H(m). \tag{45}$$

In the presence of a variable diagonal loading factor $\eta(m)$, we have

$$\hat{\mathbf{R}}_y(m)+\eta(m)\mathbf{I}=\frac{m-1}{m}\left[\hat{\mathbf{R}}_y(m-1)+\eta(m-1)\mathbf{I}\right]+\frac{1}{m}\mathbf{y}(m)\mathbf{y}^H(m)+\left(\sqrt{\frac{m-1}{m}}-\frac{m-1}{m}\right)\eta(m-1)\mathbf{I} \quad (46)$$

where we used the fact that $\eta(m)=\sqrt{(m-1)/m}\eta(m-1)$ [cf. (42)]. It facilitates obtaining a recursive expression for calculating $\hat{\mathbf{R}}_y(m)$ if $\sqrt{(m-1)/m}$ is replaced by $(m-1)/m$. Doing so has little impact on the performance as long as m is not too small (recall that $m>M$). As such, we can drop the last term at the right-hand side of (46), which is equivalent to slightly changing the size of the uncertainty set associated with the sample covariance matrix. Note that the proposed robust detectors are insensitive to such minor adjustment of the uncertainty set size, as shown in the numerical examples in Section VI. Therefore

$$\hat{\mathbf{R}}_y(m)+\eta(m)\mathbf{I}\approx\frac{m-1}{m}\left[\hat{\mathbf{R}}_y(m-1)+\eta(m-1)\mathbf{I}\right]+\frac{1}{m}\mathbf{y}(m)\mathbf{y}^H(m). \quad (47)$$

Let $\kappa(m)\triangleq(m-1)/m$ and $\breve{\mathbf{R}}_y(m)=\hat{\mathbf{R}}_y(m)+\eta(m)\mathbf{I}$. Then

$$\breve{\mathbf{R}}_y(m)=\kappa(m)\breve{\mathbf{R}}_y(m-1)+[1-\kappa(m)]\mathbf{y}(m)\mathbf{y}^H(m) \quad (48)$$

where $\kappa(m)$ is a variable forgetting factor, changing from symbol to symbol. It is clear that the standard recursive least square (RLS) like iteration can be used (using the matrix inversion lemma) [35]

$$\breve{\mathbf{R}}_y^{-1}(m)=\frac{1}{\kappa}\breve{\mathbf{R}}_y^{-1}(m-1)-\frac{1-\kappa}{\kappa^2}\frac{\breve{\mathbf{R}}_y^{-1}(m-1)\mathbf{y}(m)\mathbf{y}^H(m)\breve{\mathbf{R}}_y^{-1}(m-1)}{1+(1-\kappa)\mathbf{y}^H(m)\breve{\mathbf{R}}_y^{-1}(m-1)\mathbf{y}(m)}. \quad (49)$$

In summary, our adaptive implementations for the proposed robust MUD consist of the following steps.

- Step 1**:** Initialization :
  - 1.a : Form $\mathbf{y}=[\mathbf{y}^T(1),\ldots,\mathbf{y}^T(M)]^T$ from M training symbols, and use **(6)** to compute $\hat{\mathbf{h}}$.
  - 1.b : Compute $\hat{\mathbf{R}}_y(M)$ by **(41)**, $\eta(m)$ by **(42)**, and matrix inverse $[\hat{\mathbf{R}}_y(M)+\eta(M)\mathbf{I}]^{-1}$.
- Step 2**:** Adaptive estimation and MUD :

  for $m=M+1:M+N_y$
  - 2.a : Compute $[\hat{\mathbf{R}}_y(m)+\eta(m)\mathbf{I}]^{-1}$, i.e., $\breve{\mathbf{R}}_y^{-1}(m)$, by **(49)**.
  - 2.b : Compute $\mathbf{\Phi}_k=\boldsymbol{\mathcal{F}}^H\mathbf{C}_k[\hat{\mathbf{R}}_y(m)+\eta(m)\mathbf{I}]^{-1}\mathbf{C}_k\boldsymbol{\mathcal{F}}$.
  - 2.c : Perform the EVD of $\mathbf{\Phi}_k=\mathbf{U}_k\mathbf{\Gamma}_k^{-1}\mathbf{U}_k^H$.
  - 2.d : Solve **(34)** for λ by a Newton search scheme (e.g., by Matlab function fminbnd).
  - 2.e : Compute the robust channel estimate $\tilde{\mathbf{h}}_k$ by **(37)** and form $\tilde{\mathbf{S}}$.
  - 2.f : Compute the robust MUD detectors $\tilde{\mathbf{W}}_{\text{robust ZF-MV}}(m)$ by **(43)** or $\tilde{\mathbf{w}}_{\text{robust MV}}(m)$ by **(44)**.

  end.

