Noncoherent Multiuser Communications: Multistage Detection and Selective Filtering

Rajnish Sinha
WINLAB, ECE Dept.
Rutgers University
sinhar@winlab.rutgers.edu

Aylin Yener
EE Dept.
The Pennsylvania State University
yener@ee.psu.edu

Roy D. Yates
WINLAB, ECE Dept.
Rutgers University
ryates@winlab.rutgers.edu

June 3, 2002

Submitted to EURASIP Journal on Applied Signal Processing, Special Issue on Multiuser Detection and Blind Estimation

Abstract

We consider noncoherent multiuser detection techniques for a system employing nonlinear modulation of non-orthogonal signals. Our aim is to investigate near-optimum noncoherent multiuser detection techniques that utilize the received signal structure while retaining reasonable complexity. Near-optimum approximations of the maximum likelihood detector are investigated where the signal structure is reflected in the approximation techniques explored. Several implementations of noncoherent soft interference cancelers are proposed and investigated, each of which exploits the signal structure in a specific way. We propose a class of detectors that employ *selective* filtering, a technique that exploits the *a priori* information that each user selects one of M signals for transmission. We show that selective filtering offers improved performance over the noncoherent counterparts of the existing near-optimum multiuser detectors. Both deterministic as well as blind adaptive implementations of selective filtering are considered. Numerical comparisons are provided to demonstrate the near-optimum performance of the proposed detectors.

1 Introduction

Nonlinear M-ary modulation with noncoherent detection is often necessary when phase estimation is difficult due to rapid changes in the channel conditions [1]. In a multiuser setting, the correlated waveforms that are used to transmit the users' messages give rise to interference issues since the receiver observes the superposition of all users' transmissions. Similar to its coherent, linear modulation counterpart [2], the maximum likelihood (ML) detector for multiuser noncoherent communications with nonlinear modulation that estimates all users' messages jointly, has prohibitive complexity. The computational complexity of the ML multiuser detector, for the cases of both linear and nonlinear modulation, is a direct result of the fact that the maximum likelihood bit estimate is the optimizer of an NP hard problem for which no efficient solution methodology is known [3].

For noncoherent systems, the complexity of optimal detection has directed attention toward suboptimal interference suppression techniques [4–8]. The pioneering work [4] introduced a pseudo-linear representation in which the signal space is spanned by MK signals corresponding to the M possible messages for each of the K users. This approach led naturally to two-step detectors in which decorrelative [4, 7] or MMSE [6] linear filtering for user separation is followed by noncoherent single-user detection. Alternatively, [5, 8] have employed decision-directed methods that use prior decisions for suppression of the interference. The approach of [5] is to decorrelate against all possible interfering signals. Prior decisions reduce the space of possible interfering signals in that if a decision is made that user k transmitted signal m, then there is no need to decorrelate against the other $M-1$ possible transmissions of user k. In [8], the approach is to decorrelate against the known interfering signals identified by previous decisions. The resulting output is then passed to a bank of single-user matched filters and a maximum magnitude detector to determine another symbol.

In this paper, we follow the general spirit of [5, 8] and examine approaches that combine filtering with decision-directed methods. We propose low-complexity, sub-optimal noncoherent detectors that take advantage of certain *a priori* information available regarding the multiuser signaling. We incorporate this structure into the algorithms of three detector classes: *constrained* detectors, soft interference cancelers, and *selective filtering* detectors [9]. The constrained detectors embed maximum amplitude information for the received signal components as constraints for nonlinear programming relaxations of the ML multiuser detector. In the class of interference cancelers, we explore three variations that arise due to the noncoherent nature of the detection scheme - the noncoherent serial, clipped, and parallel soft-ICs (Soft Interference Cancelers). These soft-IC detectors employ the same fundamental multistage detection approach as their linear modulation and coherent detection counterparts, e.g. [10]. Each of these cancelers embeds the multiuser signal structure in its detection algorithm in a different way. We further improve the performance of the noncoherent multiuser detectors by exploiting additional information and introduce the concept of *selective filters*. The selective filters use the *a priori* information that the desired user selects for transmission only one of the M messages available in its constellation. Unlike the non-selective filters of [4, 6, 7], selective filtering for the desired user attempts to suppress only the possible signals of the interfering users. For the most part, our results will show that the soft-ICs yield better performance than the non-selective decorrelating and MMSE filters, especially in near-far scenarios. To illustrate the feasibility of the selective filters in scenarios where limited information regarding the interferers is available, e.g. a CDMA down-link, a blind adaptive implementation of the selective MMSE detector is also presented.

The rest of this paper is organized as follows: Section 2 establishes notation for the additive-noise, synchronous CDMA system model and discusses the ML detector. Section 3 discusses prior work on non-selective decorrelator and MMSE detectors proposed in the literature. This section also introduces the constrained detectors as well as the noncoherent detectors based on soft-ICs. Section 4 applies selective filtering to some of the detectors discussed in Section 3. This section also discusses a blind adaptive implementation of the selective MMSE detector as well as a successive interference suppression (SIS) scheme. Section 5 discusses the numerical results and concluding remarks are presented in Section 6. The Appendix contains developments for certain results in Sections 3-4.

2 System Model

We consider a synchronous CDMA system with K active users, processing gain N, and a signaling scheme where each user transmits one of M signals. A discrete-time model can be obtained by projecting the received signal onto an N-dimensional orthonormal basis. Using the pseudo-linear representation introduced in [4], we view the signal space as being an expanded signal space spanned by the MK signals: M messages for each of the K users. We concentrate on cases where the possible waveforms for all messages of all users are linearly independent. The channel is assumed to be AWGN, and the receiver observes a superposition of the K signals.

For user k, the $N\times 1$ vector $\mathbf{s}_{k,m}$ denotes the signature corresponding to message m while the $N\times M$ matrix $\mathbf{S}_k \stackrel{\Delta}{=} [\mathbf{s}_{k,1}\cdots\mathbf{s}_{k,M}]$ denotes the signature matrix. It is assumed that the signatures in $\mathbf{S}_k$ are complex-valued, have unit norm, and are time limited. The amplitude and phase of message m of user k are denoted by $A_{k,m}$ and $\phi_{k,m}$, with corresponding $M\times M$ diagonal matrices $\mathbf{A}_k \stackrel{\Delta}{=} \text{diag}[A_{k,1},\ldots,A_{k,M}]$ and $\Phi_k \stackrel{\Delta}{=} \text{diag}[e^{j\phi_{k,1}},\ldots,e^{j\phi_{k,M}}]$. The phases are assumed to be independent and uniformly distributed over $[0,2\pi]$. Let m_k be the transmitted message of user k. We define the vector $\mathbf{b}_k = [b_{k,1}\cdots b_{k,M}]^\top$ where

$$b_{k,m} = \begin{cases} 1 & m = m_k \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

We note that $\mathbf{b}_k$ belongs to the set $F = \{\mathbf{e}_1,\ldots,\mathbf{e}_M\}$ where $\mathbf{e}_k \in \{0,1\}^M$ has kth entry $e_{k,k} = 1$ and zero for all other entries. It is assumed that the M messages of a user are equiprobable. The received vector at the output of the bank of chip-matched filters can be written as

$$\mathbf{r} = \sum_{k=1}^{K} \mathbf{S}_k\mathbf{A}_k\Phi_k\mathbf{b}_k + \mathbf{n} \tag{2}$$

where $\mathbf{n}$ is the AWGN vector. Further, $\mathbf{r}$ can be expressed in terms of the MK-length vector $\mathbf{b} = [\mathbf{b}_1^\top\cdots\mathbf{b}_K^\top]^\top$, the $N\times MK$ matrix $\mathbf{S} \stackrel{\Delta}{=} [\mathbf{S}_1\cdots\mathbf{S}_K]$, the $MK\times MK$ matrices $\mathbf{A} \stackrel{\Delta}{=} \text{diag}[\mathbf{A}_1,\ldots,\mathbf{A}_K]$ and $\Phi \stackrel{\Delta}{=} \text{diag}[\Phi_1,\ldots,\Phi_K]$ as

$$\mathbf{r} = \mathbf{S}\mathbf{A}\Phi\mathbf{b} + \mathbf{n}. \tag{3}$$

The aim of the multiuser detector is to recover the message vector $\mathbf{b} \in F^K$. For a given $\mathbf{b}$, let the vector $\phi = [e^{j\phi_1}\cdots e^{j\phi_K}]^\top$ represent the phases corresponding to the K nonzero entries of $\mathbf{b}$. With known $\mathbf{A}$ and Φ, the ML estimate of $\mathbf{b}$ given $\mathbf{r}$, is the solution to the *optimum* multiuser detector [2]. The estimate may be written as

$$\hat{\mathbf{b}} = \arg\max_{\mathbf{b}\in F^K} f(\mathbf{r}|\mathbf{b},\phi,\mathbf{A}) \tag{4}$$

In the AWGN channel, the optimization (4) becomes the familiar distance minimization problem

$$\hat{\mathbf{b}} = \arg\min_{\mathbf{b}\in F^K} ||\mathbf{r}-\mathbf{S}\mathbf{A}\Phi\mathbf{b}||^2. \tag{5}$$

Note that (4) and (5) describe a *coherent* detector since knowledge of Φ is assumed.

Next, consider the case where the amplitudes, $\mathbf{A}$, are known at the receiver as in (4), but both Φ and $\mathbf{b}$ are unknown. Since each element of ϕ must belong to the unit circle C_1, the joint ML estimate of $\mathbf{b}$ and ϕ is

$$(\hat{\mathbf{b}}, \hat{\phi}) = \arg \max_{\mathbf{b} \in F^K} \max_{\phi \in C_1^K} f(\mathbf{r}|\mathbf{b}, \phi, \mathbf{A}). \tag{6}$$

The implementation of this detector requires an exhaustive search over possible vectors $\mathbf{b}$. Further, since each of the elements of ϕ lie on a unit circle, the inner maximization in (6) is over a non-convex set and hence there is no guarantee of finding the global minimum in (5). However, relaxing the constraints and allowing each of the elements of ϕ to lie on the unit disk $\overline{C}_1$ yields a convex set for the inner optimization. The resulting detector is

$$(\hat{\mathbf{b}}, \hat{\phi}) = \arg \max_{\mathbf{b} \in F^K} \max_{\phi \in \overline{C}_1^K} f(\mathbf{r}|\mathbf{b}, \phi, \mathbf{A}). \tag{7}$$

This detector will be referred to as the *joint* detector. The joint detector effectively assumes that $\mathbf{A}$ characterizes the maximum amplitudes of the signals. Both the detector (6) and the joint detector (7) are generalized likelihood ratio test (GLRT) detectors that differ in their assumptions regarding the received signal amplitudes. In particular, when all elements of $\mathbf{A}$ become large, the maximum amplitude constraint of the joint detector becomes trivial and the joint detector approaches the GLRT detector in [8] which treats the signal amplitudes as unknown.

