Submitted to IEEE Trans. Comm. 10/98

On Multipath Channel Estimation for CDMA Systems Using Multiple Sensors

Chaitali Sengupta Joseph R. Cavallaro Behnaam Aazhang

Abstract

This paper focuses on the design of a multiuser receiver structure for a Code Division Multiple Access (CDMA) communication system, in the presence of multipath effects and using an antenna array at the base station receiver. The algorithm presented in this paper solves the complex multidimensional problem of channel estimation in this complex scenario using a maximum likelihood approach. Once a composite channel impulse response of each user is estimated, it is directly used in the detection process instead of first extracting the individual channel parameters, such as, path delays and attenuation factors. The paper presents a framework that facilitates a computationally efficient solution to the combined problem of channel estimation and detection in a scenario involving multiple users, multiple paths, and multiple sensors at the receiver.

Keywords: CDMA, channel estimation, multipath, multiple sensors

---

The authors are with the Department of Electrical and Computer Engineering (MS 366), Rice University, Houston, Texas 77005. Email: {chaitali,cavallar,aaz}@rice.edu
This work was supported by Nokia Inc., by the Texas Advanced Technology Program under grants 1995-#003604-049 and 1997-#003604-044, and by NSF under grants NCR 9506681 and CDA-9617383.
This paper was presented in part at the 1998 Conference on Information Sciences and Systems, Princeton, USA, and at the 9th IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications, Boston, 1998.

1 Introduction

In a *Code Division Multiple Access (CDMA)* communication system, a communication channel with a given bandwidth is accessed by all the users simultaneously. The different mobile users are distinguished at the base station receiver by the unique *spreading code* assigned to the users to modulate their signals. Hence, the CDMA signal transmitted by any given user consists of that user's data which modulates the unique spreading code assigned to that user, which in turn modulates a carrier using any well-known modulation scheme such as *binary phase shift keying (BPSK)*. The frequency of this carrier is the same for all users. The receiver receives a linear superposition of the signals transmitted by all the users, attenuated by arbitrary factors and delayed by an arbitrary amount. The goal of channel parameter estimation is to determine these unknown and time varying attenuation factors and delays by processing the received signal, to facilitate recovery of the data transmitted by each user.

Channel estimation is one of the major problems in radio communications particularly when the mobile system is subject to multipath fading, that is, the transmission channel consists of more than one distinct propagation path for each user's signal. Moreover, when the CDMA technique is used to allow multiple users access to a single channel, the system is susceptible to the *near-far* effect. The near-far problem arises when the signals from the different users arrive at the receiver with widely varying power levels. The near-far problem has been shown to severely degrade the performance of standard single user techniques (e.g., matched filters, correlators, etc.) in conventional CDMA systems.

Conventional CDMA systems try to limit the near-far problem with power control. However, even a small amount of the near-far effect can drastically degrade the performance of conventional receivers. For many years this was thought to be an inherent limitation of CDMA until Verdú developed the optimum multiuser detector [1]. Verdú's work was followed by many suboptimal schemes of lower computational complexity [2, 3, 4], all of which are near-far resistant. However, these methods deal only with detection and assume that the timing of the spreading waveforms is known.

Most of the initial work done on timing acquisition for CDMA systems focused on jointly estimating the necessary parameters for all users [5, 6]. While these techniques produce excellent results, they can be computationally intense since they involve solving a multidimensional optimization problem for a large number of parameters.

Subspace-based techniques [7, 8, 9, 10] provide a method for decomposing the multidimensional parameter search into a series of one dimensional optimization problems. Such algorithms are near-far resistant and effective in the presence of multiple propagation paths. In addition they do not require any preamble or training sequences. The subspace-based technique has also been extended to joint channel estimation and detection [11, 12]. However, the primary drawback of such algorithms is the computational requirement of the signal or noise subspace estimation process, and the limitation on the number of users that can be handled.

Recently, a number of channel estimation algorithms based on the maximum likelihood technique (ML) [13] have emerged [14, 15, 16, 17, 18, 19]. The ML technique is an attractive alternative as the ML based algorithms have been shown to be computationally efficient and can handle a large number of users. It has been also shown [17, 18, 20], in the single path case, that the application of the ML technique in the presence of a number of sensors at the receiver dramatically improves the performance of channel estimators as well as linear detectors.

Several forms of the ML technique have been proposed in literature. These algorithms work in different situations, such as single path and single sensor [14, 16] or single path and multiple sensors [15, 17, 18] or single sensor and multiple paths [19]. Out of these, [16, 17, 18] fall under a distinct class which can be easily separated from our proposed algorithm. All these three algorithms model all the interfering users as colored noise, thus reducing the size of the spreading code matrix and training sequence vector and hence the required computation. Also the algorithm in [17] uses a new structure for the observation vectors while forming the joint conditional density functions. This structure exploits the spatial diversity provided by the multiple sensors in such a way that a very short training sequence (about 20 bits) can be used to achieve acquisition. However it should be noted that none of these algorithms handle the problem of multipath channel parameters, using multiple sensors.

The problem of using multiple sensors to detect multipath channel parameters of multiple users [21], is mathematically complex, as it involves the estimation of a large number of parameters. In this paper, we present a maximum likelihood (ML) approach that solves the complex problem of channel estimation in the presence of multiple propagation paths and multiple sensors, in a computationally efficient manner.

We define our channel and system model such that the effect of both multiple paths and multiple sensors can be incorporated in an efficient manner. The model is then used to develop an ML based algorithm that estimates a composite channel impulse response for each user. This composite channel impulse response is directly used in the detection process to appropriately modify the spreading code of the user. Thus, we define a process by which both the channel estimation and detection process can benefit from exploiting multiple paths and multiple sensors at the base station receiver.

The proposed algorithm assumes the transmission of training sequences by all the users being acquired but makes no assumption on the individual delays or on the antenna array structure. The delays are estimated modulo N, where N is the length of each spreading code. The additive noise is assumed to be a circularly complex zero mean Gaussian random vector, but no *a priori* assumption is made on its covariance.

The paper is organized into the following sections. Section 2 presents the CDMA system under study and a model for the multipath channel. In Section 3 we describe the maximum likelihood algorithm and Section 4 describes the use of the estimated channel impulse response directly in the detection process. Section 5 determines the complexity of the channel estimation algorithm and Section 6 presents the results of our simulations. We conclude the paper in Section 7 with a summary and directions for future research.

2 CDMA System Model : Multipath and Multiple Sensors

In this section, we develop a system model that incorporates the effect of both multiple paths and multiple sensors in an efficient manner. The model does not assume chip synchronicity, that is the delays of the signals, introduced by the channel are not integral multiples of chips (Figure 1).

We assume a K-user direct sequence CDMA system with BPSK (Binary Phase Shift Keying) modulation with each transmitted signal selected from a binary alphabet and limited to $[0, T]$, where T is the symbol period. Each user transmits a zero mean stationary bit sequence with i.i.d. components.

The complex baseband representation of the k^{th} user's transmitted signal is given by

$$s_k(t) = \sqrt{P_k} \sum_{i=1}^{L} b_{k,i} c_k(t - iT) \ , \tag{1}$$

where P_k is the transmitted power, $b_{k,i} \in \{+1, \ -1\}$ is the i^{th} transmitted bit and $c_k(t)$ is the spreading waveform. The parameter L denotes the number of bits being considered. The spreading or code waveform is composed of N chips and if we assume BPSK for the spreading modulation we have $c_k(t) = \sum_{n=0}^{N-1} c_{k,n} \Pi(t - nT_c)$, where $c_{k,n} \in \{+1, \ -1\}$ and the chip pulse waveform $\Pi(t)$ is a rectangular pulse of duration T_c. We will assume that the extent of the spreading code is one bit period and hence we have $T = NT_c$.

The frontend of the receiver consists of an antenna array of M sensors arranged in a specific geometry. The corresponding array response vector, which is the response to a propagating plane wave impinging on the array at an angle θ, with the vertical is :

$$[\rho^{(1)}(\theta), \cdots, \rho^{(m)}(\theta), \cdots, \rho^{(M)}(\theta)]. \tag{2}$$

and is determined by the geometry of the array. In this paper, no *a priori* assumption is made on the geometry of the array. However it is assumed that the time taken by the signal to traverse the physical array is much smaller than the inverse of the message bandwidth and hence, the envelope characteristics of the signal do not vary across the array (Figure 2).

We assume that the channel for each user consists of P distinct and resolvable propagation paths [7]. The impulse response $h_k(t)$ of the channel seen by user k is given by:

$$h_k(t) = \sum_{p=1}^{P} w_{k,p}\delta(t - \tau_{k,p}) \tag{3}$$

where, $w_{k,p}$ is the complex amplitude with which the p^{th} path of the k^{th} user is received and includes contributions from the channel attenuation and the phase offset and $\tau_{k,p}$ is the relative delay with respect to a reference at the receiver. The channel parameters are assumed to be unknown but constant in the time taken to estimate them. The multipath spread of the channel T_m is the maximum difference between propagation delays, that is, $T_m = \max_{p,p'} |\tau_{k,p} - \tau_{k,p'}|$. In our model, we assume that the multipath spread is less than half the symbol period ($T_m < \frac{1}{2}T$).