The complexity of the adaptive implementations can be analyzed in the same way as the one done for the batch versions. The results are similar except that the inverse matrix computation $[\hat{\mathbf{R}}_y+\eta\mathbf{I}]^{-1}$ is simplified to $O(P^2)$ flops by the RLS implementation.

VI. Numerical Results

We consider a K-user MC-CDMA (uplink) system with binary phase shift key (BPSK) modulation, $P=32$ subcarriers, $L=7$ for $\mathbf{h}_k$ and random spreading codes $\mathbf{c}_k$. The channel vector $\mathbf{h}_k$ is generated as $L\times 1$ independent complex Gaussian random variables with zero mean and identical variance $1/L$, and varied independently from trial to trial. In the following examples, the number of active users is ten.

We have two sets of multiuser detectors under consideration. The first set consists of our proposed robust ZF-MV detector in (39), which assumes knowledge of the spreading codes and training symbols of multiple users, and the robust MV detector in (40), which assumes the spreading code and training symbols of only the desired user. The other set includes the standard ZF-MV detector in (3) and the MV detector in (4), which ignore the estimation errors in the prior estimates of the CSI and the data covariance matrix. For both the conventional ZF-MV and robust ZF-MV detectors, we further consider two different scenarios: one has knowledge of the spreading codes and training symbols of all ten users; the other has knowledge of six out of ten users, i.e., we assume the presence of ICI with four out of ten users coming from a different cell unknown to the detector at the basestation (of the current cell). Therefore, we have a total of six different MUD detectors, namely our **robust ZF-MV (39)**, **robust ZF-MV (with ICI) (39)**, **robust MV (40)**, **standard ZF-MV (3)**, **standard ZF-MV (with ICI) (3)**, and **standard MV (4)** detectors for comparison.

We first examine the impact of the size ϵ_k of the uncertainty set pertaining to the initial channel estimate [cf. (15)] when $M=50$, SNR $=10$ dB, and $\eta/E\{\|\mathbf{R}_y\|\}=0.2$. Fig. 1 depicts the receiver output signal-to-interference-and-noise ratio (SINR) for user 1 as a function of the normalized $\epsilon_k/E\{\|\mathbf{h}_k\|^2\}$ when the variance σ^2 [cf. (8a)] of the initial channel estimation error is fixed[2]. Since the conventional MUD detectors ignore the prior estimation errors, they are independent of ϵ_k. While the robust detectors require a choice of ϵ_k, they are insensitive to the choice. Notable improvements are observed for all robust detectors relative to their nonrobust counterparts. Compared with the

[2]The results are obtained over 200 independent channel realizations

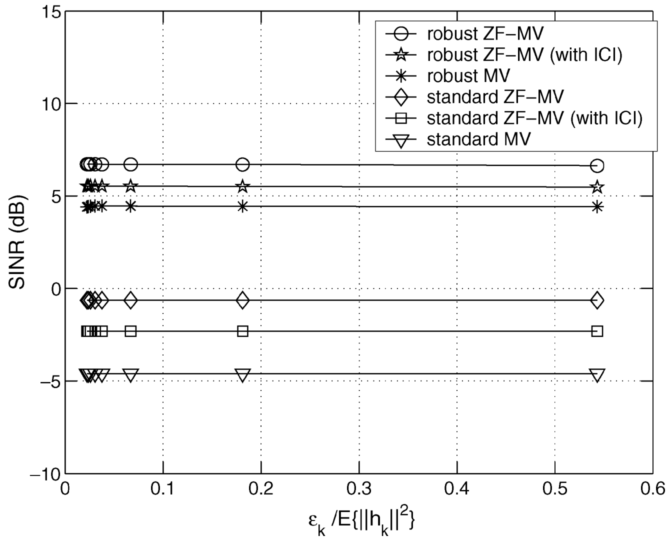


Fig. 1. Receiver output SINR performance versus the normalized channel uncertainty $\epsilon_k/E\{\|\mathbf{h}_k\|^2\}$ when $P = 32, L = 7, K = 10, M = 50$, $\eta/E\{\|\mathbf{R}_y\|\} = 0.2$, and SNR $= 10$ dB.