3 Non-selective Detection

Recently, several detection methods with reasonable complexity have been formulated that *approximate* the solution of the NP hard ML multiuser detection problem [11, 12]. Further results using nonlinear programming techniques to approximate the ML multiuser detector for linear modulation can be found in [13–15]. In this section, similar to the linear modulation counterparts considered in [11–13], we *relax* the constraint set of the ML multiuser detection problem. We represent the structure of the signal in the form of a constraint set and explore various detectors with the same objective function yet different constraint sets.

3.1 Prior Work

To place the constrained multiuser detectors in proper context, we start by examining the decorrelative and the MMSE two-stage detectors proposed in [4, 6, 7]. These detectors combine linear filtering first stages with single user detection second stages. To describe these detectors, let $\mathbf{z} = \widehat{\mathbf{Ax}}$ denote the estimate of the desired vector $\mathbf{Ax}$ and let the output of the matched filters be

$$\mathbf{y} = \mathbf{S}^H \mathbf{r} = \mathbf{RAx} + \mathbf{S}^H \mathbf{n} \tag{8}$$

where $\mathbf{R} = \mathbf{S}^H \mathbf{S}$ is the cross-correlation matrix. The first stage of the decorrelative detector [4] applies the decorrelating filter $\mathbf{R}^{-1}$ to $\mathbf{y}$ to obtain

$$\mathbf{z} = \mathbf{R}^{-1} \mathbf{y} = \mathbf{Ax} + \mathbf{R}^{-1} \mathbf{S}^H \mathbf{n}. \tag{9}$$

If the signals are linearly dependent, we can replace $\mathbf{R}^{-1}$ by the Moore-Penrose generalized inverse [16]; however, for simplicity, we will assume that the signals are linearly independent.

For the MMSE filter, we apply the matrix transformation $\mathbf{C}^H$ to the output $\mathbf{r}$ to obtain the estimate $\mathbf{z} = \widehat{\mathbf{Ax}} = \mathbf{C}^H\mathbf{r}$ that minimizes the mean-squared error $E\left[||\mathbf{C}^H\mathbf{r} - \mathbf{Ax}||^2\right]$. The solution is given in [6] as

$$\mathbf{C} = \mathbf{H}^{-1}\mathbf{SE} \tag{10}$$

where $\mathbf{H} = E[\mathbf{rr}^H] = \mathbf{SES}^H + \sigma^2\mathbf{I}_N$, $\mathbf{E} = E[\mathbf{Axx}^H\mathbf{A}] = (1/M)\mathbf{A}^2$, and $\mathbf{I}_n$ is the identity matrix of dimension n. Equivalently, if the MMSE filter is applied to the matched filter output $\mathbf{y}$ in (8) instead of $\mathbf{r}$, then $\mathbf{z} = \bar{\mathbf{C}}^H\mathbf{y}$, where

$$\bar{\mathbf{C}} = (\mathbf{R} + \sigma^2\mathbf{E}^{-1})^{-1}. \tag{11}$$

Note that in case of linear modulation, $\mathbf{E} = \mathbf{A}^2$ and the familiar expression $\mathbf{z} = (\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1}\mathbf{y}$ is obtained [2, 17].

For both decorrelative and MMSE filtering, the filter output $\mathbf{z}$ is an MK-length vector that is the concatenation $\mathbf{z} = [\mathbf{z}_1^\top, \ldots, \mathbf{z}_K^\top]^\top$ of M-length component vectors $\mathbf{z}_k$. To describe decision rules for particular users, we adopt the general convention that $\mathbf{z}_k = [x_{M(k-1)+1}, \ldots, x_{Mk}]^\top$ denotes *the kth vector component* of an MK-length vector $\mathbf{z}$. To address the specific elements of $\mathbf{z}_k$, we write $\mathbf{z}_k = [z_{k,1}, \ldots, z_{k,M}]^\top$.

In the second stage, references [6, 7] employ the kth component vector $\mathbf{z}_k$ as a decoupled decision statistic to obtain an estimate $\hat{m}_k$ of the kth user's message. The simplest such method is the maximum magnitude (MM) rule, denoted $\mu(\mathbf{z}_k)$, and defined by

$$\hat{m}_k = \mu(\mathbf{z}_k) \stackrel{\Delta}{=} \arg \max_{m\in\{1,\ldots,M\}} |z_{k,m}|^2. \tag{12}$$

In the event of ties, the MM rule arbitrarily chooses one of the maximizing entries. For orthogonal signaling over a single user AWGN channel, the MM rule is optimum; however, since the decorrelative and MMSE filters introduce correlation in the additive noise and/or interference components of $\mathbf{z}_k$, the MM rule is merely a heuristic. Single user decoding rules for user k that exploit the correlation structure are developed in [4, 6].

3.2 Constrained Noncoherent Multiuser Detection

Our starting point is the ML detector (6) in which the amplitudes $\mathbf{A}$ are known but the symbols $\mathbf{b}$ and phases ϕ, or equivalently Φ, are unknown. In this case, we estimate them jointly as $\mathbf{x} = \Phi\mathbf{b}$. We define the set

$$G = \left\{e^{j\phi}\mathbf{e} | \mathbf{e} \in F, 0 \le \phi \le 2\pi\right\} \tag{13}$$

and observe that $\mathbf{x} \in G^K$. Rewriting (6), the jointly optimal estimate is

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}\in G^K} ||\mathbf{r} - \mathbf{SAx}||^2 \tag{14}$$

We observe that the minimization (14) is difficult because G^K is a non-convex constraint set. Due to the high complexity associated with the ML detector, reduced complexity approximations can be obtained by solving a relaxation of the original problem [18]. If we relax the constraint set such that the new constraint set is convex, then the optimizer of the quadratic objective function in (14) can be found efficiently via a variety of nonlinear programming methods. This observation is the key towards formulation of the approximate solutions presented in the remainder of this section.

We start with the case where the vector $\mathbf{x}$ containing *all* the users' messages is constrained. Since the constraint $\mathbf{x}^H\mathbf{x} = K$ is non-convex, a relaxation of the form $\mathbf{x}^H\mathbf{x} \leq K$ results in a convex set. The estimate $\hat{\mathbf{x}}$ is the solution to the optimization problem

$$\begin{aligned} \text{minimize} \quad & ||\mathbf{r} - \mathbf{S}\mathbf{A}\mathbf{x}||^2 \\ \text{subject to} \quad & ||\mathbf{x}||^2 \leq K. \end{aligned} \tag{15}$$

The convex set $||\mathbf{x}||^2 \leq K$ can be thought of as the interior of an MK dimensional hypersphere of radius $\sqrt{K}$. The solution to the above problem, derived in [11,12] in the context of linear modulation, is the *generalized MMSE* detector

$$\hat{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{R} + \lambda^* \mathbf{A}^{-2})^{-1}\mathbf{y} \tag{16}$$

where λ^* is the optimum Lagrange multiplier corresponding to the global constraint (15). Note that (16) reduces to the MMSE solution [17] for $\lambda^* = \sigma^2$ the variance of the AWGN process. We apply the MM rule to the filter output $\hat{\mathbf{x}}$ to obtain the symbl decision $\hat{m}_k = \mu(\hat{\mathbf{x}}_k)$. The resulting detector, consisting of the filter (16) followed by the MM rule, will be referred to as the *global* constrained detector.

Now we consider local constraints for each individual user k. Recalling that $\mathbf{x} = \Phi\mathbf{b}$, the kth vector component of $\mathbf{x}$ is $\mathbf{x}_k = \Phi_k\mathbf{b}_k$. Since $\mathbf{x}_k \in G$, $\mathbf{x}_k^H\mathbf{x}_k = 1$ for all k. If we relax the local constraint $\mathbf{x}_k^H\mathbf{x}_k = 1$ to be the convex set $\mathbf{x}_k^H\mathbf{x}_k \leq 1$, which represents the interior of an M dimensional hypersphere of unit radius,, then the estimate $\hat{\mathbf{x}}$ is the solution to

$$\begin{aligned} \text{minimize} \quad & ||\mathbf{r} - \mathbf{S}\mathbf{A}\mathbf{x}||^2 \\ \text{subject to} \quad & ||\mathbf{x}_k||^2 \leq 1 \qquad k = 1, \ldots, K. \end{aligned} \tag{17}$$

The solution to (17) is derived in Appendix A to be

$$\hat{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{R} + \Lambda^* \mathbf{A}^{-2})^{-1}\mathbf{y} \tag{18}$$

where Λ^* is a diagonal matrix containing the Lagrange multipliers. We then apply the MM rule described in (12) to the kth component vector $\hat{\mathbf{x}}_k$ to obtain $\hat{m}_k = \mu(\hat{\mathbf{x}}_k)$. This detector will be referred to as the *local* constrained detector. Note that the local constrained detector is not the same as the joint detector (7). Although both detectors are obtained by enforcing a maximum amplitude constraint on each user k, the joint detector searches only over vectors $\mathbf{x}$ for which each component vector $\mathbf{x}_k$ is of the form $ae^{j\phi_k}\mathbf{b}_k$ where $\mathbf{b}_k \in F$ and $0 \leq a \leq 1$.

Note that there may be other sub-optimal schemes with more constraints that yield better performance with lower complexity compared to the detectors proposed here. Also, it is not clear whether using a more properly constrained search space with more limited search is better than the expanded search space we have considered with virtually all magnitudes and phases that satisfy a maximum energy bound (local and global constraints). These issues require further research.