Accordingly, the received signal at the base station is a superposition of multiple copies of attenuated and delayed signals transmitted by all the K users. Therefore, the signal at the m^{th} sensor is given by:

$$r^{(m)}(t) = \sum_{k=1}^{K}\sum_{p=1}^{P} w_{k,p}s_k(t - \tau_{k,p})\rho^{(m)}(\theta_{k,p}) + \nu^{(m)}(t) \tag{4}$$

where $\theta_{k,p}$ is the direction of arrival of the p^{th} path of user k with respect to an axis in the plane of the array.

The additive noise is assumed to have some correlation across the array. We use the following model for the noise [15]: $\nu^{(m)}(t) = \beta\eta(t) + (1-\beta)\eta^{(m)}(t)$. The component $\eta(t)$ is completely correlated across the array and can arise from other-cell interference - this component is taken to be the same at each sensor; $\eta^{(m)}(t)$ is the noise element independent from sensor to sensor, arising from a combination of receiver noise and the uncorrelated contributions from other-cell transmissions ($\beta \in [0, 1)$). The noise components $\eta(t)$ and $\eta^{(m)}(t)$ are uncorrelated random processes and each is assumed to be white and Gaussian with zero-mean and double-sided spectral density of $\mathcal{N}_0/2$. Modeling the other-cell interference as Gaussian is valid in light of the number of such interferers and the applicability of the central limit theorem.

The continuous time signal at each sensor is then discretized [7, 20] by sampling the output of a chip-matched filter, at the chip rate. The chip-matched filtering is a simple integrate and dump operation, over a time interval equal to the chip period:

$$r^{(m)}[n] = \frac{1}{T_c} \int\limits_{nT_c}^{(n+1)T_c} r^{(m)}(t)dt \tag{5}$$

Observation vectors $\mathbf{r}_i^{(m)}$ are formed by collecting together N successive outputs $r^{(m)}[n]$. Hence, each observation vector corresponds to a time interval equal to the bit period. Since each spreading vector is periodic with period N, $r^{(m)}[n]$ is wide sense cyclostationary, and the observation vectors $\mathbf{r}_i^{(m)}$ are wide sense stationary. The observation vector, $\mathbf{r}_i^{(m)} \in \mathbb{C}^N$, at time i is

$$\mathbf{r}_i^{(m)} = \left[r^{(m)}[iN]\,,\ r^{(m)}[iN+1]\,,\ \ldots\,,\ r^{(m)}[iN+N-1]\right]^\top. \tag{6}$$

The system is asynchronous and the receiver has an arbitrary timing reference which will not be aligned to actual transmitted bit boundaries. Hence, each observation vector can be viewed as a linear combination of $2K$ *signal vectors* - 2 components from each user due to the past and current bits as shown in Figure 1. In Figure 1, the received vector $\mathbf{r}_i^{(m)}$ overlaps the two adjacent transmitted bits b_{i-1} and b_i. So the contribution of the spreading code of the user to vector $\mathbf{r}_i$ appears in two parts - the right part (shown as R) of bit b_{i-1} and the left part (shown as L) of bit b_i. Expressing $\mathbf{r}_i^{(m)}$ in terms of its various components:

$$\mathbf{r}_i^{(m)} = \mathbf{A}^{(m)}\mathbf{b}_i + \boldsymbol{\nu}_i^{(m)}, \quad \boldsymbol{\nu}_i^{(m)} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^{(m)}). \tag{7}$$

The $2K$ length vector $\mathbf{b}_i$ is of the form : $\mathbf{b}_i = [b_{1,i-1}, b_{1,i}, \ldots, b_{k,i-1}, b_{k,i}, \ldots, b_{K,i-1}, b_{K,i}]'$, where $b_{k,i}$ is the i^{th} bit of the k^{th} user.

Let us consider the columns of the matrix $\mathbf{A}^{(m)}$. As explained in Figure 1, the contribution of the spreading code of the user to vector $\mathbf{r}_i$ appears in two parts - R and L. Hence, matrix $\mathbf{A}^{(m)}$ has columns corresponding to these two parts: denoted by the superscripts R and L:

$$\mathbf{A}^{(m)} = [\mathbf{a}_1^R \quad \mathbf{a}_1^L \quad \cdots \quad \mathbf{a}_K^R \quad \mathbf{a}_K^L]. \tag{8}$$

Each of the columns $\mathbf{a}_k^R$ and $\mathbf{a}_k^L$ for each user k, are functions of the corresponding delays ($\tau_{k,1}$ to $\tau_{k,P}$), attenuation factors ($w_{k,1}$ to $w_{k,P}$), and array responses ($\rho_{k,1}^{(m)}$ to $\rho_{k,P}^{(m)}$). Here, $\rho_{k,p}^{(m)} = \rho^{(m)}(\theta_{k,p})$.

If T_c is the chip period, let $\tau_{k,p}/T_c = q + \gamma$, $q \in \{0, 1, \cdots, N-1\}, \gamma \in [0, 1)$, we have [7]:

$$\begin{aligned} \mathbf{a}_k^R &= \sum_{p=1}^{P} [w_{k,p}\{(1-\gamma)\mathbf{c}_k^R[q] + \gamma\mathbf{c}_k^R[q+1]\}\rho_{k,p}^{(m)}] \\ \mathbf{a}_k^L &= \sum_{p=1}^{P} [w_{k,p}\{(1-\gamma)\mathbf{c}_k^L[q] + \gamma\mathbf{c}_k^L[q+1]\}\rho_{k,p}^{(m)}] \end{aligned} \tag{9}$$

where $\mathbf{c}_K^R[q]$ and $\mathbf{c}_K^L[q]$ are the spreading codes shifted by integer (multiples of chips) delays.

$$\mathbf{c}_k^R[q] = [c_{k,N-q} \cdots c_{k,N-1}\, 0 \cdots 0]'$$
$$\mathbf{c}_k^L[q] = [0 \cdots 0\, c_{k,0} \cdots c_{k,N-q-1}]' \tag{10}$$

Note that (8) and (9) show that the columns of $\mathbf{A}^{(m)}$ are linear combinations of $\mathbf{c}_k[q]$'s. Let $\mathbf{U}_k^R \in \mathbb{C}^{N\times N}$ and $\mathbf{U}_k^L \in \mathbb{C}^{N\times N}$ be matrices formed from the spreading codes, delayed by all possible integer delays. Let $\mathbf{z}_k^{(m)} \in \mathbb{C}^{N\times 1}$ be the composite channel impulse response vector corresponding to the m^{th} sensor, consisting of the weight vectors in the linear combinations of the $\mathbf{c}_K[q]$'s, expressed in (9). Hence, we can write $\mathbf{U}_k^R$, $\mathbf{U}_k^L$, and $\mathbf{z}_k^{(m)}$ as

$$\mathbf{U}_k^{(R)} = [\mathbf{c}_k^{(R)}[0] \cdots \mathbf{c}_k^{(R)}[N-1]],$$

$$\mathbf{U}_k^{(L)} = [\mathbf{c}_k^{(L)}[0] \cdots \mathbf{c}_k^{(L)}[N-1]], \tag{11}$$

$$\mathbf{z}_k^{(m)} = \mathbf{p}_k^{(m)} \circ \mathbf{h}_k, \tag{12}$$

where,

$$\mathbf{h}_k = \begin{bmatrix} 0 \\ \vdots \\ w_{k,1}(1-\gamma_{k,1}) \\ w_{k,1}\gamma_{k,1} \\ \vdots \\ w_{k,P}(1-\gamma_{k,P}) \\ w_{k,P}\gamma_{k,P} \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow \quad q_{k,1}^{th} element \\ \leftarrow \quad (q_{k,1}+1)^{th} element \\ \\ \leftarrow \quad q_{k,P}^{th} element \\ \leftarrow \quad (q_{k,P}+1)^{th} element \\ \\ \\ \end{array} \tag{13}$$

and where,

$$\mathbf{p}_k^{(m)} = \begin{bmatrix} 0 \\ \vdots \\ \rho_{k,1}^{(m)} \\ \rho_{k,1}^{(m)} \\ \vdots \\ \rho_{k,P}^{(m)} \\ \rho_{k,P}^{(m)} \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow \quad q_{k,1}^{th} element \\ \leftarrow \quad (q_{k,1}+1)^{th} element \\ \\ \leftarrow \quad q_{k,P}^{th} element \\ \leftarrow \quad (q_{k,P}+1)^{th} element \\ \\ \\ \end{array} \tag{14}$$

Here the operator '$\circ$' denotes the element-wise multiplication of the two vector operands. Now, the columns of $\mathbf{A}^{(m)}$ can be expressed in terms of $\mathbf{U}_k^R$, $\mathbf{U}_k^L$, and $\mathbf{z}_k^{(m)}$ as

$$\mathbf{a}_k^R = \mathbf{U}_k^R \mathbf{z}_k^{(m)}, \quad \mathbf{a}_k^L = \mathbf{U}_k^L \mathbf{z}_k^{(m)}. \tag{15}$$