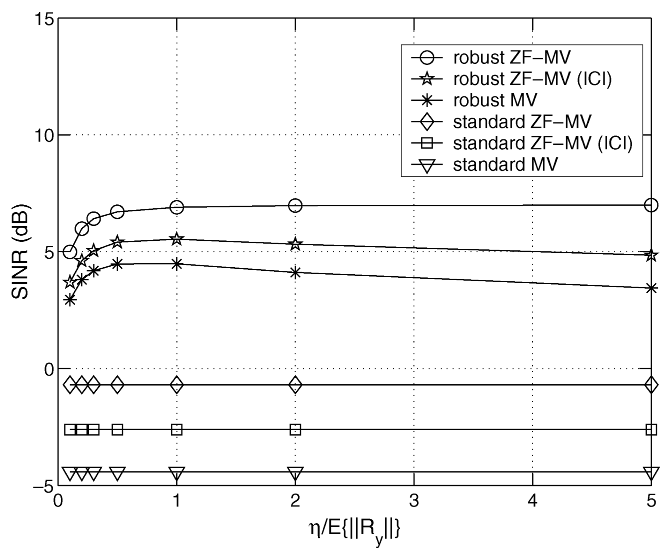


Fig. 2. Receiver output SINR versus the normalized covariance matrix uncertainty $\eta/E\{\|\mathbf{R}_y\|\}$ when $P = 32, L = 7, K = 10, M = 50$, $\epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$, and SNR $= 10$ dB.

standard nonrobust detectors, the robust detectors obtain around 5–8-dB improvements.

We next investigate the impact of the size η of the uncertainty set related to the initial sample covariance matrix estimate [cf. (18)]. We keep the same set of simulation parameters as in the previous example except that $\epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$ and the size of the uncertainty set, $\eta/E\{\|\mathbf{R}_y\|\}$, is varying. The results are shown in Fig. 2. Again, the proposed robust schemes are not sensitive to the selections of η and, compared to the standard nonrobust MUD detectors, obtain significant improvements.

One may wonder how the proposed robust detectors compare to the standard detectors when initial estimates are sufficiently accurate. To find out an answer, we consider a case where the amount of the prior estimation error varies but the assumed size of the uncertainty set is fixed. In particular, Fig. 3(a) depicts the performance as a function of the channel estimation error $E\{\|\Delta\mathbf{h}_k\|^2/\|\mathbf{h}_k\|^2\}$ when $M = 50$, $\epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$, and $\eta/E\{\|\mathbf{R}_y\|\} = 0.2$. It may seem strange that even with perfect channel estimate (i.e., zero estimation error), the robust detectors still outperform the standard ones by over 5 dB. The performance gap is caused by the error in the sample covariance matrix obtained with $M = 50$ symbols. We next replace the sample covariance matrix with the *true* covariance matrix, so that channel estimation error is the only type of error that affects the performance, and repeat the above simulation. The results are shown in Fig. 3(b). It is seen that with zero estimation error, both the robust and standard detectors perform almost identically.[3] Meanwhile, even with small estimation error, the standard detectors degrade quickly, and the gap between the standard and our robust detectors increases as the estimation error increases.