3.3 Soft Interference Cancelers

Multistage detectors, also referred to as *multistage interference cancelers*, fall in the class of decision directed multiuser detectors and are viable alternatives to popular linear detectors such as the decorrelator and MMSE, due to their excellent BER performance and reasonably low complexity [2]. Several versions of multistage detectors, serial and parallel implementations, as well as versions where the bit estimates of the interferers can be hard or soft, have been proposed in the literature for linear modulation and coherent detection [10, 12, 19–23]. The contributions of this section are twofold. First, the detectors proposed here are noncoherent realizations of the decision directed, nonlinear detectors proposed in [10, 22]. Second, new techniques are proposed to incorporate the signal structure into the decision algorithms. In particular, we propose three detectors: the serial soft-IC, the clipped soft-IC, and the parallel soft-IC.

In this section, a stage refers to a single pass through the detectors of all users. All implementations here use the decorrelator outputs in the first stage, followed by multiple stages of processing of these outputs. The goal once again is to obtain $\hat{\mathbf{x}}$, the vector that contains the estimates of all users' transmitted messages. To obtain the estimate $\hat{\mathbf{x}}_k$ for the kth user's message, soft estimates are used to reconstruct the interference that is then subtracted from that user's matched filter output. The differences between these detectors will arise in their implementation, i.e. serial or parallel, as well as the type of decisions that are communicated between the users' detectors.

In the serial soft-IC detector, the first step is to determine sequentially the estimates $\tilde{x}_{k,1}, \ldots, \tilde{x}_{k,M}$ of the M possible messages of user k. In the second step, only the entry $\tilde{x}_{k,m}$ with the largest magnitude is retained while the other $M-1$ entries are forced to 0. This estimated-and-mapped vector for user k is denoted by $\hat{\mathbf{x}}_k$. The mapping ensures that $\hat{\mathbf{x}}_k$ has a structure similar to that of $\mathbf{x}_k$. Following from (8), we verify in Appendix B that the estimate for message m of user k is

$$\begin{aligned}\tilde{x}_{k,m} = \frac{y_{k,m}}{A_{k,m}} &- \frac{1}{A_{k,m}} \sum_{i=1}^{k-1} \sum_{j=1}^{M} \mathbf{s}_{k,m}^H \mathbf{s}_{i,j} A_{i,j} \hat{x}_{i,j} - \frac{1}{A_{k,m}} \sum_{j=1}^{m-1} \mathbf{s}_{k,m}^H \mathbf{s}_{k,j} A_{k,j} \tilde{x}_{k,j} \\ &- \frac{1}{A_{k,m}} \sum_{j=m+1}^{M} \mathbf{s}_{k,m}^H \mathbf{s}_{k,j} A_{k,j} x_{k,j} - \frac{1}{A_{k,m}} \sum_{i=k+1}^{K} \sum_{j=1}^{M} \mathbf{s}_{k,m}^H \mathbf{s}_{i,j} A_{i,j} x_{i,j}. \end{aligned} \tag{19}$$

where the components on the right side of (19) are: (from left to right) the matched filter output, the estimates of the previous $k-1$ users' messages, the previously detected estimates of messages of user k, the not-yet-detected messages of user k, and the not-yet-detected messages of the other users. After the M entries of user k are determined, the estimated vector $\tilde{\mathbf{x}}_k$ is then mapped to $\hat{\mathbf{x}}_k$ using the maximum magnitude rule:

$$\hat{\mathbf{x}}_k = \mathbf{e}_{\mu(\tilde{\mathbf{x}}_k)} \tilde{x}_{k,\mu(\tilde{\mathbf{x}}_k)}. \tag{20}$$

This vector estimate $\hat{\mathbf{x}}_k$ is then used by user $k+1$ in (19) above for estimating its vector, and so on. The whole procedure mentioned above can then be repeated for multiple stages to refine the estimates.

The implementation of the clipped soft-IC detector employs the same first step (19). In the second step, however, we incorporate the relaxed constraint $|x_{k,i}| \leq 1$ by clipping in accordance with the following rule:

$$\hat{x}_{k,i} = \begin{cases} \tilde{x}_{k,i} & |\tilde{x}_{k,i}| \leq 1 \\ \tilde{x}_{k,i}/|\tilde{x}_{k,i}| & \text{otherwise.} \end{cases} \tag{21}$$

Thus, the difference between the serial soft-IC and the clipped soft-IC lies in the type of decision fed between the users.

Lastly, the parallel soft-IC differs from the serial soft-IC only in the first step determination of the kth user's estimated vector $\tilde{\mathbf{x}}_k$. Instead of serial estimation of each element $\tilde{x}_{k,m}$ in the first step, the parallel soft-IC estimates all elements of $\tilde{\mathbf{x}}_k = [\tilde{x}_{k,1}, \ldots, \tilde{x}_{k,M}]^\top$ in parallel. From (2), we can write the received signal $\mathbf{r}$ in terms of the components $\mathbf{x}_k = \Phi_k \mathbf{b}_k$ as

$$\mathbf{r} = \sum_{j=1}^{K} \mathbf{S}_j \mathbf{A}_j \mathbf{x}_j + \mathbf{n}. \tag{22}$$

The matched filter vector output for user k is

$$\mathbf{y}_k = \mathbf{S}_k^H \mathbf{r} = \mathbf{R}_{kk}\mathbf{A}_k\mathbf{x}_k + \sum_{j\neq k} \mathbf{R}_{kj}\mathbf{A}_j\mathbf{x}_j + \mathbf{S}_k^H \mathbf{n}, \tag{23}$$

where $\mathbf{R}_{kj} = \mathbf{S}_k^H \mathbf{S}_j$. Therefore, $\mathbf{x}_k$ can be estimated as

$$\tilde{\mathbf{x}}_k = (\mathbf{R}_{kk}\mathbf{A}_k)^{-1} \left(\mathbf{y}_k - \sum_{j=1}^{k-1} \mathbf{R}_{kj}\mathbf{A}_j\hat{\mathbf{x}}_j - \sum_{j=k+1}^{K} \mathbf{R}_{kj}\mathbf{A}_j\mathbf{x}_j \right). \tag{24}$$

The components on the right side of (24) are: (from left to right) the matched filter output, the $k-1$ processed users with their estimated-and-mapped vectors $\hat{\mathbf{x}}_j$, and the users that are yet to be processed. In the second step, the parallel soft-IC obtains the users' message decisions using the same maximum magnitude mapping rule (20) as the serial soft-IC detector.

Since the serial soft-IC estimates message elements $x_{k,m}$ one at a time, its granularity is finer than that of the parallel soft-IC which estimates the entire vector $\mathbf{x}_k$ in one step. Hence, it is to be expected that the serial soft-IC will perform slightly better. Also, note that the serial and parallel soft-ICs can be implemented without the knowledge of the individual amplitudes $A_{k,m}$. Instead of estimating just $x_{k,m}$ in (19), the element $A_{k,m}x_{k,m}$ can be jointly estimated, followed by (20). Since the maximum-magnitude rule uses only the *relative* magnitudes, the individual amplitudes do not have to be known explicitly. It is easy to observe from (19) and (21) that this is not the case for the *clipped* soft-IC for which amplitude values must be known.

4 Selective Filtering

To detect whether user k transmitted message m the non-selective detectors of Section 3 consider all possible signals of interferers $j \neq k$, as well as the other $M-1$ possible signals of user k as sources of interference. However, it is known *a priori* that user k transmits precisely one of its M messages. Therefore, for $m \in \{1,\ldots,M\}$, one and only one of the $x_{k,m}$ is nonzero for each user $k \in \{1,\ldots,K\}$. Selective filtering makes use of this observation. Note that if the desired user's signatures (associated with the M messages) are mutually orthogonal, then the selective and non-selective detectors for this user yield identical performance. In this section, we will examine selective implementations of the decorrelator, the MMSE detector followed by its blind implementation, and soft interference canceler for the desired user. To further enhance the performance of the selective detectors, a successive interference suppression (SIS) scheme is also proposed.

In the following, $d(i)$ will denote the ith element of a vector $\mathbf{d}$, while $D(i,j)$ and $D(i,:)$ will denote the (i,j)th element and ith row of a matrix $\mathbf{D}$, respectively. For notational convenience, all vectors and matrices associated with selective filtering will be denoted by a bar above the entry. Assume $k=1$ to be the desired user. Then, in selective detection of message $x_{1,m}$, first the vector $\overline{\mathbf{y}}_m$ is constructed which only contains $y_{1,m}$ and the $M(K-1)$ entries of $\mathbf{y}$ belonging to the interferers. Similarly, the selective signature set $\overline{\mathbf{S}}_m$ consists of $\mathbf{s}_{1,m}$ and the $M(K-1)$ signatures corresponding to all the interferers' signals. The matrix $\overline{\mathbf{A}}_m$ may be interpreted in a similar manner.