We can now rewrite matrix $\mathbf{A}^{(m)}$ as

$$\mathbf{A}^{(m)} = [\mathbf{U}_1^R \mathbf{z}_1^{(m)} \;\; \mathbf{U}_1^L \mathbf{z}_1^{(m)} \cdots \mathbf{U}_K^R \mathbf{z}_K^{(m)} \;\; \mathbf{U}_K^L \mathbf{z}_K^{(m)}]. \tag{16}$$

Now (7) is expressed as $\mathbf{A}^{(m)} = [\mathbf{U}_1^R \mathbf{z}_1^{(m)} \;\; \mathbf{U}_1^L \mathbf{z}_1^{(m)} \cdots \mathbf{U}_K^R \mathbf{z}_K^{(m)} \;\; \mathbf{U}_K^L \mathbf{z}_K^{(m)}]\mathbf{b}_i + \boldsymbol{\nu}_i^{(m)}$. Hence, the observation vector $\mathbf{r}_i$, of length MN, across the array, $\mathbf{r}_i = [\mathbf{r}_i^{(1)'}, \cdots \mathbf{r}_i^{(k)'}, \cdots \mathbf{r}_i^{(K)'}]'$ is

$$\mathbf{r}_i = \begin{bmatrix} \mathbf{U}_1^R \mathbf{z}_1^{(1)} & \mathbf{U}_1^L \mathbf{z}_1^{(1)} \cdots \mathbf{U}_K^R \mathbf{z}_K^{(1)} & \mathbf{U}_K^L \mathbf{z}_K^{(1)} \\ & \vdots & \\ \mathbf{U}_1^R \mathbf{z}_1^{(M)} & \mathbf{U}_1^L \mathbf{z}_1^{(M)} \cdots \mathbf{U}_K^L \mathbf{z}_K^{(M)} & \mathbf{U}_K^L \mathbf{z}_K^{(M)} \end{bmatrix} \mathbf{b}_i + \boldsymbol{\nu}_i. \tag{17}$$

where the noise vector $\boldsymbol{\nu}_i$ is formed from the components $\boldsymbol{\nu}_i^{(m)}$ and is assumed to have an unknown covariance of $\mathbf{K}$.

Let $\mathcal{U}_k^R \in \mathbb{C}^{MN \times MN}$ and $\mathcal{U}_k^L \in \mathbb{C}^{MN \times MN}$ be matrices defined as follows

$$\mathcal{U}_k^R = I_M \otimes \mathbf{U}_k^R, \qquad \mathcal{U}_k^L = I_M \otimes \mathbf{U}_k^L. \tag{18}$$

Here the operator '$\otimes$' represents the Kronecker product of two matrices. Also let $\mathbf{z}_k \in \mathbb{C}^{MN \times 1}$ be a vector defined as

$$\mathbf{z}_k = [\mathbf{z}_k^{(1)'}, \cdots, \mathbf{z}_k^{(m)'}, \cdots, \mathbf{z}_k^{(M)'}]'. \tag{19}$$

Using these matrices defined above, the columns of the combined code and channel matrix, corresponding to user k, in the expression for $\mathbf{r}_i$ (17) can be written as $\mathcal{U}_k^R \mathbf{z}_k$ and $\mathcal{U}_k^L \mathbf{z}_k$. This helps us to rewrite the expression for $\mathbf{r}_i$ as

$$\mathbf{r}_i = [\mathcal{U}_1^R \mathbf{z}_1 \quad \mathcal{U}_1^L \mathbf{z}_1 \cdots \mathcal{U}_K^R \mathbf{z}_K \quad \mathcal{U}_K^L \mathbf{z}_K]\mathbf{b}_i + \boldsymbol{\nu}_i$$

$$\mathbf{r}_i = \mathcal{U}\mathbf{Z}\mathbf{b}_i + \boldsymbol{\nu}_i \tag{20}$$

where, $\mathcal{U} = [\mathcal{U}_1^R \;\; \mathcal{U}_1^L \cdots \mathcal{U}_k^R \;\; \mathcal{U}_k^L \cdots \mathcal{U}_K^R \;\; \mathcal{U}_K^L]$ and $\mathbf{Z} = diag(\mathbf{z}_1, \mathbf{z}_1, \cdots, \mathbf{z}_k, \mathbf{z}_k, \cdots, \mathbf{z}_K, \mathbf{z}_K)$.

The advantage to be gained from expressing $\mathbf{r}_i$ as (20) is that it allows easy modeling of multiple propagation paths without increasing the size of any of the matrices involved. Increasing the sizes of the matrices involved is directly related to an increase in the computational load of the algorithm. An alternative method to model multiple paths within (7), is to treat the multiple paths as different 'virtual users' and have two additional columns in matrix $\mathbf{A}$ for each extra path of each user. Hence, as the number of paths increases the size of matrix $\mathbf{A}$ also increases. However in (20), the size of matrices $\mathcal{U}$ and $\mathbf{Z}$ does not increase, as P increases; instead matrix $\mathbf{Z}$ becomes more dense. In effect, we discretize the multipath delay axis of the impulse response [22] into N equal time delay segments, each corresponding to a chip interval. Any number of multipath signals received within the n^{th} bin are represented by a single resolvable multipath component.

Now $\mathbf{Z}$ has all the unknown parameters of all the paths of all the users from all the sensors. In the next section we will describe an algorithm in which $\mathbf{Z}$ will be estimated from the observations $\mathbf{r}_i$'s, the known sequence of transmitted bits (e.g., a preamble at the beginning of each frame), and the knowledge of the spreading codes of the different users.

3 Channel Estimation Algorithm

In this section, we will develop a ML based technique for the estimation of the channel impulse response vector $\mathbf{z}_k$ for each user. The algorithm applies the ML technique for angle-of-arrival estimation presented in [23] and for single path, multiple sensor CDMA channel estimation presented in [15]. We apply the same technique to the estimation of a composite channel impulse response vector for each user in the presence of multiple propagation paths using multiple sensors. We solve the complex problem involving both multiple paths as well as multiple sensors in a computationally efficient manner, without increasing the size of the matrices involved as the number of propagation paths increases.

When an observation vector $\mathbf{r}_i$ depends on a parameter vector $\boldsymbol{\xi}$, that is either deterministic but unknown or whose a priori statistics are unknown, the maximum likelihood estimate of the parameter $\boldsymbol{\xi}$ is often used. The maximum likelihood estimate of the parameter is given by [13]

$$\hat{\boldsymbol{\xi}}_{ML} = \arg\max_{\boldsymbol{\xi}} p(\mathbf{r}_i|\boldsymbol{\xi}). \tag{21}$$

In our problem, $\mathbf{r}_i$ is a function of the channel vectors $\mathbf{z}_k$, the noise covariance matrix $\mathbf{K}$, which is assumed unknown and the transmitted bits $\mathbf{b}_i$. We assume that the bits $\mathbf{b}_i$ are known since, otherwise, the maximization of the likelihood function as a function of all the above unknowns is an ill-posed problem. This is accomplished in the acquisition phase by requiring that all the users transmit training sequences or in the tracking phase by the receiver operating in a decision-directed mode.

Given L observations $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_L$, these are conditionally independent given $\mathbf{b} = [\mathbf{b}'_0, \mathbf{b}'_1, \ldots, \mathbf{b}'_L]'$, the transmitted bits. We form their joint conditional probability density function as

$$p(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_L \mid \mathcal{U}, \mathbf{Z}, \mathbf{b}, \mathbf{K}) = \frac{1}{\pi^{MNL}|\mathbf{K}|^L} \exp\left\{-\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H \mathbf{K}^{-1}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)\right\}, \tag{22}$$

where $|\cdot|$ represents the determinant operator. The corresponding log-likelihood function Λ is

$$\Lambda = -\ln|\mathbf{K}| - \operatorname{tr}\left\{\frac{1}{L}\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H \mathbf{K}^{-1}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)\right\} \tag{23}$$

where $\operatorname{tr}(\cdot)$ represents the trace operator.

As the first step, maximization of the log-likelihood function is to be carried out with respect to $\mathbf{K}$. The maximization is achieved by the following value of $\mathbf{K}$ [13]:

$$\widehat{\mathbf{K}}(\mathcal{U}, \mathbf{Z}) = \frac{1}{L}\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H. \tag{24}$$

Substituting this into (23), in the next step, we find that we need to maximize $-\ln|\widehat{\mathbf{K}}|$ or minimize $|\widehat{\mathbf{K}}|$ over all $\{\mathbf{Z}\}$. The cost function that we need to minimize with respect to $\mathbf{Z}$ is now:

$$\mathcal{L} = |\widehat{\mathbf{K}}| = |\frac{1}{L}\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H|. \tag{25}$$

Details regarding derivation of the cost function (25) by maximizing the log likelihood function (23) with respect to $\mathbf{K}$ is outlined in Appendix A.

Direct minimization of (25) with respect to $\mathbf{Z}$ is rather intractable and hence, it is carried out indirectly in two steps [23, 15]:

(i) Capture the effect of all the unknowns in a single complex matrix $\mathcal{Y} \stackrel{\triangle}{=} \mathcal{U}\mathbf{Z}$. Form the unconstrained ML estimate of $\mathcal{Y}$, given by $\widehat{\mathcal{Y}}$.