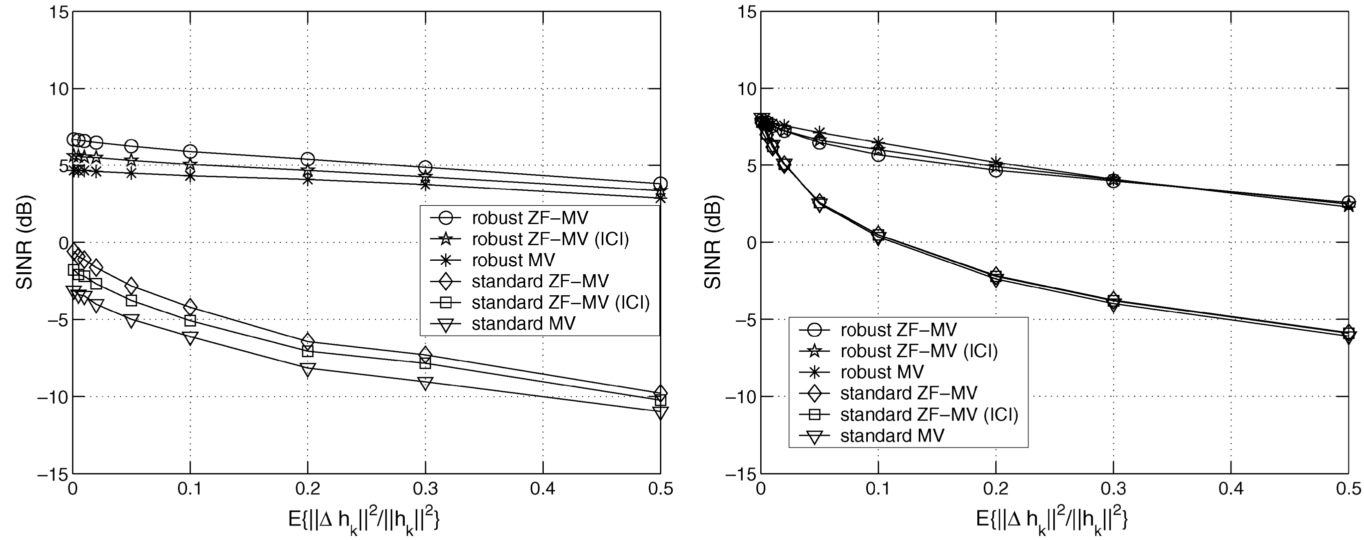
We proceed to compare the output SINR performance of the proposed robust and the standard nonrobust detectors as a function of the input SNR, when the number of training symbols $M = 50$, $\epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$ and $\eta/E\{\|\mathbf{R}_y\|\} = 0.2$. In Fig. 4, it is seen that the robust detectors outperform the standard nonrobust detectors substantially, especially in the higher SNR regime. The robust ZF-MV detector, which utilize the spreading codes and training for all ten users, achieves the best performance. The robust ZF-MV detector with ICI, which utilizes knowledge of six out of all ten users' spreading codes and training, shows some degradation relative to the robust ZF-MV detector without ICI. Finally, the robust MV detector utilizes knowledge of a single user information and, as expected, is the worst of the three robust detectors.

In the last example, we examine the average bit-error rate (BER) performance of all the six detectors versus the input SNR, where we consider adaptive implementations of the proposed and standard detectors. We use $M = 100$ training symbols, while η and ϵ_k are chosen as the same in the previous example. As shown in Fig. 5, when the input SNR varies from 2 to 12 dB, the performance gap between the robust detectors and the nonrobust detectors increases. Moreover, the conventional ZF-MV with ICI and MV detectors have an irreducible error floor due to poor initial channel estimates.

VII. Conclusion

In this paper, we have proposed a class of robust MUD detectors to seek enhanced robustness against prior estimation errors in the initial channel estimate and the sample covariance matrix. Specifically, robust MUD detectors have been obtained by optimizing the worst case performance over two separate bounded uncertainty sets pertaining to the aforementioned estimation errors. We have shown that, although the prior estimation errors are generally unbounded, it is beneficial to optimize the worst case performance over properly chosen bounded uncertainty sets which are determined by a bounding probability. Numerical results show that the proposed robust schemes yield improved performance over those that ignore the prior channel estimation and sample covariance matrix estimation errors.

[3]Note that in this case, the Standard MV detector is optimum since it reduces to the linear MMSE receiver.