4.1 Selective Decorrelator

To formulate the selective decorrelator, we define H_m as the hypothesis that user 1 transmitted signal $\mathbf{s}_{1,m}$. Our problem is to determine which hypothesis among $\{H_1,\ldots,H_M\}$ is correct. From (22), the received signal under hypothesis H_m is

$$H_m: \quad \mathbf{r} = \mathbf{s}_{1,m} A_{1,m} x_{1,m} + \sum_{k=2}^{K} \mathbf{S}_k \mathbf{A}_k \mathbf{x}_k + \mathbf{n} \tag{25}$$

and the decorrelating transformation to suppress all users $k \neq 1$ is given by $(\bar{\mathbf{S}}_m^H \bar{\mathbf{S}}_m)^{-1} \bar{\mathbf{S}}_m^H$ [2]. Thus we first construct

$$\bar{\mathbf{y}}_m = \bar{\mathbf{S}}_m^H \mathbf{r} = \bar{\mathbf{S}}_m^H \mathbf{S}\mathbf{A}\mathbf{x} + \bar{\mathbf{S}}_m^H \mathbf{n} \tag{26}$$

followed by the *selective* transformation

$$\bar{\mathbf{z}}_m = \bar{\mathbf{R}}_m^{-1} \bar{\mathbf{y}}_m = \bar{\mathbf{R}}_m^{-1} \bar{\mathbf{S}}_m^H \mathbf{S}\mathbf{A}\mathbf{x} + \bar{\eta}_m \tag{27}$$

where $\bar{\mathbf{R}}_m = \bar{\mathbf{S}}_m^H \bar{\mathbf{S}}_m$, and $\bar{\eta}_m = \bar{\mathbf{R}}_m^{-1} \bar{\mathbf{S}}_m^H \mathbf{n}$ is zero-mean Gaussian with covariance matrix $\sigma^2 \bar{\mathbf{R}}_m^{-1}$. Since only the first entry of $\bar{\mathbf{z}}_m$ contains the information regarding user 1's information, we construct a new vector $\mathbf{v}$ consisting of only the first entries of the M selective filter outputs, $\bar{\mathbf{z}}_1, \cdots, \bar{\mathbf{z}}_M$, as follows

$$\mathbf{v} = \begin{bmatrix} \bar{\mathbf{z}}_1(1) \\ \vdots \\ \bar{\mathbf{z}}_M(1) \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_1^{-1}(1,:)\bar{\mathbf{S}}_1^H \mathbf{S}\mathbf{A}\mathbf{x} + \bar{\eta}_1(1) \\ \vdots \\ \bar{\mathbf{R}}_M^{-1}(1,:)\bar{\mathbf{S}}_M^H \mathbf{S}\mathbf{A}\mathbf{x} + \bar{\eta}_M(1) \end{bmatrix} = \mathbf{U}\mathbf{S}\mathbf{A}\mathbf{x} + \bar{\eta} \tag{28}$$

where the matrix $\mathbf{U}$ consists of $\bar{\mathbf{R}}_m^{-1}(1,:)\bar{\mathbf{S}}_m^H$ as its mth row. Note that $\mathbf{v}$ is also the estimate $\widehat{\mathbf{A}_1\mathbf{x}_1}$. Therefore, the MM rule is applied to $\mathbf{v}$ to obtain $\hat{m}_1$.

4.2 Selective MMSE Detector

The MMSE detector is popular due to its amenability to adaptive implementation. Blind adaptive implementations of detectors are useful since they only require the signature and timing of the desired user. They are especially attractive for the CDMA downlink, where due to the dynamic environment, it may be difficult for a mobile user to obtain accurate information regarding signatures and timings of other active users in the system [24–26]. In this section, we will first discuss the selective version of the MMSE filter (10) and then formulate its blind adaptive implementation.

The selective MMSE filter for user 1 is obtained similarly to the decorrelator, by applying an MMSE transformation to the received signal (25), under each hypothesis $m \in \{1, \ldots, M\}$. From (10), the selective MMSE filter corresponding to the mth signature is

$$\bar{\mathbf{C}}_m = \bar{\mathbf{H}}_m^{-1} \bar{\mathbf{S}}_m \bar{\mathbf{E}}_m \tag{29}$$

where $\bar{\mathbf{E}}_m = (1/M)\bar{\mathbf{A}}_m^2$ and $\bar{\mathbf{H}}_m = \bar{\mathbf{S}}_m \bar{\mathbf{E}}_m \bar{\mathbf{S}}_m^H + \sigma^2 \mathbf{I}_N$. The filter vector $\bar{\mathbf{c}}_{1,m}$ corresponding to the mth signature of user 1 is the first column of $\bar{\mathbf{C}}_m$, i.e. $\bar{\mathbf{c}}_{1,m} = \bar{\mathbf{C}}_m(:,1)$.

We will now discuss a blind adaptive implementation of (29) above. A blind adaptive implementation of the noncoherent non-selective MMSE detector was proposed in [6]. We extend the algorithm developed in [6] to implement a blind adaptive version of the selective MMSE detector. Since user 1 is the one of interest, the filter coefficients of only this user will be adaptively varied. Representing the mth diagonal entry of $\mathbf{E}$ by E_m, the filter vector $\mathbf{c}_{1,m}$ corresponding to the mth signature of user 1 can be obtained as

$$\mathbf{c}_{1,m} = (\bar{\mathbf{S}}_m \bar{\mathbf{E}}_m \bar{\mathbf{S}}_m^H + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{s}_{1,m} E_m. \tag{30}$$

Note that $\mathbf{c}_{1,m}$ corresponds to the first column of $\bar{\mathbf{C}}_m$. If we denote

$$\bar{\mathbf{r}}_m = \bar{\mathbf{S}}_m \bar{\mathbf{A}}_m \bar{\mathbf{x}}_m + \mathbf{n} \tag{31}$$

then,

$$\mathbf{c}_{1,m} = \left(E[\bar{\mathbf{r}}_m \bar{\mathbf{r}}_m^H]\right)^{-1} \mathbf{s}_{1,m} E_m. \tag{32}$$

Note that in the non-selective version in [6], the filter (32) involves the term $\mathbf{r}\mathbf{r}^H$ which is readily available. In contrast, here, only a subset of that information $\bar{\mathbf{r}}_m$ is needed which cannot be obtained explicitly due to a lack of knowledge of the signature set $\bar{\mathbf{S}}_m$. This problem can be circumvented by writing $\bar{\mathbf{r}}_m$ as

$$\begin{aligned} \bar{\mathbf{r}}_m &= \mathbf{r} - \sum_{i \neq m} \mathbf{s}_{1,i} A_{1,i} x_{1,i} && (33) \\ &= \mathbf{r} - \tilde{\mathbf{S}}_m \tilde{\mathbf{A}}_m \tilde{\mathbf{x}}_m && (34) \end{aligned}$$

where $\tilde{\mathbf{S}}_m$ is the signature matrix of user 1 without the mth signature. Hence, its dimension is $N \times (M-1)$. The terms $\tilde{\mathbf{A}}_m$, $\tilde{\mathbf{E}}_m$ and $\tilde{\mathbf{x}}_m$ may be interpreted in a similar manner. It can be shown that

$$E[\bar{\mathbf{r}}_m \bar{\mathbf{r}}_m^H] = E[\mathbf{r}\mathbf{r}^H] - \tilde{\mathbf{S}}_m \tilde{\mathbf{E}}_m \tilde{\mathbf{S}}_m^H. \tag{35}$$

Since the receiver knows the signatures of the desired user, it can construct $\tilde{\mathbf{S}}_m \tilde{\mathbf{E}}_m \tilde{\mathbf{S}}_m^H$ and extract $\bar{\mathbf{r}}_m$ from the received signal $\mathbf{r}$. Extending the stochastic gradient algorithm in [6], the adaptation for the mth filter vector may then be expressed as

$$\mathbf{c}_{1,m}[n+1] = \mathbf{c}_{1,m}[n] - \mu \left[(\mathbf{r}[n]\mathbf{r}^H[n] - \tilde{\mathbf{S}}_m \tilde{\mathbf{E}}_m \tilde{\mathbf{S}}_m^H) \mathbf{c}_{1,m}[n] - E_m \mathbf{s}_{1,m} \right]. \tag{36}$$

We use the Normalized Squared Error (NSE) criterion [6] to study the convergence properties of the filter coefficients. The NSE at the nth iteration is defined as

$$\text{NSE}[n] = \frac{1}{M} \sum_{m=1}^{M} \frac{||\bar{\mathbf{c}}_{1,m} - \mathbf{c}_{1,m}[n]||^2}{||\bar{\mathbf{c}}_{1,m}||^2}. \tag{37}$$

Note that since the structures of the non-selective and selective MMSE detectors are similar, the convergence analysis of the former [6] can be easily extended to the latter to obtain the upper bound on the step size μ to ensure convergence.

4.3 Selective Soft-IC

Next, we consider the selective implementation of the serial soft-IC scheme described in subsection 3.3. Here, $\hat{x}_{k,m}$ is obtained through a *selective* version of (19), namely

$$\tilde{x}_{k,m} = \frac{1}{A_{k,m}} \left(y_{k,m} - \sum_{i=1}^{k-1} \sum_{j=1}^{M} \mathbf{s}_{k,m}^H \mathbf{s}_{i,j} A_{i,j} \hat{x}_{i,j} - \sum_{i=k+1}^{K} \sum_{j=1}^{M} \mathbf{s}_{k,m}^H \mathbf{s}_{i,j} A_{i,j} x_{i,j} \right) \tag{38}$$

followed by (20). Note that in going from (19) to (38) above, selective filtering has suppressed the terms containing the other $M-1$ messages of user k. Once all M soft-outputs of user k are obtained in this manner, the MM rule is applied to obtain the message decision $\hat{m}_k$.

4.4 Selective Filtering with Successive Interference Suppression

Although it is expected that the selective filters shall yield performance improvements over their non-selective counterparts, further improvements are possible through the use of successive decisions. We call the resulting technique selective filtering with successive interference suppression (SIS). For a user whose message has already been decoded, we need only to suppress the signal corresponding to the decoded message. This is analogous to the successive interference cancelation (SIC) scheme in [27] where a decoded user's signal is reconstructed and explicitly subtracted from the received signal $\mathbf{r}$. The algorithm for SIS is as follows:

1. Corresponding to the M messages of user i, select the maximum magnitude matched filter output.

2. Sort the users in order of decreasing maximum-magnitudes.

3. **for** $k = 1 : K$,

   - Perform selective filtering for the kth user in the sorted list.
   - Assuming that the kth user has been decoded correctly and the decoded message is $\hat{m}_k$, retain only signature $\mathbf{s}_{k,\hat{m}_k}$ of user k in the selective filter matrix of user $k+1$.

In the above algorithm, we can potentially employ any of the selective detectors proposed in Section 4. We will present performance results for the selective decorrelator with SIS in the next section. Note that the selective decorrelator with SIS and the noncoherent decision feedback detector proposed in [5] share the similarity that for the users whose messages have been decoded, both schemes decorrelate only against the signatures corresponding to the decoded messages. However, they differ in that [5] performs non-selective decorrelation against the M-1 signatures of the desired user and uses a second stage single user GLRT detector instead of the MM rule for symbol decisions.