(ii) Obtain the estimates $\hat{\mathbf{z}_k}$ by minimizing the weighted least squares fit between $\mathcal{Y} = \mathcal{U}\mathbf{Z}$ and its unstructured estimate, $\widehat{\mathcal{Y}}$.

3.1 Step 1: Covariance approximation

Following the algorithm development in [15, 23] $\widehat{\mathcal{Y}}$ which denotes the unconstrained ML estimate of $\mathcal{Y}$, can be derived by minimizing the cost function (25), with respect to $\mathcal{Y}$.

Such a minimization yields expressions for $\widehat{\mathcal{Y}}$ and $\widehat{\mathbf{K}}(\mathcal{Y})$ in terms of certain sample correlation matrices as

$$\widehat{\mathcal{Y}} = \widehat{\mathbf{R}}_{rb}\widehat{\mathbf{R}}_{bb}^{-1}, \tag{26}$$

$$\widehat{\mathbf{K}}(\widehat{\mathcal{Y}}) = \widehat{\mathbf{R}}_{rr} - \widehat{\mathcal{Y}}\widehat{\mathbf{R}}_{rb}^{H} \tag{27}$$

The sample correlation matrices used in (26) and (27) are defined as

$$\widehat{\mathbf{R}}_{rr} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{r}_i\mathbf{r}_i^H, \qquad \widehat{\mathbf{R}}_{br} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{b}_i\mathbf{r}_i^H, \qquad \widehat{\mathbf{R}}_{bb} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{b}_i\mathbf{b}_i'. \tag{28}$$

Thus the unconstrained estimates are derived completely from the received observations and a knowledge of the training sequences. Details regarding derivation of expressions for $\widehat{\mathcal{Y}}$ and $\widehat{\mathbf{K}}(\mathcal{Y})$ by minimizing the cost function (25), are outlined in Appendix B.

3.2 Step 2: Channel impulse response estimation

The expressions for $\widehat{\mathcal{Y}}$ and $\widehat{\mathbf{K}}$ are substituted into the cost function. The modified cost function is then optimized with respect to $\mathbf{Z}$. This process yields an elegant expression for the estimate of the channel impulse response of each user [15, 23] as follows

$$\hat{\mathbf{z}}_k = \arg\min_{\mathbf{z}_k}\left[(\mathbf{y}_{2k-1} - \hat{\mathbf{y}}_{2k-1})^H\widehat{\mathbf{K}}^{-1}(\mathbf{y}_{2k-1} - \hat{\mathbf{y}}_{2k-1}) + (\mathbf{y}_{2k} - \hat{\mathbf{y}}_{2k})^H\widehat{\mathbf{K}}^{-1}(\mathbf{y}_{2k} - \hat{\mathbf{y}}_{2k})\right] \tag{29}$$

For details regarding derivation of (29), please see Appendix C.

For each user we have contributions for right and left signal vectors, i.e., the $2k^{th}$ and $(2k-1)^{th}$ columns of $\mathcal{Y}$. Let us recall that the columns of $\mathcal{Y}$ are nothing else but the columns of matrix $\mathcal{U}\mathbf{Z}$ as expressed in (20): $\mathbf{y}_{2k-1} = \mathcal{U}_k^R\mathbf{z}_k$, $\mathbf{y}_{2k} = \mathcal{U}_k^L\mathbf{z}_k$ Hence, the estimate of $\mathbf{z}_k$ is :

$$\widehat{\mathbf{z}_k} = \arg\min_{\mathbf{z}_k}\left[(\mathcal{U}_k^R\mathbf{z}_k - \hat{\mathbf{y}}_{2k-1})^H\widehat{\mathbf{K}}^{-1}(\mathcal{U}_k^R\mathbf{z}_k - \hat{\mathbf{y}}_{2k-1}) + (\mathcal{U}_k^L\mathbf{z}_k - \hat{\mathbf{y}}_{2k})^H\widehat{\mathbf{K}}^{-1}(\mathcal{U}_k^L\mathbf{z}_k - \hat{\mathbf{y}}_{2k})\right] \tag{30}$$

Solving this optimization problem using Lemma B.1, we obtain an expression for $\hat{\mathbf{z}}_k$ as follows

$$\hat{\mathbf{z}}_k^H = (\hat{\mathbf{y}}_{2k-1}^H\widehat{\mathbf{K}}^{-1}\mathcal{U}_k^R + \hat{\mathbf{y}}_{2k}^H\widehat{\mathbf{K}}^{-1}\mathcal{U}_k^L)(\mathcal{U}_k^{R'}\widehat{\mathbf{K}}^{-1}\mathcal{U}_k^R + \mathcal{U}_k^{L'}\widehat{\mathbf{K}}^{-1}\mathcal{U}_k^L)^{-1} \tag{31}$$

So, we have a closed form expression for the channel impulse response vector $\hat{\mathbf{z}}_k$ for each user. In the next section we will show how this estimated channel impulse response vector can be used directly in the detection process, exploiting both multipath and spatial diversity.

3.3 Extraction of channel parameters

Most existing literature [7, 15, 16] on channel estimation techniques focus on estimating individual parameters for each user, such as, delay and attenuation factors for one or multiple paths. For our model, in the *single sensor* scenario, extraction of individual parameters from the estimated channel impulse response (31) is a relatively easy task [24]. In Appendix D, we provide an algorithm, following [7], for extraction of individual channel parameters (delays and amplitudes) from the estimated channel impulse response vector, *with a single sensor*.

However when multiple sensors are used, and the various multipath components arrive at the receiver from different directions, the problem of fitting the estimate $\hat{\mathbf{z}_k}$ to the parametric channel model is much more complex. To simplify the computation, in this paper, we propose to use the estimate $\hat{\mathbf{z}_k}$ directly in detection algorithms. Using this approach it will be possible to exploit the benefits of spatial diversity from using multiple sensors, without having to extract the individual delays, amplitudes, and directions of arrival of all the paths of each user.

In the next section we show how the estimated composite multipath channel impulse response from multiple sensors can be used directly in the detection process in an efficient manner. This will not only avoid the computation involved in extracting individual parameters but also avoid any error arising from the extraction process. Also, this technique facilitates the design of a base station receiver which will estimate the channel and detect the information bits while exploiting and accounting for multiple paths and multiple sensors.

4 Use of the Channel Impulse Response Vector in Detection

In this section, we explore the use of the composite channel impulse response, from the multiple sensors, in the conventional detector and in a linear multiuser detector. We propose to use the channel impulse response vector directly in the detection algorithms, instead of first extracting individual parameters. We restrict ourselves to *single-shot* detectors; however the gains from using multiple sensors in this case can also be extended to sequence detection.

The conventional detector for the received signal at a single sensor, in a single path environment, is a bank of K code-matched filters. For the k^{th} user, the matched filter matched to the single path is

$$y_k^{(1)}(i) = \int_{iT}^{(i+1)T} \mathrm{w}_k r^{(1)}(t) c_k(t - iT - \tau_k) dt, \tag{32}$$

where $\mathbf{y}_k^{(1)}(i)$ is the matched filter output at a single sensor for the k^{th} user at the i^{th} time step. The outputs of the matched filters yield *soft* estimates of the transmitted bits. The final *hard* decisions are the sign of these *soft* estimates.

The discretized version of the received signal at a single sensor, for a single path for each user, is expressed as $\mathbf{r}_i^{(1)} = \mathbf{A}^{(1)}\mathbf{b}_i + \boldsymbol{\nu}_i^{(1)}$, where the superscript (1) always denotes the instance $M = 1$. Using this discretized signal model, the equation for the matched filter output, for all the users, is expressed in vector form as :

$$\mathbf{y}_i^{(1)} = (\mathbf{A})^{(1)^H} \mathbf{r}_i^{(1)}. \tag{33}$$

It follows from Section 2 that, in a single path environment, and for a single sensor ($M = 1$), $\mathbf{A}^{(1)} = (\mathcal{U}\mathbf{Z})^{(1)}$. Using this expression for $\mathbf{A}^{(1)}$ we can rewrite the above equation in terms of the channel impulse response matrix:

$$\mathbf{y}_i^{(1)} = (\mathcal{U}\mathbf{Z})^{(1)^H} \mathbf{r}_i^{(1)}. \tag{34}$$

It can be shown quite easily that (34) continues to hold for the matched filter output in the multipath case, when the received signal at each sensor is given by (4). As in Section 2, $\mathbf{Z}$ incorporates the information about all the paths of all the users.

In this model, (34) is similar to the notion of *received effective signature waveform*, described in [9, 25]. The received effective signature waveform is the spreading code of a user modified with the corresponding channel parameters, which is exactly the result of $\mathbf{UH}$, for all the users.

Extending the same idea to multiple sensors, the combined output of a matched filter receiver at each of the sensors, is given by:

$$\mathbf{y}_i = (\mathcal{U}\mathbf{Z})^H \mathbf{r}_i. \tag{35}$$

where $\mathbf{y}_i \in \mathbb{C}^{2K}$. Since all the delay as well as the angle of arrival information is captured in $\mathbf{Z}$, and $\mathbf{z}_k^{(m)} = \mathbf{p}_k^{(m)} \circ \mathbf{h}_k$, (35) can be interpreted as a code-matched filter operation at each sensor, followed by an 'angle-matched filter' or conventional beamformer. Thus the compound matched filters at each sensor are matched to the users' delays as well as direction of arrivals, which would be optimum for the multisensor case for a single user only, in a added white Gaussian noise (AWGN) channel.