(a) (b)

Fig. 3. Receiver output SINR versus the normalized initial channel estimation error $E\{\|\Delta \mathbf{h}_k\|^2/\|\mathbf{h}_k\|^2\}$ when $P = 32, L = 7, K = 10, \epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$, and $\eta/E\{\|\mathbf{R}_y\|\} = 0.2$. (a) Using a sample covariance matrix $\hat{\mathbf{R}}_y$ obtained with $M = 50$; (b) Using the true $\mathbf{R}_y$.

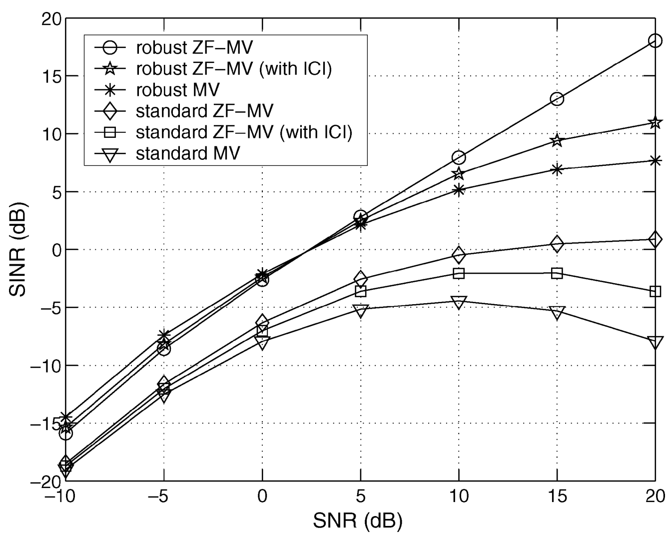


Fig. 4. Receiver output SINR versus the input SNR when $P = 32, L = 7, K = 10, M = 50, \epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$, and $\eta/E\{\|\mathbf{R}_y\|\} = 0.2$.

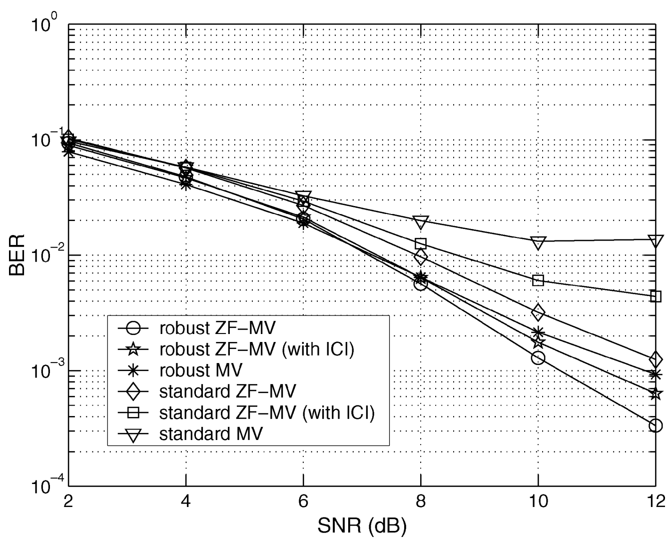


Fig. 5. Average BER versus the input SNR when $P = 32, L = 7, K = 10, M = 100, \epsilon_k/E\{\|\mathbf{h}_k\|^2\} = 0.1$, and $\eta/E\{\|\mathbf{R}_y\|\} = 0.2$.

APPENDIX
DERIVATIONS OF $E\{\gamma\}$ AND $\mathrm{var}\{\gamma\}$

The sample covariance matrix $\hat{\mathbf{R}}_y$ is a unbiased estimate of $\mathbf{R}_y$, i.e., $E\{\hat{\mathbf{R}}_y\} = \mathbf{R}_y$. Let $y_i(m)$ be the ith element of $\mathbf{y}(m)$. The unbiasedness can be easily checked

$$E\{\hat{R}_y(i,j)\} = \frac{1}{M}\sum_{m=1}^{M} E\{y_i(m)y_j^*(m)\} = R_y(i,j). \quad (50)$$