5 Numerical Results

In this section we evaluate the performance of the proposed constrained detectors, the soft-ICs, and selective detectors and compare with the non-selective detectors proposed in [4, 6]. Since the exact symbol error rate expressions are cumbersome or intractable for each of the detectors considered in the paper, we resort to simulations for performance evaluations.

In all simulations, we used complex random signatures. The signatures are linearly independent and hence the inverse of the cross-correlation matrix $\mathbf{R}$ exists. In all figures, user 1 is assumed to be the desired user and P_s represents the probability of symbol error of user 1. We also assume that the M messages of user k are received with equal power, or that $A_{k,m} = A_k$. The SNR of user k is defined as $A_k^2/2\sigma^2$. In near-far scenarios, all interferers are assumed to be at the same SNR.

Figure 1 shows P_s versus the SNR for $K = 2$ users, $M = 4$ messages per user, and a processing gain of $N = 20$ for the detectors studied in Section 3. The parameters K and M were chosen to be small due to the implementation complexity of the *joint* detector in (7). However, it is worth noting that although the number of users is small, KM itself is a sizeable fraction of the processing gain N. Experiments with larger processing gains are considered later on for the detectors proposed in this paper. It is interesting to note that the *global* (16) and *local* (18) constrained detectors perform very close to the MMSE detector (10). A similar observation has been made in [12] as well, and this may be attributed to the resemblance of the analytical solutions of constrained optimization problems to the generalized MMSE solution. Figure 2 shows the symbol error probability P_s of the desired user versus the SNR of the interferer in a near-far scenario. Since the *local* constrained detector performs only slightly better than the *global* detector, it has been omitted from this figure.

Figure 3 compares the performance of the various soft-ICs proposed in this paper (19)-(24) to the non-selective decorrelative and MMSE detectors in a near-far scenario. In all the soft-ICs, a decorrelative first stage was used followed by two more stages of processing of the matched filter outputs. Interestingly, the non-selective MMSE detector (10), (11) does not converge to the decorrelator in the high interferer power region in contrast with multiuser systems that employ linear modulation. This is a direct consequence of the fact that, in the near-far situation, the powers of the interferers are high compared to the powers associated with all possible messages of the desired user, and that the non-selective detectors take the undesired $M-1$ messages of the desired user (with relatively low powers) as well as all interferers' signals (with high powers) into account in decoding the desired user's message. Thus, unlike the decorrelator, the non-selective MMSE filter does not zero-force the contributions of the $M-1$ undesired messages of the desired user, resulting in a performance improvement in near-far scenarios. Note that this issue does not arise for selective filters, and the selective MMSE and decorrelative detectors do converge in the near-far situations.

In Figure 4, the upper graph compares the performance of the selective and non-selective filters for a lightly loaded system with $K = 2$ users, $M = 4$ signals per user, and $N = 20$ dimensions; the lower graph compares the selective and non-selective implementations of the decorrelator and MMSE for a fully loaded system with $K = 5$, $M = 4$, and $N = 20$. Note that the non-selective decorrelator and MMSE curves compare well with those of Figure 3(a) in [6]. Next, we increase both the processing gain N as well as M, the number of messages per user. The upper and lower graphs of Figure 5 show the relative performance of the detectors for a moderately loaded and a fully loaded system, respectively. It can be seen that the selective detectors consistently outperform the non-selective detectors at all values of SNR. Among the selective detectors, the serial soft-IC is better able to cancel interferers at higher powers and hence the cross-over in the lower graph.

The upper graph in Figure 6 shows the NSE (37) of the blind adaptive selective (36) and non-selective MMSE detectors [6] averaged over 10 runs, for different step sizes μ. The limiting MSE (mean squared error) of the detector is proportional to the value of NSE to which the filter coefficients converge. The step size impacts both the rate of convergence as well as the limiting MSE, and the tradeoff between the two is apparent from the figure. It can be seen that a larger μ brings about faster convergence but at the cost of a higher limiting MSE. The selective detector converges to a lower value of NSE compared to its non-selective counterpart at $\mu = 0.0001$ and vice-versa at $\mu = 0.001$. The performance of the blind selective MMSE detector is illustrated in the lower graph in Figure 6 and compared with that of the blind non-selective MMSE (Figure 7(b) of [6]) as well as the non-blind selective and non-selective MMSE detectors (Figure 3(a) of [6]). For lower values of SNR, $\mu = 0.001$ was used, whereas for the higher values of SNR, $\mu = 0.0001$ had to be used to ensure proper convergence. The ability of the blind filter coefficients to track their non-blind deterministic counterparts is reflected in the similarity between the error probability curves for the blind and the non-blind implementations as shown in the figure.

Figure 7 illustrates the performance improvement for the selective decorrelator gained by using the SIS scheme. Since the non-selective decorrelator is near-far resistant [4], the probability of symbol error for the desired user remains unchanged with interference power. The selective decorrelator also exhibits a similar behavior since it projects the received signal onto a space orthogonal to the interferers' subspace which remains unaffected by a change in the interferers' SNR. With SIS however, the situation is different. When the interferers' SNRs are lower than the desired user's SNR, the desired user is decoded first and it does not benefit from the SIS scheme. Hence its performance is similar to that of the selective decorrelator without SIS in the low SNR regions. In the high SNR regime, the desired user is decoded last with very high probability and hence it benefits the most from the SIS scheme (due to a reduction in the space of possible interfering signals) yielding an improvement in symbol error rate of around two orders of magnitude over the selective decorrelator without SIS. The SIS curve finally flattens out in the high SNR region because the dimensionality of the interference subspace remains unaffected by a change in the interferers' SNRs.

For comparison, the successive interference canceler (SIC) of [27] is also included in Figure 7. Due to the high correlation between the signatures of the users in our case, the MAI residual terms remain significant for the subsequent users even after the high energy users are canceled out. That is why the performance of the desired user deteriorates as the SNR of the interferers is increased. On the other hand, when the SNR of the interferers is low, the desired user is the first to get decoded and hence it does not suffer from the excessive MAI. Note that for some other instance of the cross-correlation matrix, the SIC may yield results that are qualitatively similar to that of the SIS scheme.

6 Conclusion

This paper illustrates how the judicious use of a priori knowledge of the users' selective transmission mechanism can yield improved performance over the noncoherent multiuser detectors proposed in the literature [4, 6]. To this end, we proposed and investigated three categories of detectors.

First, a nonlinear programming approach to noncoherent multiuser detection was explored, where the structure of the multiuser signal reflected in the various constraint sets analyzed. Using this technique, the *joint* detector (7) was derived to provide a benchmark (a lower bound) for evaluating the performance of different detectors. The *global* and *local* detectors were also derived as relaxations of the ML detector in (16) and (18), respectively, but with different constraint sets. These two detectors were shown to resemble the solution to the *generalized* MMSE detector, as also observed in [12] for linear modulation and coherent detection.

Second, motivated by the ability of the soft-IC based coherent detectors to perform well in near-far scenarios, the serial, clipped, and parallel implementations of noncoherent soft-ICs were suggested and investigated. The three detectors mainly differ in the manner they incorporate the *a priori* information available regarding the structure of the signal. It was observed that the serial soft-IC not only outperforms MMSE and decorrelative detectors in near-far scenarios, but that it does so in equal-received-powers situations as well. A selective implementation of the serial soft-IC was also considered.

Third, we proposed and implemented a class of detectors that employ selective filtering. Unlike their non-selective counterparts, these detectors make use of the *a priori* information that of the M signals available to a user, only *one* is transmitted. The decorrelative, MMSE and soft-IC selective detectors have been shown to outperform their non-selective counterparts in all cases. To illustrate the feasibility of the selective detectors in scenarios where limited information regarding the interferers is available, e.g. a CDMA down-link, a blind adaptive implementation of the selective MMSE detector was presented. Finally, an approach to improve the performance of the selective detectors based on decision-directed successive user suppression was presented in this paper.

Our results indicate that incorporating the information regarding the signal structure offers performance improvements. In particular, detectors employing selective filtering with their excellent performance emerge as viable solutions in a variety of system conditions.

A Derivation of the local constrained detector

Here, we derive the solution to the optimization problem in (17). The objective function in (17) can be expanded in terms of $\mathbf{y} = \mathbf{S}^H\mathbf{r}$ as

$$||\mathbf{r} - \mathbf{SAx}||^2 = \mathbf{r}^H\mathbf{r} - 2\text{Re}[\mathbf{y}^H\mathbf{Ax}] + \mathbf{x}^H\mathbf{Tx} \quad (39)$$

where $\mathbf{T} = \mathbf{ARA}$. Since (17) involves the minimization of a convex function over a convex set, it has a unique minimum over this constraint set which can be found using a variety of iterative algorithms, e.g. the gradient descent algorithm [28]. In addition, the convex duality theorem [28] ensures that no duality gap exists and one can solve for the dual problem instead. Since (17) has K constraints, there are K dual variables. In terms of $\mathbf{T}_{kj} = \mathbf{A}_k\mathbf{S}_k^H\mathbf{S}_j\mathbf{A}_j$, the Lagrangian dual function of (39) can be expressed as

$$\begin{aligned}\mathcal{L}(\mathbf{x},\lambda) &= \mathbf{x}^H\mathbf{Tx} - 2\text{Re}[\mathbf{y}^H\mathbf{Ax}] + \sum_{i=1}^{K}\lambda_i(\mathbf{x}_i^H\mathbf{x}_i - 1) \quad (40)\\ &= \sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{x}_i^H\mathbf{T}_{ij}\mathbf{x}_j - 2\text{Re}\left[\sum_{i=1}^{K}\mathbf{y}_i^H\mathbf{A}_i\mathbf{x}_i\right] + \sum_{i=1}^{K}\lambda_i(\mathbf{x}_i^H\mathbf{x}_i - 1) \quad (41)\end{aligned}$$

which is to be maximized over $\mathbf{x}$ and $\lambda \geq 0$, where $\lambda = [\lambda_1 \cdots \lambda_K]^\top$. The gradient vector associated with $\mathcal{L}(\mathbf{x},\lambda)$ is $\nabla\mathcal{L}(\mathbf{x},\lambda) = [\nabla_{\mathbf{x}_1}^\top\mathcal{L}(\mathbf{x},\lambda)\cdots\nabla_{\mathbf{x}_K}^\top\mathcal{L}(\mathbf{x},\lambda)]^\top$, where

$$\nabla_{\mathbf{x}_k}\mathcal{L}(\mathbf{x},\lambda) = 2\left(\sum_{j=1}^{K}\mathbf{T}_{kj}\mathbf{x}_j - \mathbf{A}_k\mathbf{y}_k + \lambda_k\mathbf{x}_k\right). \quad (42)$$