The idea of using the composite channel impulse response vector in detection can also be applied to linear multiuser detectors, with multiple sensors [20]. Linear multiuser detectors apply a linear mapping G, to the soft output of the matched filters to reduce the multiple access interference seen by each user.

Therefore, the vector of decision statistics for all the users, from a linear *single-shot* detector, $G \in \mathbb{C}^{2K \times 2K}$, is given by :

$$\mathbf{x}_i = G\mathbf{y}_i = G(\mathcal{U}\mathbf{Z})^H \mathbf{r}_i. \tag{36}$$

The required bit estimates are the hard decision: $\hat{\mathbf{b}}_i = sign(\mathbf{x}_i)$. According to our system model, $\mathbf{b}_i$ contains consecutive bits of each user. Hence the decision statistic for each user will consist of two *soft* estimates for each bit of the user. The two *soft* estimates can be combined such that the output signal-to-interference ratio is maximized [20], to obtain a *hard* decision for each bit.

The decorrelating detector, G_{dec} provides the maximum likelihood linear estimate of the received bits, given the output of the code-matched filters [2]. Hence, it follows that:

$$G_{dec} = ((\mathcal{U}\mathbf{Z})^H(\mathcal{U}\mathbf{Z}))^{-1}. \tag{37}$$

The benefits of incorporating multiple sensors in the decorrelating detector has been analyzed in [20] in the single path case. In a following section, we will report results of simulations that show the performance gains obtained from such a detector, using the composite channel impulse response vector, in the multipath case.

5 Computational Complexity

The computational complexity of the various steps of the ML channel estimation algorithm are :

- Covariance approximation
  - Calculation of sample correlation matrices in (28) - $O(M^2N^2)$, $O(NMK)$, $O(K^2)$
  - Calculation of $\widehat{\mathcal{Y}}$, $\widehat{\mathbf{K}}(\widehat{\mathcal{Y}})$ in (26) and (27) - $O(K^3)$, $O(MNK^2)$, $O(KM^2N^2)$
- Channel impulse response estimation in (31) - $O(M^3N^3)$
- Extraction of parameters, for single sensor case (Section 3.3) - $O(N)$

From the list of the complexity of all the steps in the algorithm, shown above, it is evident that the overall complexity of the algorithm is $O(M^3N^3)$. However the complexity will be reduced further if $\widehat{\mathbf{K}}$ is assumed to be identity (the noise being additive white Gaussian) instead of explicitly calculating the covariance. This whiteness assumption is practical not only because thermal noise is often assumed white Gaussian but also because the residual noise due to other-cell interference can be considered spatially and temporally white.

Other than eliminating the computation of $\widehat{\mathbf{K}}$, this will also reduce the complexity of the channel impulse response estimation step to $O(M^2N^2)$. The inverse of $\widehat{\mathbf{R}}_{bb}$ can be estimated using matrix inversion update algorithms ($O(K^2)$), using the Sherman-Morrison-Woodbury formula [26], or can be pre-calculated as the preamble is a known sequence of bits. After these computation-saving approximations, the dominant computation will be the matrix-vector multiplication of complexity $O(M^2N^2)$ which can benefit greatly from well-known parallelization techniques on a number of processors [26].

It has been shown that the parallelization of matrix-vector multiplication is scalable with the number of processors used, that is, a speedup of a factor of C can be obtained by using C processors. It can be shown easily that it is possible to multiply an MN length vector and an $MN \times MN$ matrix in $O(MN)$ time, using a systolic array of $O(MN)$ processor elements.

6 Simulation Results

In this section, we describe the simulations that we conducted to evaluate the performance of the proposed estimators. Gold codes of length $N = 31$ were used as spreading codes, in all the simulations. The delays of all the users were assumed uniformly distributed in $[1, 31)$ chips. The multiple access interference presented by each interferer, which is the ratio of the interferer's and desired user's received energies, was uniformly distributed in $[0, MAI]$dB. The number of paths for each user is denoted by P, the number of observations by L, the number of users by K, and the signal-to-noise ratio of the background noise at each sensor by SNR.

6.1 Comparison with other estimators - single sensor case

We compare the proposed algorithm with previously proposed channel parameter estimation algorithms for a single sensor, in a multipath environment. The single sensor scenario has been chosen as the other techniques used in this experiment do not address the problem of channel estimation in the presence of multiple paths as well as multiple sensors. The performance of our algorithm in this complex scenario is presented later.

In this experiment, channel parameter estimation is performed using the proposed ML algorithm, a subspace-based algorithm presented in [7], and the conventional sliding correlator [27]. The channel estimates from the various sources as well as the exact value of the channel parameters used in the simulations are then passed to a matched filter receiver to detect the bits. The bit error rates obtained from this experiment is shown in Figure 3. The proposed algorithm performs consistently better than the other strategies. However, the subspace based algorithm does not require a preamble but is more computationally complex and suffers from being dimension-limited, in terms of the number of users it can handle.

6.2 Use of the Channel Impulse Response Vector in Detection

In this section, we illustrate the performance benefits to be obtained from using the estimated channel impulse response vector, directly in the detection step, instead of first extracting individual parameters.

Figure 4 shows the bit error rate (BER) obtained from the matched filter receiver as the SNR is increased. The results for the single sensor case show that the performance actually improves if the channel impulse response is directly used in the detector as in (34) (plot labeled 'z directly, M=1') instead of first extracting the individual parameters from the composite channel impulse response and then using them as in (32) (plot labeled 'parameters, M=1'). The technique outlined in Appendix D was used to extract the individual delays and amplitudes of all the paths of all the users from the estimated channel impulse response of each user.

Figure 4 also shows that the performance of the ML algorithm followed by matched filter detection improves dramatically as the number of sensors are increased. In the experiments described in this section and the next, the direction of arrivals were assumed to be uniformly distributed in $[-60^\circ, 60^\circ]$, corresponding to one sector of 120° in a cell. The values of the rest of the system parameters are specified along with the graphs.

6.3 White noise assumption

As mentioned in Section 2, the additive noise at each sensor is assumed to have some correlation across the array, and has two components. The component $\eta(t)$ is completely correlated across the array and the component $\eta^{(m)}(t)$ is independent from sensor to sensor.

The maximum likelihood algorithm developed in this paper, provides a mechanism for estimating the covariance $\widehat{\mathbf{K}}$ of this non-white noise. However, as discussed in Section 5, estimation of this covariance is one of the most computationally complex steps in the algorithm. It was suggested that to reduce complexity, $\widehat{\mathbf{K}}$ may be assumed to be identity. This implies that the algorithm will assume that the noise is additive white Gaussian, even though in reality, it will have some correlation across the array.

In this section, we present a simulation study (Figure 5), which shows the performance loss that will result from this assumption. The parameter β which defines the extent of the correlation across the array was specified to be quite high ($\beta = 0.9999$). However the simulation shows that the performance loss due to this assumption is very small. Also, the performance loss is evident only if a large window size is used ($L = 1000$ in Figure 5(a)), so that the noise covariance can be estimated with greater accuracy.

6.4 Performance of multiple sensors and multipath case

In this section, we focus on detailed evaluation of the channel estimation and linear detection strategy developed in the paper. Multiple propagation paths are considered and both the detection and channel estimation steps exploit the benefits of multiple sensors.

Figure 6 shows the performance of the decorrelating detector using multiple sensors and compares it to the conventional or the matched filter detector, as the different system parameters are varied. For both the decorrelating detector and the matched filter detector, channel parameter estimates from the proposed maximum likelihood algorithm under the same system configuration have been used.

In Figure 6, the default set of system parameters are such that, the system is heavily loaded (25 users with 3 paths per user) and the interference and noise is quite high (multiple access interference is 20dB and SNR is 6dB). Figure 6(a), with SNR in the x-axis, shows that for the heavily loaded situation, the single-shot decorrelating detector performs quite poorly with only one sensor. However, by increasing the number of sensors, the performance improves dramatically to achieve acceptable bit-error-rates at relatively low SNRs.

It is well known that the performance of the decorrelating detector, with one sensor, drops significantly as the number of users is increased (Figure 6(b)), with number of users in the x-axis). Geometrically speaking, since the decorrelating detector attempts to orthogonalize the users, with more than $N/2$ users in the system, they cannot be orthogonalized all at once and hence the performance degradation. Again, the use of even one extra sensor significantly improves the performance in the presence of a large number of users. The same plot shows that, the matched filter detector performs worse than the decorrelator because of the high multiple access interference. Even then, the matched filter detector performs better as the number of sensors are increased, especially when the number of users and hence multiple access interference is small.

Figure 6(c), with L in the x-axis, shows that a relatively short preamble (less than 100 bits) can be used to achieve optimum performance. It may be noted that for the proposed algorithm to work, at least MN observations are required. This is because, for $L < MN$, the covariance matrix, which has dimension $MN \times MN$, is rank deficient. This observation is substantiated by this plot.