Using (50), we can rewrite the mean of γ as follows:

$$\begin{aligned} &E\{\gamma\} \\ &= \sum_{i=1}^{P}\sum_{j=1}^{P} E\left\{\left[\hat{R}_y(i,j) - R_y(i,j)\right]\left[\hat{R}_y(i,j) - R_y(i,j)\right]^*\right\} \\ &= \sum_{i=1}^{P}\sum_{j=1}^{P} E\{\hat{R}_y(i,j)\hat{R}_y(j,i)\} - R_y(i,j)R_y(j,i). \quad (51) \end{aligned}$$

Note that in (51), we have utilized the fact both $\hat{\mathbf{R}}_y$ and $\mathbf{R}_y$ are Hermitian matrices and, therefore, $\hat{R}_y^*(i,j) = \hat{R}_y(j,i)$ and $R_y^*(i,j) = R_y(j,i)$. Furthermore

$$\begin{aligned} &E\{\hat{R}_y(i,j)\hat{R}_y(j,i)\} \\ &= \frac{1}{M^2}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M} E\left\{y_i(m_1)y_j^*(m_1)y_j(m_2)y_i^*(m_2)\right\} \\ &= \frac{1}{M^2}\sum_{m_1=1}^{M}\sum_{m_2\neq m_1} E\left\{y_i(m_1)y_j^*(m_1)\right\}E\left\{y_j(m_2)y_i^*(m_2)\right\} \\ &\quad + \frac{1}{M^2}\sum_{m_1=1}^{M} E\left\{y_i(m_1)y_j^*(m_1)y_j(m_1)y_i^*(m_1)\right\} \\ &= \frac{M^2 - M}{M^2}R_y(i,j)R_y(j,i) \\ &\quad + \frac{1}{M}\left[R_y(i,j)R_y(j,i) + R_y(i,i)R_y(j,j)\right] \\ &= R_y(i,j)R_y(j,i) + \frac{1}{M}R_y(i,i)R_y(j,j) \quad (52) \end{aligned}$$

$$\begin{aligned} E\{\gamma^2\} &= E\Big\{\sum_{i=1}^{P}\sum_{j=1}^{P} v(i,j)v^*(i,j)\sum_{p=1}^{P}\sum_{q=1}^{P} v(p,q)v^*(p,q)\Big\} \\ &= \sum_{i=1}^{P}\sum_{j=1}^{P}\sum_{p=1}^{P}\sum_{q=1}^{P}\Big[E\{v(i,j)v^*(i,j)\}E\{v(p,q)v^*(p,q)\} + E\{v(i,j)v(p,q)\}E\{v^*(i,j)v^*(p,q)\} \\ &\quad + E\{v(i,j)v^*(p,q)\}E\{v^*(i,j)v(p,q)\}\Big] \end{aligned} \tag{54}$$

$$\begin{aligned} E\{\gamma^2\} &= \frac{1}{M^2}\sum_{i=1}^{P}\sum_{j=1}^{P}\sum_{p=1}^{P}\sum_{q=1}^{P}\Big[R_y(i,i)R_y(j,j)R_y(p,p)R_y(q,q) + |R_y(i,q)R_y(p,j)|^2 + |R_y(i,p)R_y(q,j)|^2\Big] \\ &= \frac{1}{M^2}\left[\mathrm{tr}^4\{\mathbf{R}_y\} + 2\|\mathbf{R}_y\|^4\right] \end{aligned} \tag{58}$$

where we have used the standard result on the fourth-order moment of Gaussian random variables in the second equality [31]. Substituting (52) into (51), we have

$$\begin{aligned} E\{\gamma\} &= \frac{1}{M}\sum_{i=1}^{P}\sum_{j=1}^{P} R_y(i,i)R_y(j,j) \\ &= \frac{1}{M}\mathrm{tr}^2\{\mathbf{R}_y\}. \end{aligned} \tag{53}$$