Consequently,

$$\nabla\mathcal{L}(\mathbf{x},\lambda) = 2\left(\mathbf{Tx} - \mathbf{Ay} + \Lambda\mathbf{x}\right) \quad (43)$$

where

$$\Lambda = \text{diag}[\underbrace{\lambda_1,\ldots,\lambda_1}_{M \text{ terms}},\ldots,\underbrace{\lambda_K,\ldots,\lambda_K}_{M \text{ terms}}]$$

is an $MK\times MK$ diagonal matrix. Let $\hat{\mathbf{x}}(\lambda)$ be the solution to $\nabla\mathcal{L}(\hat{\mathbf{x}}(\lambda),\lambda) = 0$. Solving for $\hat{\mathbf{x}}(\lambda)$, we get

$$\hat{\mathbf{x}}(\lambda) = (\mathbf{T} + \Lambda)^{-1}\mathbf{Ay} \quad (44)$$

where $\hat{\mathbf{x}}(\lambda) = [\hat{\mathbf{x}}_1^\top(\lambda_1)\cdots\hat{\mathbf{x}}_K^\top(\lambda_K)]^\top$. On substituting $\hat{\mathbf{x}}(\lambda)$ back into (40), we arrive at

$$\max_{\lambda\geq\mathbf{0}} \quad \mathcal{L}(\hat{\mathbf{x}}(\lambda),\lambda) = \hat{\mathbf{x}}(\lambda)^H\mathbf{T}\hat{\mathbf{x}}(\lambda) - 2\text{Re}[\mathbf{y}^H\mathbf{A}\hat{\mathbf{x}}(\lambda)] + \sum_{i=1}^{K}\lambda_i(\hat{\mathbf{x}}_i(\lambda_i)^H\hat{\mathbf{x}}_i(\lambda_i) - 1). \quad (45)$$

Simple unconstrained gradient descent algorithms can be used to iteratively determine each element of λ as follows

$$\bar{\lambda}_k(t+1) = \bar{\lambda}_k(t) - \mu_k\left[\nabla\mathcal{L}_{\lambda_k}(\hat{\mathbf{x}}(\lambda),\lambda)\right] \quad (46)$$

which converges to $\bar{\lambda}_k$. The maximizer of (45) is given by $\lambda^* = [\lambda_1^*\cdots\lambda_K^*]^T$ where, $\lambda_k^* = \max(0,\bar{\lambda}_k)$. Then, from (44), the unique minimizer of (17) can be re-written as

$$\hat{\mathbf{x}}(\lambda^*) = \mathbf{A}^{-1}(\mathbf{R} + \Lambda^*\mathbf{A}^{-2})^{-1}\mathbf{y}. \quad (47)$$

B Serial Soft-IC

From the matched filter of (8), the element of $\mathbf{y}$ corresponding to message k of user m can be written as

$$y_{k,m} = \sum_{i=1}^{K} \sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{i,j} A_{i,j} x_{i,j} + \mathbf{s}_{k,m}^{H} \mathbf{n}. \tag{48}$$

Decomposing (48) to isolate the terms containing the kth user's messages, we have

$$y_{k,m} = \sum_{i=1}^{k-1} \sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{i,j} A_{i,j} x_{i,j} + \sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,j} A_{k,j} x_{k,j} + \sum_{i=k+1}^{K} \sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{i,j} A_{i,j} x_{i,j} + \mathbf{s}_{k,m}^{H} \mathbf{n}. \tag{49}$$

Note that the second term on the right side of (49) can be decomposed further to isolate the term representing the current message $x_{k,m}$ as

$$\sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,j} A_{k,j} x_{k,j} = \sum_{j=1}^{m-1} \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,j} A_{k,j} x_{k,j} + \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,m} A_{k,m} x_{k,m} + \sum_{j=m+1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,j} A_{k,j} x_{k,j}. \tag{50}$$

Since the signatures have unit norm, $\mathbf{s}_{k,m}^{H} \mathbf{s}_{k,m} = 1$, substituting (50) in (49) and solving for $x_{k,m}$ yields

$$\begin{aligned} x_{k,m} = \frac{y_{k,m}}{A_{k,m}} & - \frac{1}{A_{k,m}} \sum_{i=1}^{k-1} \sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{i,j} A_{i,j} x_{i,j} - \frac{1}{A_{k,m}} \sum_{j=1}^{m-1} \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,j} A_{k,j} x_{k,j} \\ & - \frac{1}{A_{k,m}} \sum_{j=m+1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{k,j} A_{k,j} x_{k,j} - \frac{1}{A_{k,m}} \sum_{i=k+1}^{K} \sum_{j=1}^{M} \mathbf{s}_{k,m}^{H} \mathbf{s}_{i,j} A_{i,j} x_{i,j}. \end{aligned} \tag{51}$$

In the *serial* soft-IC, the current message estimate $\tilde{x}_{k,m}$ is determined using the previous decisions to reconstruct the interference which is then subtracted from the matched filter output $x_{k,m}$. By rewriting the right side of (51) to distinguish those $x_{k,m}$ which have already been estimated, we obtain (19).

References

[1] J. Proakis. *Digital Communications*. McGraw Hill, 3rd edition, 1995.

[2] S. Verdú. *Multiuser Detection*. Cambridge University Press, 1998.

[3] S. Verdú. Computational complexity of multiuser detection. *Algorithmica*, 4(4):303–312, 1989.

[4] M. K. Varanasi and A. Russ. Noncoherent decorrelative detection for nonorthogonal multipulse modulation over the multiuser gaussian channel. *IEEE Transactions on Communications*, 46(12):1675–1684, December 1998.

[5] M. K. Varanasi and D. Das. Noncoherent decision-feedback multiuser detection. *IEEE Transactions on Communications*, 48(2):259–269, February 2000.

[6] A. Kapur, D. Das, and M. K. Varanasi. Noncoherent MMSE multiuser receivers and their blind adaptive implementations. *IEEE Transactions on Communications*, 50(3):503–513, March 2002.

[7] M. L. McCloud and L. L. Scharf. Asymptotic analysis of the MMSE multiuser detector for non-orthogonal multipulse modulation. *IEEE Transactions on Communications*. Submitted September 1999.

[8] E. Visotsky and U. Madhow. Noncoherent multiuser detection for CDMA systems with nonlinear modulation: a non-Bayesian approach. *IEEE Transactions on Information Theory*, 47(4):1352–1367, May 2001.

[9] R. Sinha, A. Yener, and R. D. Yates. Constrained detection for noncoherent nonlinear multiuser communications. In *Proceedings of 34th Asilomar Conference on Signals, Systems and Computers*, October 2000.

[10] M. K. Varanasi and B. Aazhang. Multistage detection in asynchronous CDMA communications. *IEEE Transactions on Communications*, 38(4):509–519, April 1990.

[11] A. Yener, R. D. Yates, and S. Ulukus. A nonlinear programming approach to CDMA multiuser detection. In *Proceedings of 33rd Asilomar Conference on Signals, Systems and Computers*, October 1999.

[12] A. Yener, R. D. Yates, and S. Ulukus. CDMA multiuser detection: A nonlinear programming approach. *IEEE Transactions on Communications*, 50(6), June 2002.

[13] P. H. Tan, L. K. Larsmussen, and T. J. Lim. Constrained maximum-likelihood detection in CDMA. *IEEE Transactions on Communications*, 49(1):142–153, January 2001.

[14] W. Ma, T. N. Davidson, K. M. Wong, Z. Q. Luo, and P. C. Ching. Efficient quasi-maximum-likelihood multiuser detection by semi-definite relaxation. In *IEEE International Conference on Communications*, pages 6–10, June 2001.

[15] P. Spasojevic and A. Yener. Constrained slowest descent detectors for multiuser CDMA systems. In *IEEE International Symposium on Information Theory, ISIT'01*, June 2001.

[16] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[17] U. Madhow and M. L. Honig. MMSE interference suppression for direct-sequence spread-spectrum CDMA. *IEEE Transactions on Communications*, 42(12):3178–3188, December 1994.

[18] P. Hansen. Methods of nonlinear 0-1 programming. *Annals of Discrete Mathematics 5: Discrete Optimization II*. P.L. Hammer, E.L. Johnson and B.H. Korte, Eds. Noth Holland, 1979.

[19] M. K. Varanasi and B. Aazhang. Near-optimum detection in synchronous CDMA systems. *IEEE Transactions on Communications*, 39(5):725–736, May 1991.

[20] D. Divsalar, M. K. Simon, and D. Rapheli. Improved parallel interference cancellation for CDMA. *IEEE Transactions on Communications*, 46(2):258–268, February 1998.

[21] X. Zhang and D. Brady. Asymptotic multiuser efficiencies for decision-directed multiuser detectors. *IEEE Transactions on Information Theory*, 44(2):502–515, March 1998.

[22] L. B. Nelson and H. V. Poor. Iterative multiuser receivers for CDMA channels: An EM-based approach. *IEEE Transactions on Communications*, 44(12):1700–1710, December 1996.