The last plot Figure 6(d) demonstrates the near-far resistant capability of the algorithm, which improves with increasing the number of sensors. It may be noted that the slight increase in BER at high MAI is due to the use of the *single-shot* decorrelator.

7 Conclusion

The paper presents a framework that facilitates an efficient solution to the combined problem of multipath channel estimation and detection for a set of transmitting users in the reverse link of a wireless CDMA communication system, when an antenna array is used at the receiver. Using the framework we have developed a maximum likelihood algorithm for channel estimation. The algorithm elegantly decomposes the multiuser problem into a series of single-user ones and estimates a composite channel impulse response encapsulating all the channel parameters. This estimated channel impulse response was then directly used in the detection process.

The additive noise in the system is assumed to be zero-mean, Gaussian but no assumption is made on its covariance, which is estimated within the algorithm. Our simulations verify that the algorithm is near-far resistant. Also, the estimators are not dimensionally limited; making it ideal for acquisition of a large number of users. Furthermore, the preamble required is not prohibitively large.

Our simulations show that considerable performance benefits are obtained by using multiple sensors at the receiver, in the presence of multipath effects and multiple access interference. We have shown that it is possible and indeed beneficial to use the estimated channel impulse response directly in detection to appropriately steer the spreading codes, both in terms of performance as well as reduced computational load.

In future, the proposed algorithm will be extended to tracking the channel [10] in a decision-directed mode after the initial estimation during the preamble period. Ultimately this work is aimed towards the design of a multiuser receiver for a time varying and fading multipath environment, exploiting spatial as well as temporal diversity, that will be suitable for implementation in next generation CDMA systems.

A Derivation of (25)

The log-likelihood function (23) is

$$\Lambda = -\ln|\mathbf{K}| - \operatorname{tr}\left\{\frac{1}{L}\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H\mathbf{K}^{-1}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)\right\}$$

$$= -\ln|\mathbf{K}| - \operatorname{tr}\left\{\mathbf{K}^{-1}\frac{1}{L}\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H\right\}. \tag{38}$$

The above equality follows from the following property of the trace: $\operatorname{tr}\{\mathbf{ABC}\} = \operatorname{tr}\{\mathbf{CAB}\} = \operatorname{tr}\{\mathbf{BCA}\}$. As the first step, maximization of the log-likelihood function is to be carried out with respect to $\mathbf{K}$. We use the following fact, ([13]) to find the maximum with respect to $\mathbf{K}$:

Fact: If $\mathbf{B}$ is a positive definite Hermitian matrix, then $f(\mathbf{K}) = -\ln|\mathbf{K}| - \mathrm{tr}\{\mathbf{K}^{-1}\mathbf{B}\}$ is maximized by $\widehat{\mathbf{K}} = \mathbf{B}$.

Hence we have,

$$\widehat{\mathbf{K}}(\mathcal{U}, \mathbf{Z}) = \frac{1}{L}\sum_{i=1}^{L}(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)(\mathbf{r}_i - \mathcal{U}\mathbf{Z}\mathbf{b}_i)^H. \tag{39}$$

Substituting this value of $\widehat{\mathbf{K}}$ into (38), we obtain the cost function (25).

B Derivation of (26) and (27)

In terms of the sample correlation matrices, $\widehat{\mathbf{K}}(\mathcal{U}, \mathbf{Z})$ or $\widehat{\mathbf{K}}(\mathcal{Y})$ from (24) can be written as

$$\widehat{\mathbf{K}}(\mathcal{Y}) = \widehat{\mathbf{R}}_{rr} - \mathcal{Y}\widehat{\mathbf{R}}_{rb}^H - \widehat{\mathbf{R}}_{rb}\mathcal{Y}^H + \mathcal{Y}\widehat{\mathbf{R}}_{bb}\mathcal{Y}^H. \tag{40}$$

Minimizing the cost function $\mathcal{L}$ in (25), $\widehat{\mathcal{Y}}$, the unconstrained ML estimate of $\mathcal{Y}$ is

$$\widehat{\mathcal{Y}} = \arg\min_{\mathcal{Y}} |\widehat{\mathbf{K}}(\mathcal{Y})| = \arg\min_{\mathcal{Y}} |\widehat{\mathbf{R}}_{rr} - \mathcal{Y}\widehat{\mathbf{R}}_{rb}^H - \widehat{\mathbf{R}}_{rb}\mathcal{Y}^H + \mathcal{Y}\widehat{\mathbf{R}}_{bb}\mathcal{Y}^H|. \tag{41}$$

We use the following lemma [20], stated without proof, to obtain an expression for $\widehat{\mathcal{Y}}$:

Lemma B.1: If $\mathbf{P}$, $\mathbf{C} \in \mathbb{C}^{A\times B}$, and $\mathbf{R} \in \mathbb{C}^{A\times A}$ are matrices and A and B are some integers, then for any positive definite Hermitian matrix $Q \in \mathbb{C}^{B\times B}$, we have

$$\arg\min_{\mathbf{C}} \left|\mathbf{R} - \mathbf{P}\mathbf{C}^H - \mathbf{C}\mathbf{P}^H + \mathbf{C}\mathbf{Q}\mathbf{C}^H\right| = \mathbf{P}\mathbf{Q}^{-1}. \tag{42}$$

Applying this Lemma to (41), $\widehat{\mathcal{Y}}$ can be expressed as

$$\widehat{\mathcal{Y}} = \widehat{\mathbf{R}}_{rb}\widehat{\mathbf{R}}_{bb}^{-1}. \tag{43}$$

Substituting back into the equation for $\widehat{\mathbf{K}}$,

$$\widehat{\mathbf{K}}(\widehat{\mathcal{Y}}) = \widehat{\mathbf{R}}_{rr} - \widehat{\mathcal{Y}}\widehat{\mathbf{R}}_{rb}^H \tag{44}$$

C Derivation of (29)

Using expressions for $\widehat{\mathcal{Y}}$ and $\widehat{\mathbf{K}}$ in (40, our cost function (25) now becomes

$$\begin{aligned}\mathcal{L} &= |\widehat{\mathbf{K}}(\mathcal{Y})| \\ &= |\widehat{\mathbf{R}}_{rr} - \mathcal{Y}\widehat{\mathbf{R}}_{rb}^H - \widehat{\mathbf{R}}_{rb}\mathcal{Y}^H + \mathcal{Y}\widehat{\mathbf{R}}_{bb}\mathcal{Y}^H| \\ &= \left|\widehat{\mathbf{K}}\right| \cdot \left|\mathbf{I} + \widehat{\mathbf{K}}^{-1}(\mathcal{Y} - \widehat{\mathcal{Y}})\widehat{\mathbf{R}}_{bb}(\mathcal{Y} - \widehat{\mathcal{Y}})^H\right|.\end{aligned} \tag{45}$$

We note that minimizing $\left|\mathbf{I}+\widehat{\mathbf{K}}^{-1}(\mathcal{Y}-\widehat{\mathcal{Y}})\widehat{\mathbf{R}}_{bb}(\mathcal{Y}-\widehat{\mathcal{Y}})^H\right|$ is equivalent to minimizing $\ln\left|\mathbf{I}+\widehat{\mathbf{K}}^{-1}(\mathcal{Y}-\widehat{\mathcal{Y}})\widehat{\mathbf{R}}_{bb}(\mathcal{Y}-\widehat{\mathcal{Y}})^H\right|$. Using properties of eigenvalues and norms of matrices, it can be shown [20] that:

$$\min_{\mathbf{Z}} \ln\left|\mathbf{I}+\widehat{\mathbf{K}}^{-1}(\mathcal{Y}-\widehat{\mathcal{Y}})\widehat{\mathbf{R}}_{bb}(\mathcal{Y}-\widehat{\mathcal{Y}})^H\right| \stackrel{\equiv}{\longrightarrow} \min_{\mathbf{Z}} \operatorname{tr}\{\widehat{\mathbf{R}}_{bb}(\mathcal{Y}-\widehat{\mathcal{Y}})^H\widehat{\mathbf{K}}^{-1}(\mathcal{Y}-\widehat{\mathcal{Y}})\}, \tag{46}$$

where the symbol $\stackrel{\equiv}{\longrightarrow}$ denotes "is asymptotically equivalent to" (the reader is reminded of the scalar case where $\ln(1+x) \approx x$ for small x). This greatly simplifies the problem since it replaces the nonlinear functional $|\cdot|$ by a linear one $\operatorname{tr}\{\cdot\}$. The problem now reduces to

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \operatorname{tr}\left\{\widehat{\mathbf{R}}_{bb}(\mathcal{Y}-\widehat{\mathcal{Y}})^H\widehat{\mathbf{K}}^{-1}(\mathcal{Y}-\widehat{\mathcal{Y}})\right\} \tag{47}$$

Since the received signals of the different users are "uncorrelated", $\widehat{\mathbf{R}}_{bb}$ is diagonal for large L. This diagonal form helps to separate the estimation process of each user and yields the expression (29) for the estimate of the k^{th} user's channel impulse response.