Next we compute $E\{\gamma^2\}$. Let $v(i,j) \triangleq \hat{R}_y(i,j) - R_y(i,j)$. Then, using the standard result on the fourth-order moment of Gaussian random variables, we have (54), shown at the top of the page, where we have utilized the fact that $E\{v(i,j)\} = E\{\hat{R}_y(i,j)\} - R_y(i,j) = 0$. We first consider the item with the most general case in (54), i.e.,

$$\begin{aligned} &E\{v(i,j)v^*(p,q)\} \\ &= E\left\{\left[\hat{R}_y(i,j) - R_y(i,j)\right]\left[\hat{R}_y(p,q) - R_y(p,q)\right]^*\right\} \\ &= E\{\hat{R}_y(i,j)\hat{R}^*(p,q)\} + R_y(i,j)R^*(p,q) \\ &\quad - R_y(i,j)E\{\hat{R}^*(p,q)\} - E\{\hat{R}_y(i,j)\}R^*(p,q) \\ &= E\{\hat{R}_y(i,j)\hat{R}_y(q,p)\} - R_y(i,j)R_y(q,p). \end{aligned} \tag{55}$$

Here, $E\{\hat{R}_y(i,j)\hat{R}_y(q,p)\}$ can be rewritten as

$$\begin{aligned} &E\{\hat{R}_y(i,j)\hat{R}_y(q,p)\} \\ &= \frac{1}{M^2}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M} E\left\{y_i(m_1)y_j^*(m_1)y_q(m_2)y_p^*(m_2)\right\} \\ &= \frac{1}{M^2}\sum_{m_1=1}^{M}\sum_{m_2\neq m_1} E\left\{y_i(m_1)y_j^*(m_1)\right\}E\left\{y_q(m_2)y_p^*(m_2)\right\} \\ &\quad + \frac{1}{M^2}\sum_{m_1=1}^{M} E\left\{y_i(m_1)y_j^*(m_1)y_q(m_1)y_p^*(m_1)\right\} \\ &= \frac{M^2-M}{M^2}R_y(i,j)R^*(p,q) \\ &\quad + \frac{1}{M}\left[R_y(i,j)R^*(p,q) + R_y(i,p)R_y(q,j)\right] \\ &= R_y(i,j)R_y(q,p) + \frac{1}{M}R_y(i,p)R_y(q,j). \end{aligned} \tag{56}$$

Substituting (56) into (55), we have

$$E\{v(i,j)v^*(p,q)\} = \frac{1}{M}R_y(i,p)R_y(q,j). \tag{57}$$

Based on (57), (54) can be simplified as (58), shown at the top of the page. Using (53) and (58), we have

$$\mathrm{var}\{\gamma\} = E\{\gamma^2\} - E^2\{\gamma\} = \frac{2}{M^2}\|\mathbf{R}_y\|^4. \tag{59}$$

REFERENCES

[1] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, vol. 35, no. 12, pp. 126–133, Dec. 1997.

[2] N. Yee, J.-P. Linnartz, and G. Fettweis, "Multicarrier CDMA in indoor wireless radio networks," in *Proc. IEEE PIMRC*, Yokohama, Japan, Sep. 1993, pp. 109–113.

[3] S. Sigdel, K. M. Ahmed, and R. M. A. P. Rajatheva, "Performance evaluation of multicarrier CDMA uplink system with antenna array and multiuser detection over correlated multipath channel," in *Proc. 2002 IEEE Veh. Technol. Conf.*, vol. 4, Vancouver, BC, Canada, Sep. 24–28, 2002, pp. 1958–1962.

[4] W. Nabhane and H. V. Poor, "Blind joint equalization and multiuser detection in dispersive MC-CDMA/MC-DS-CDMA/MT-CDMA channels," in *Proc. 2002 Military Commun. Conf.*, vol. 2, Aneheim, CA, Oct. 7–10, 2002, pp. 814–819.

[5] F. Verde, "Subspace-based blind multiuser detection for quasi-synchronous MC-CDMA systems," *IEEE Signal Process. Lett.*, vol. 11, pp. 621–624, Jul. 2004.

[6] J. Li, K. B. Letaief, and Z. Cao, "Reduced-complexity MAP-based iterative multiuser detection for coded multicarrier CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1909–1915, Nov. 2004.

[7] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, pp. 853–863, 1974.

[8] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.

[9] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoustics Speech Signal Process.*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.

[10] D. D. Feldman and L. J. Griffiths, "A projection approach to robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 867–876, Apr. 1994.

[11] C. C. Lee and J. H. Lee, "Robust adaptive array beamforming under steering vector errors," *IEEE Trans. Antennas Propag.*, vol. 45, no. 1, pp. 168–175, Jan. 1997.