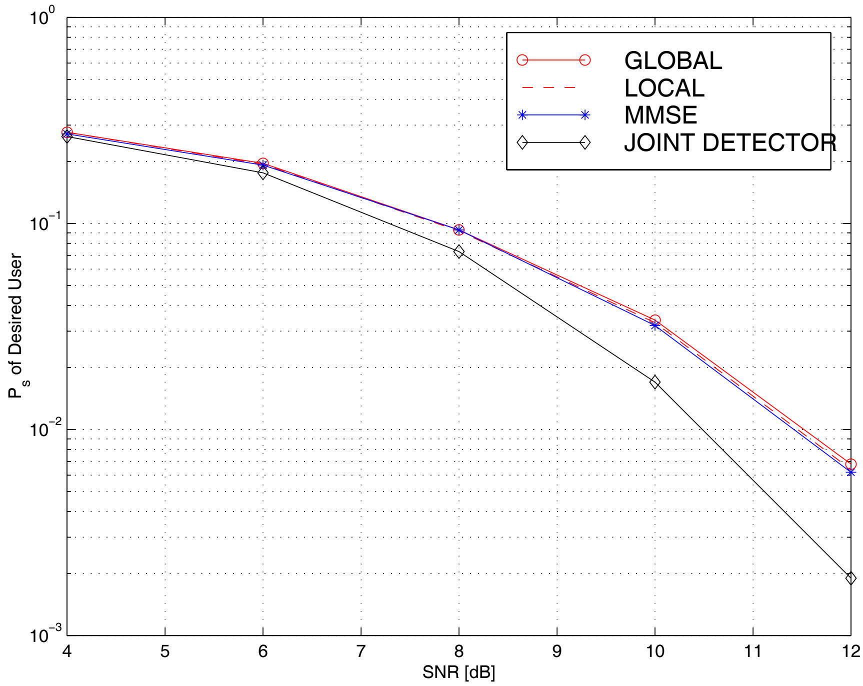


Figure 1: Comparison of various noncoherent detectors with (K, M, N)=$(2, 4, 20)$.

[23] L. G. F. Trichard, J. S. Evans, and I. B. Collings. Large system analysis of parallel interference cancellation. In *IEEE International Conference on Communications*, pages 6–10, June 2001.

[24] M. L. Honig, U. Madhow, and S. Verdu. Blind adaptive multiuser detection. *IEEE Transactions on Information Theory*, 41:994–960, July 1995.

[25] X. Wang and V. Poor. Blind multiuser detection: A subspace approach. *IEEE Transactions on Information Theory*, 44(2):677–689, March 1998.

[26] S. S. Haykin. *Adaptive Filter Theory*. Prentice Hall, 3rd edition, 1995.

[27] P. Patel and J. Holtzman. Analysis of a simple successive interference cancellation scheme in a DS/CDMA system. *IEEE Journal on Selected Areas in Communications*, 12(5):796–807, June 1994.

[28] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. Preprint, 1995.

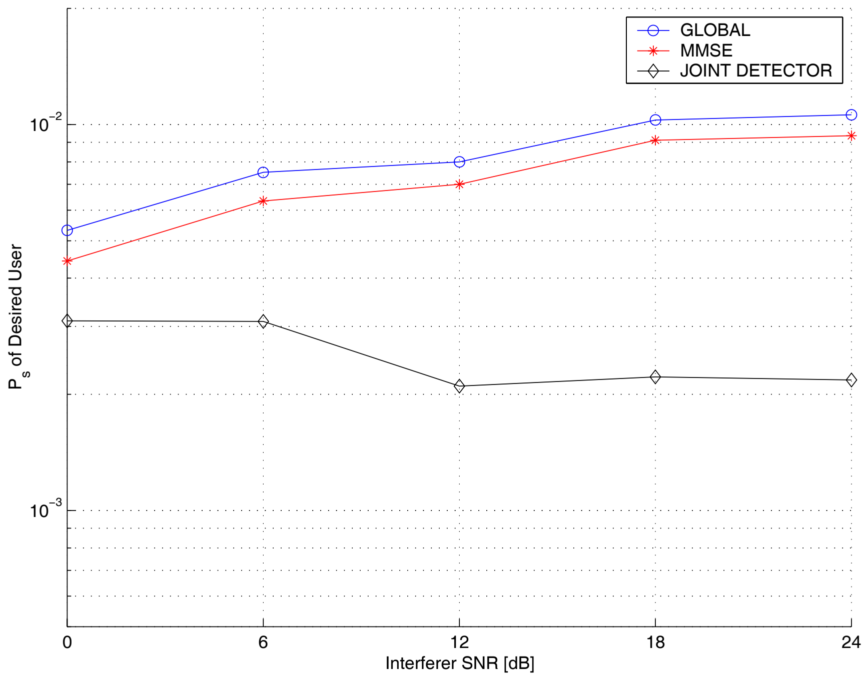


Figure 2: Performance in the near-far scenario with (K,M,N)=(2, 4, 20). Desired user's SNR= 12dB.

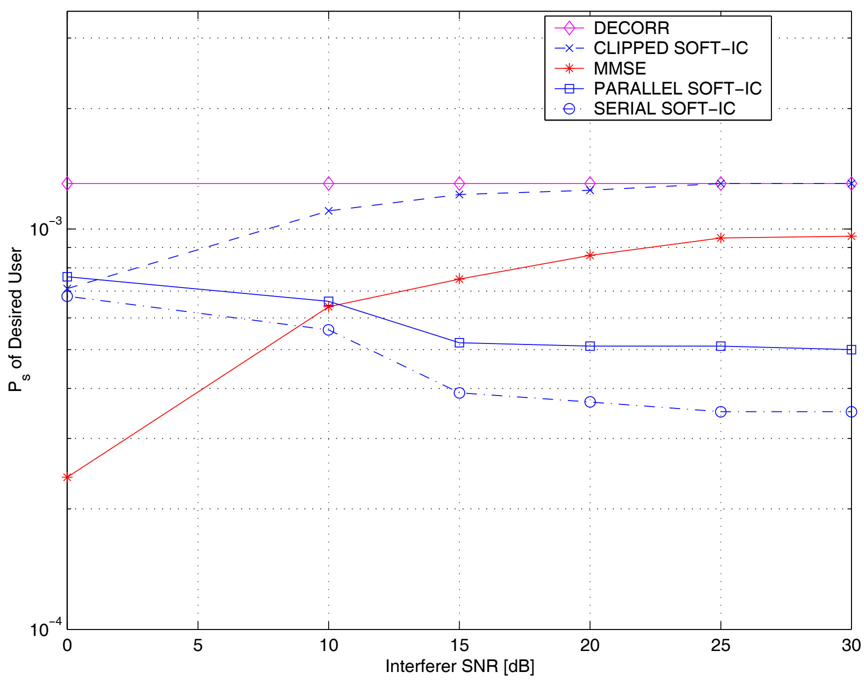


Figure 3: Comparison of soft-ICs in near-far scenario with (K,M,N)=(2, 4, 20). Desired user's SNR= 14dB.

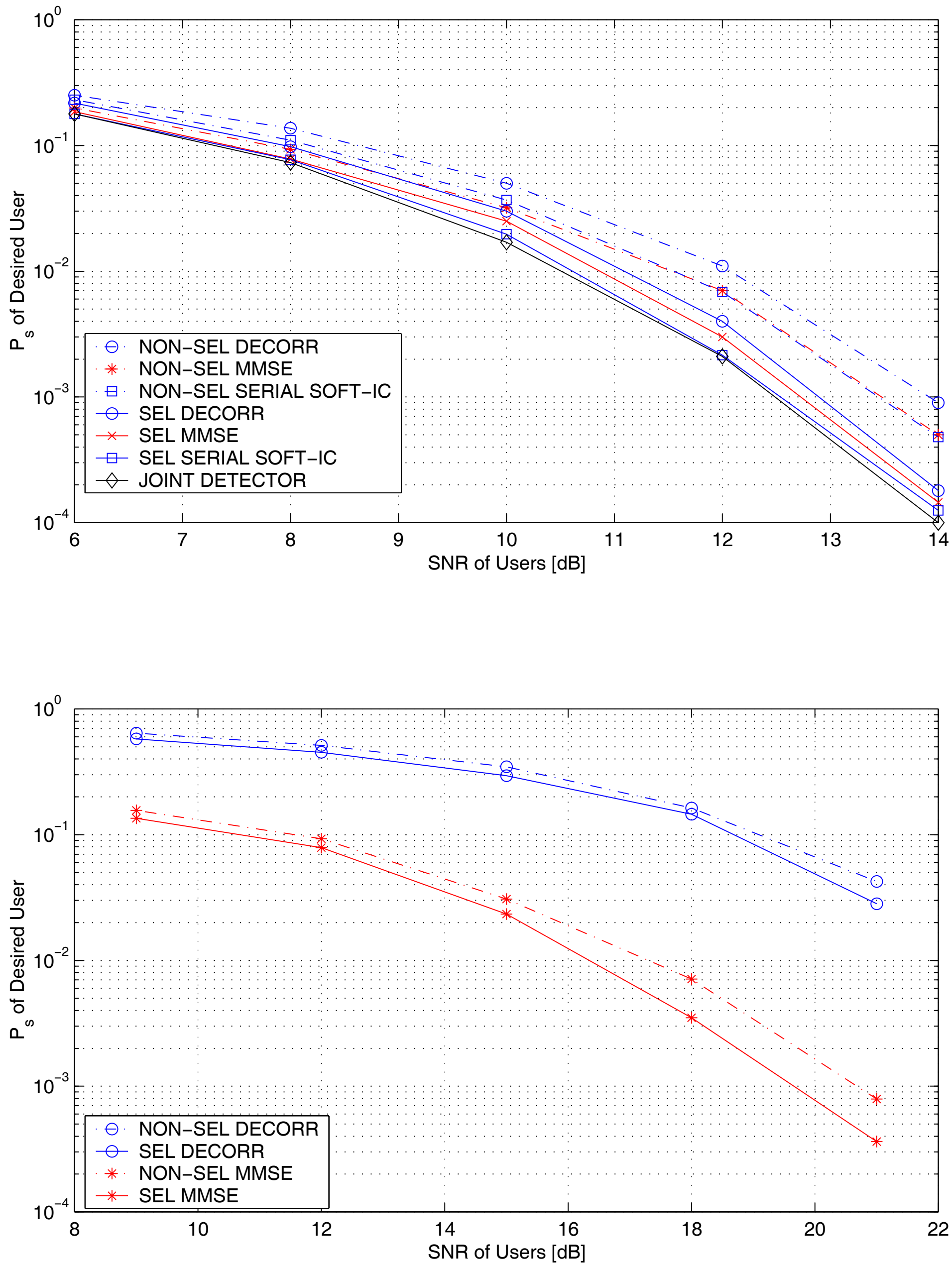
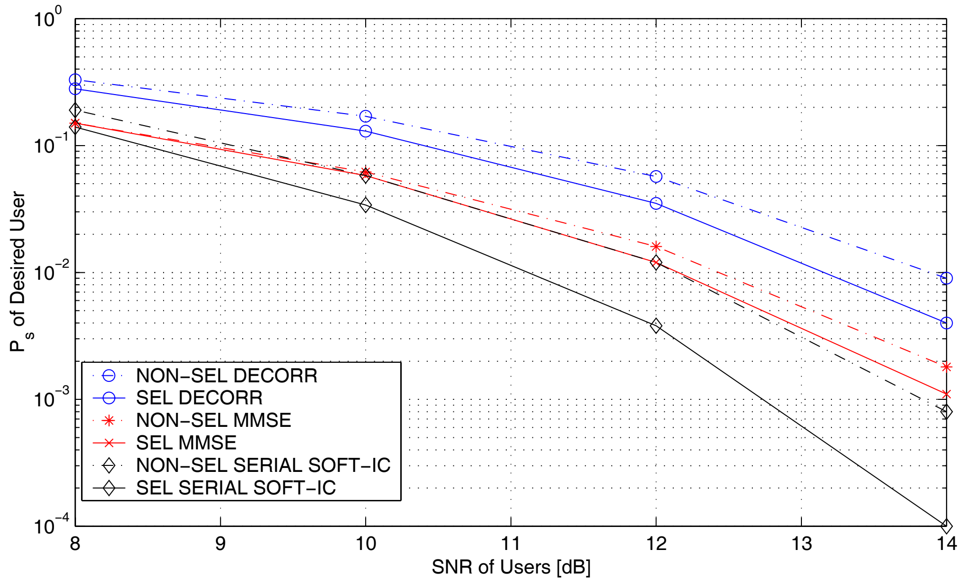