D Extraction of channel parameters from (31) in single sensor case

In the single sensor case, the estimated channel impulse response vector $\widehat{\mathbf{z}}_k$ is the same as component $\widehat{\mathbf{h}}_k$ only. We perform a least squares fit of $\widehat{\mathbf{h}}_k$ to our parametric channel model (13) to extract the strongest P paths. For each pair of adjacent coefficients of $\widehat{\mathbf{h}}_k$, we obtain *local* values of amplitudes and delays, from the following optimization:

$$[\mathrm{w}_q, \gamma_q] = \arg\min_{w\in\mathbb{C},\gamma\in[0.1)} ||\widehat{h}_{k,q}-(1-\gamma)w||^2 + ||\widehat{h}_{k,q+1}-\gamma w||^2. \tag{48}$$

We then search for the *global* maxima to obtain the strongest path :

$$\begin{aligned} \hat{q} &= \arg\max_{q\in\{0,\cdots,N-1\}} |w_q| \\ \hat{\tau} &= (\hat{q}+\gamma_{\hat{q}})T_c, \qquad \hat{w} = w_{\hat{q}}. \end{aligned} \tag{49}$$

The estimated path is subtracted from $\widehat{\mathbf{h}}_k$ and the process is repeated to find the next strongest path, until either a specified number of paths have been identified or $|\hat{w}|$ falls below some predetermined significant level.

References

[1] S. Verdú. Minimum probability of error for asynchronous Gaussian multiple-access channels. *IEEE Trans. Information Theory*, IT-32:85–96, January 1986.

[2] R. Lupas and S. Verdú. Linear multiuser detectors for synchronous code-division multiple access channels. *IEEE Trans. Information Theory*, IT-35:123–136, January 1989.

[3] M. K. Varanasi and B. Aazhang. Multistage detection in asynchronous code-divison multiple access communications. *IEEE Trans. Communications*, COM-38:509–519, April 1990.

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

[5] S. Y. Miller and S.C. Schwartz. Paramater estimation for asynchronous multiuser communication. *Conference on Information Sciences and Systems, Baltimore, MD*, pages 294–299, March 1989.

[6] M. K. Varanasi and S. Vasudevan. Multiuser detectors for synchronous CDMA communication over non-selective Rician fading channels. *IEEE Trans. Communications*, 42:711–722, February 1994.

[7] S. E. Bensley and B. Aazhang. Subspace-based channel estimation for code division multiple access communication systems. *IEEE Trans. Communications*, 44(8):1009–1020, August 1996.

[8] E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersten. DS-CDMA synchronization in time-varying fading channels. *IEEE Jounal on Selected Areas in Communication*, pages 1636–1642, October 1996.

[9] X. Wang and H. V. Poor. Subspace-based blind adaptive joint interference suppression and channel estimation in multipath CDMA channels. *IEEE International Conference on Universal Personal Communications (ICUPC), San Diego, CA*, pages 460–464, October 1997.

[10] C. Sengupta, J. R. Cavallaro, and B. Aazhang. Subspace-based tracking of multipath channel parameters for CDMA systems. *European Trans. Telecommunications, Special issue on CDMA techniques for wireless communication systems*, pages 439–447, Sep.-Oct. 1998.

[11] X. Wang and H. V. Poor. Blind equalization and multiuser detection in dispersive CDMA channels. *IEEE Trans. Signal Processing*, pages 91–103, January 1998.

[12] S. Parkvall, E. G. Strom, L. B. Milstein, and B. E. Ottersten. Asynchronous near-far resistant DS-CDMA receivers without a priori synchronization. *IEEE Trans. Communications - To appear*, 1998.

[13] D. H. Johnson and D. E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Prentice Hall, Englewood Cliffs, NJ, 1993.

[14] S. E. Bensley and B. Aazhang. Maximum likelihood synchronization of a single user for CDMA communication systems. *IEEE Trans. Communications*, 46(3):392–399, March 1998.

[15] R. Madyastha and B. Aazhang. Antenna arrays for joint maximum likelihood parameter estimation in CDMA systems. *Conference on Information Sciences and Systems, Baltimore, MD*, II:984–988, March 1997.

[16] D. Zheng, J. Li, S. L. Miller, and E. G. Strom. An efficient code-timing estimator for DS-CDMA signals. *IEEE Trans. Signal Processing*, 45(1):82–89, January 1997.

[17] Z-S Liu, J. Li, and S. L. Miller. A receiver diversity based code-timing estimator for asynchronous DS-CDMA systems. *International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Seattle, WA*, May 1998.

[18] Z-S Liu, J. Li, and S. L. Miller. An efficient code-timing estimator for receiver diversity DS-CDMA systems. *IEEE Trans. Communications*, 46(6):826–835, June 1998.

[19] E. Ertin, U. Mitra, , and S. Siwamogsatham. Maximum-likelihood based multipath channel estimation for code-division-multiple-access systems. *IEEE Trans. Communications - submitted*, 1998.

[20] R. Madyastha. *Antenna arrays for wireless CDMA communications systems*. PhD thesis, ECE Dept., Rice University, Houston, TX, 1997.

[21] C. Sengupta, J. R. Cavallaro, and B. Aazhang. Maximum likelihood multipath channel parameter estimation in CDMA systems using antenna arrays. *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Boston, MA*, September 1998.

[22] T. S. Rappaport. *Wireless communications - Principles and Practice*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1996.

[23] J. Li, B. Halder, P. Stoica, and M. Viberg. Computationally efficient angle estimation for signals with known waveforms. *IEEE Trans. Signal Processing*, 43(9):2154–2163, September 1995.

[24] C. Sengupta, A. Hottinen, J. R. Cavallaro, and B. Aazhang. Maximum likelihood multipath channel parameter estimation in CDMA systems. *Conference on Information Sciences and Systems, Princeton, NJ*, March 1998.

[25] X. Wang and H. V. Poor. Blind adaptive interference suppression for CDMA communications based on eigenspace tracking. *Conference on Information Sciences and Systems, Baltimore, MD*, pages 468–473, March 1997.

[26] Gene H. Golub and Charles F. Van Loan. *Matrix computations.* The Johns Hopkins University Press, Baltimore, 2nd edition, 1989.

[27] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread-spectrum communications - A tutorial. *IEEE Trans. Communications*, COM-30(5):855–884, May 1982.

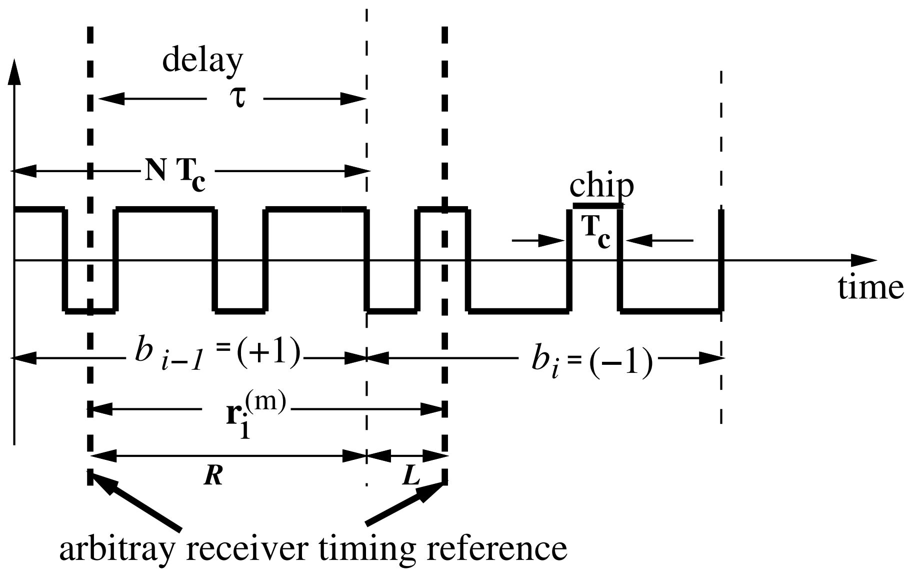

Figure 1: System model - received signal.

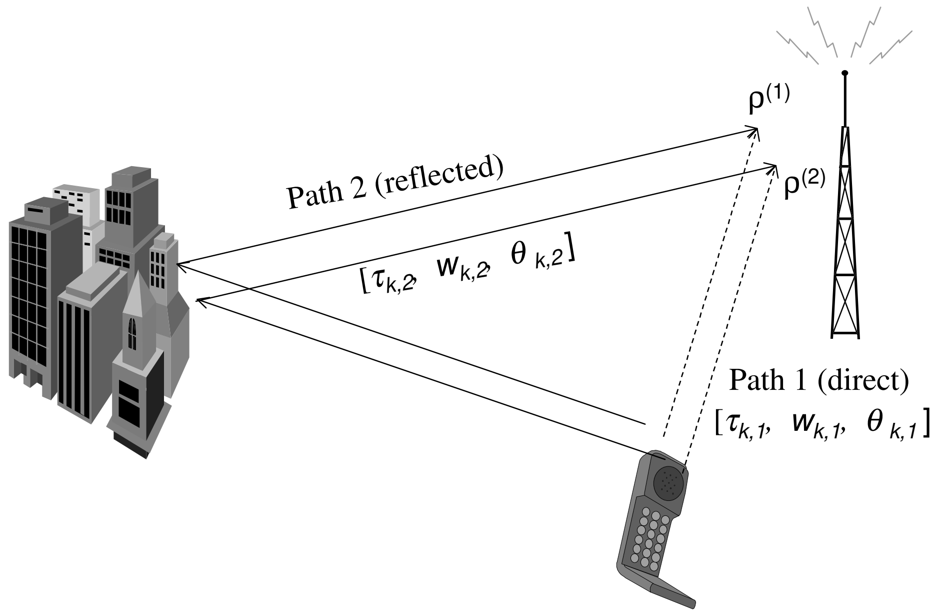

Figure 2: Two propagation paths from a user k, received by 2 sensors at the base station receiver. Each path p, has a different delay $\tau_{k,p}$, amplitude $w_{k,p}$, and direction of arrival $\theta_{k,p}$. Each sensor m, has a different response $p^{(m)}(\theta)$.