[12] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, vol. 53, pp. 305–314, Dec. 1999.

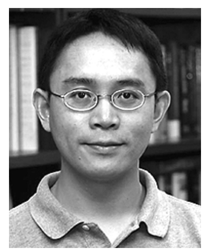
[13] S. Shahbazpanahi, A. B. Gershman, Z. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257–2269, Sep. 2003.

[14] M. H. Er and A. Cantoni, "An alternative formulation for an optimum beamformer with robustness capability," in *IEE Proc. Commun. Radar Signal Process.*, vol. 132, 1985, pp. 447–460.
[15] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
[16] ——, "Doubly constrained robust capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, Sep. 2004.
[17] S. Q. Wu and J. Y. Zhang, "A new robust beamforming method with antennae calibration errors," in *Proc. IEEE Wireless Commun. Networking Conf.*, vol. 2, New Orleans, LA, Sep. 1999, pp. 869–872.
[18] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst case performance optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.
[19] K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z. Luo, "Robust blind multiuser detection based on the worst case performance optimization of the MMSE receiver," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 295–305, Jan. 2005.
[20] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
[21] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "A robust linear receiver for uplink multi-user MIMO systems based on probability-constrained optimization and second-order cone programming," presented at the 3rd IEEE Sensor Array Multichannel Signal Process. Workshop, Barcelona, Spain, Jul. 18–21, 2004.
[22] S. Shahbazpanahi and A. B. Gershman, "Robust blind multiuser detection for synchronous CDMA systems using worst case performance optimization," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 2232–2245, Nov. 2004.
[23] M. L. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 7, pp. 944–960, Jul. 1995.
[24] M. K. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receiver," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3014–3022, Nov. 1998.
[25] Z. Xu, P. Liu, and X. Wang, "Blind multiuser detection: From MOE to subspace methods," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 510–524, Feb. 2004.
[26] J. B. Schodorf and D. B. Williams, "A constrained optimization approach to multiuser detection," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 258–262, Jan. 1997.
[27] X. Wang and A. Host-Madson, "Group-blind multiuser detection for uplink CDMA," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 9, pp. 1971–1984, Nov. 1999.
[28] O. Edfors, M. Sandell, J.-J. van de Beek, D. Landström, and F. Sjöberg, "An introduction to orthogonal frequency-division multiplexing," Luleå Univ. Technol., Luleå, Sweden, Res. Rep./1996:16, 1996.
[29] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[30] H. Li, "Blind channel estimation for multicarrier systems with narrowband interference suppression," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 326–328, Jul. 2003.
[31] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice Hall, 1997.
[32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993.
[33] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
[34] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.
[35] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 1996.

Rensheng Wang (M'05) received the B.E. and M.E. degrees from Harbin Institute of Technology, Harbin, China, in 1995 and 1997, respectively, and the Ph.D. degree from Stevens Institute of Technology, Hoboken, NJ, in 2005, all in electrical engineering.

From 1997 to 2000, he was a Researcher at the Institute of Electronics, Chinese Academy of Sciences, Beijing. He is currently a Research Associate with the Wireless Network and Security Center (WiNSeC) at Stevens Institute of Technology. His research interests lie in the general area of statistical signal processing, communication systems, and networks.

Dr. Wang received the Outstanding Research Award in 2002 and the Edward Peskin Award in 2004 from Stevens Institute of Technology.



Hongbin Li (M'99) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, the University of Florida. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory, Rome, NY, in the summers of 2003 and 2004. Since July 1999, he has been an Assistant Professor in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. His current research interests include wireless communications, statistical signal processing, and radars.

Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award in 1999 from the University of Florida. He has been an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since January 2003 and an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS since January 2005.

Tao Li received the B.S. and M.S. degrees in computer science and the Ph.D. degree in electrical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1986, 1991, and 1995, respectively.

From 1993 to 1994, he was a Visiting Scholar at the University of California, Berkeley. He is currently a Professor with the Department of Computer, Sichuan University, Chengdu, where he founded and directs the Laboratory of Computer Networks and Information Security. His research interests include computer networks, information security, artificial intelligence, neural networks, artificial immune system, computer communications, and wireless communications.