Figure 4: Comparison of selective and non-selective detectors. The upper graph shows the performance of a lightly loaded system with $(K,M,N) = (2,4,20)$, whereas the lower graph is for a fully loaded system with $(K,M,N) = (5,4,20)$.



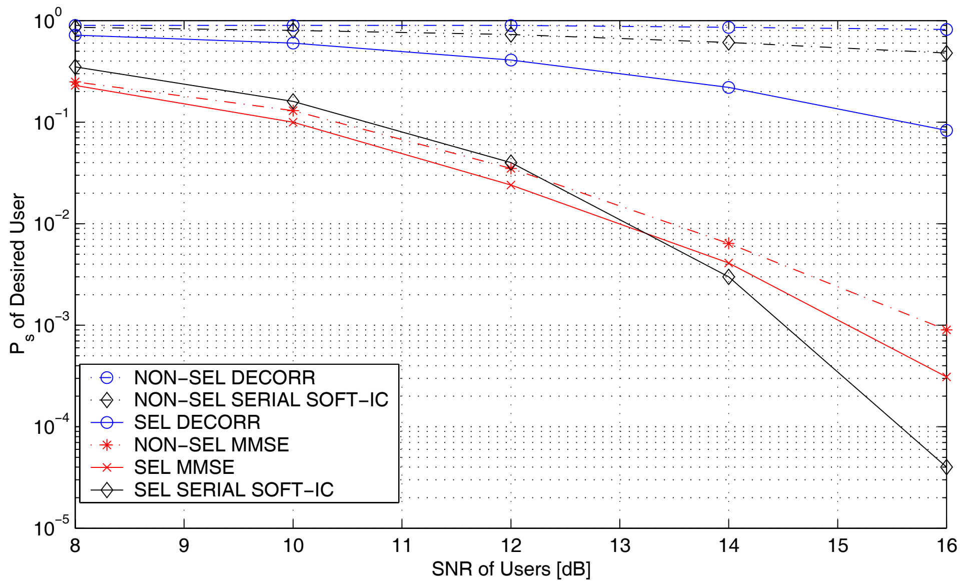


Figure 5: Comparison of selective and non-selective detectors with higher processing gain $N = 64$. The top figure represents a moderately loaded system with (K, M, N)=$(4, 8, 64)$ whereas the bottom figure represents a fully loaded system with (K, M, N)=$(4, 16, 64)$.

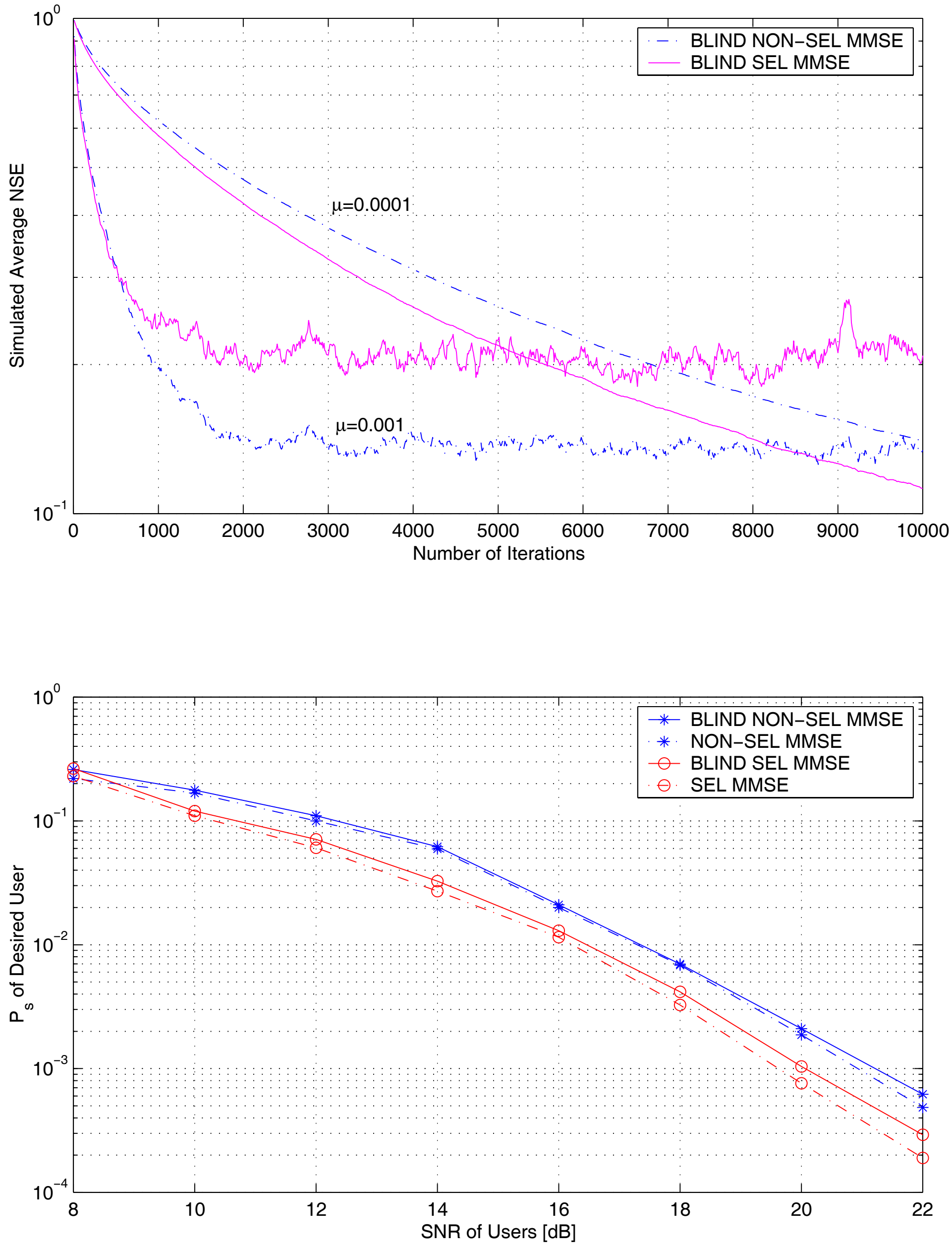


Figure 6: The top figure compares the convergence of the blind adaptive selective and non-selective MMSE filters with (K, M, N)=$(5, 4, 20)$ and variable step-size μ, whereas the bottom figure compares their performance.

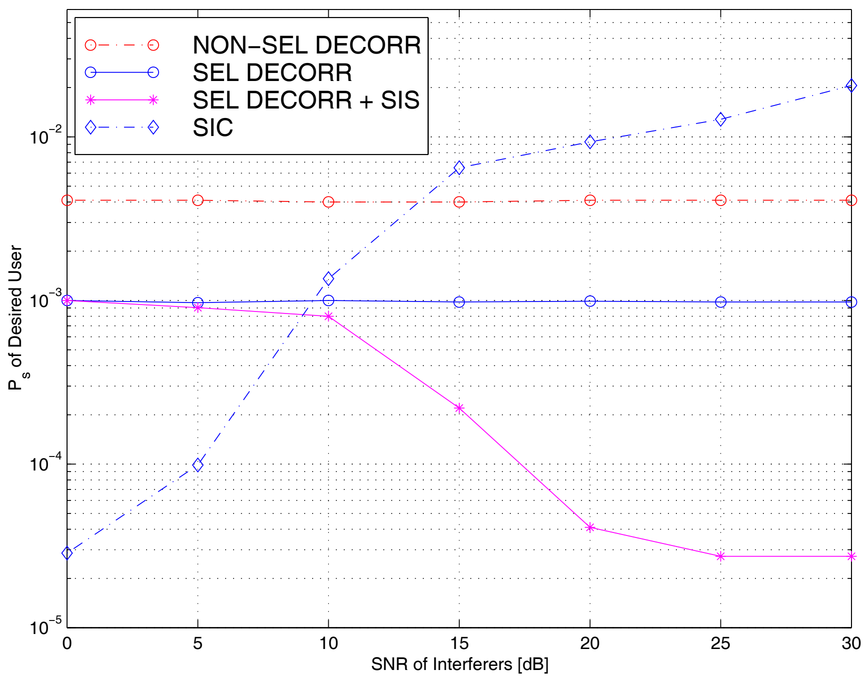


Figure 7: Performance of selective decorrelator with SIS with (K, M, N)=$(3, 4, 20)$ and desired user's SNR at 15dB.