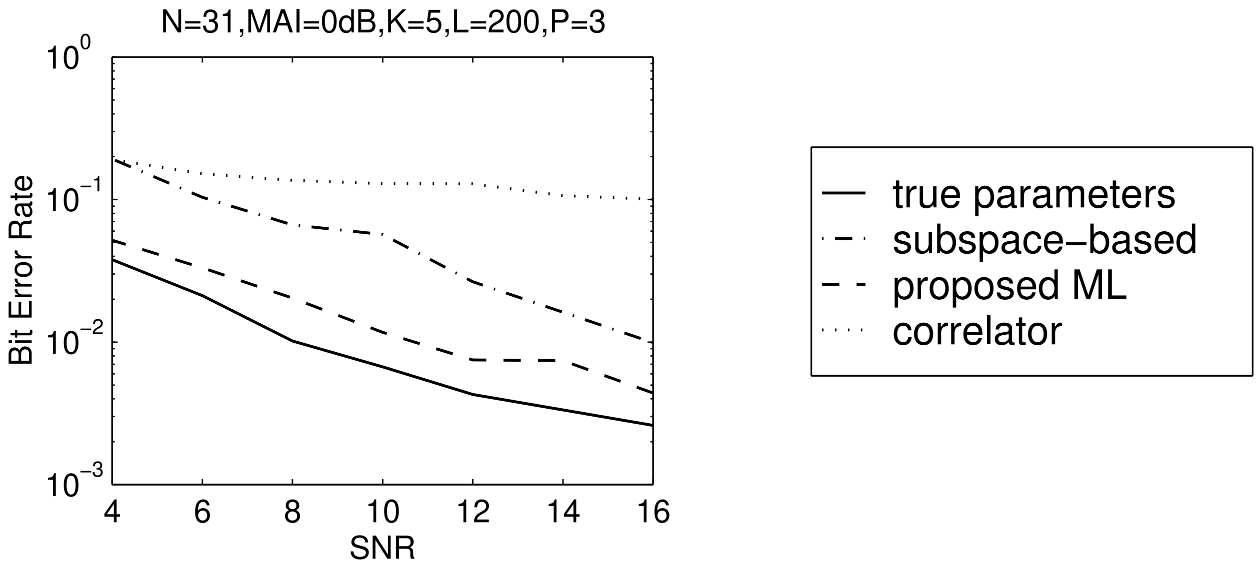

Figure 3: Multipath, single sensor case: BER from a matched filter receiver and channel parameters using (1) true parameters or perfect channel knowledge, (2) estimates from a subspace-based algorithm, (3) estimates from the proposed ML algorithm, (4) and estimates using a sliding correlator.

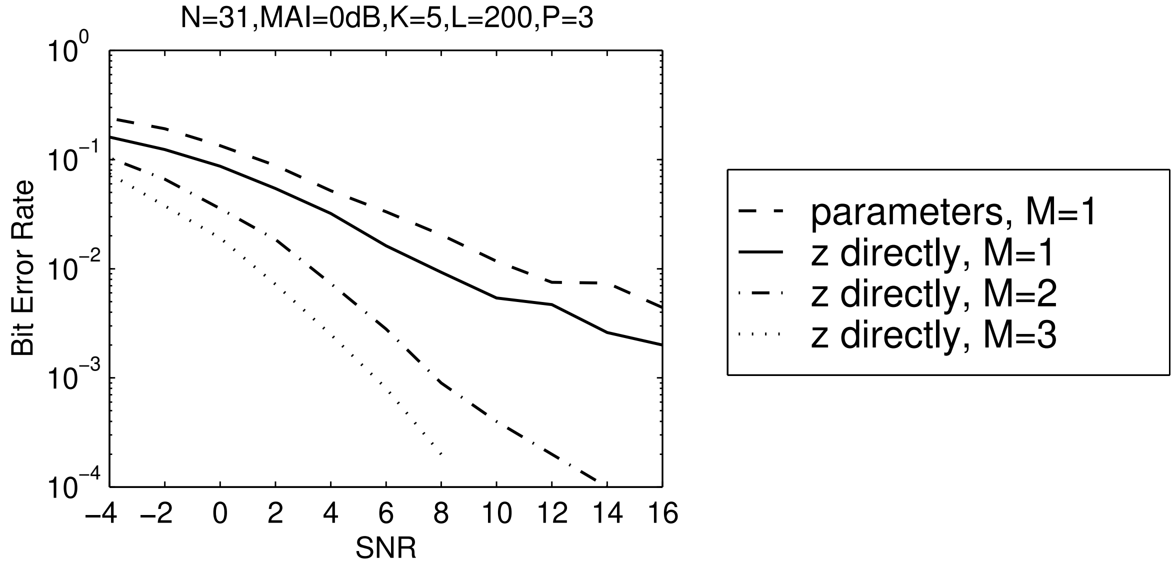

Figure 4: Multipath, multiple sensors (M) case: BER from a matched filter receiver, using composite channel impulse response ($\hat{\mathbf{z}}_k$) directly, compared to the case where the individual parameters are extracted from $\hat{\mathbf{z}}_k$ and then used in the matched filters. $\beta = 0.7$.

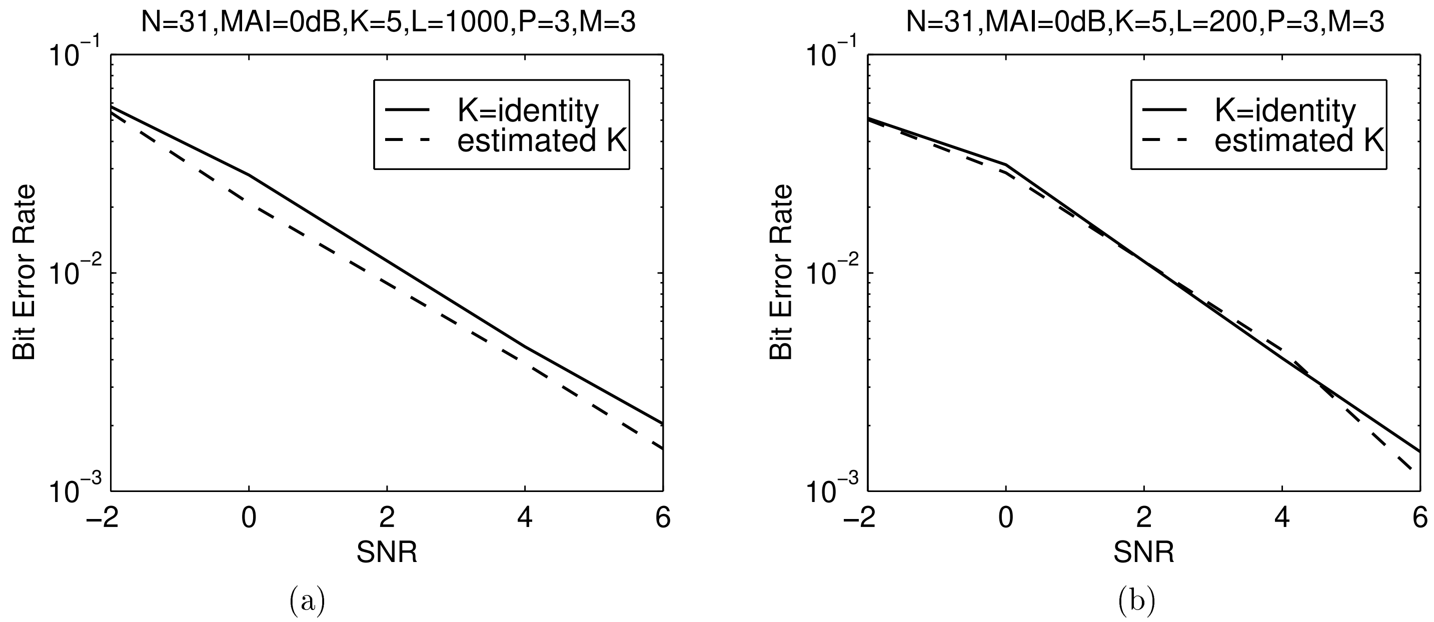

Figure 5: Multipath, multiple sensors case: BER from a matched filter receiver, using estimated noise covariance ($\mathbf{K}$) and noise covariance equal to identity. $M = 3$, $\beta = 0.9999$.

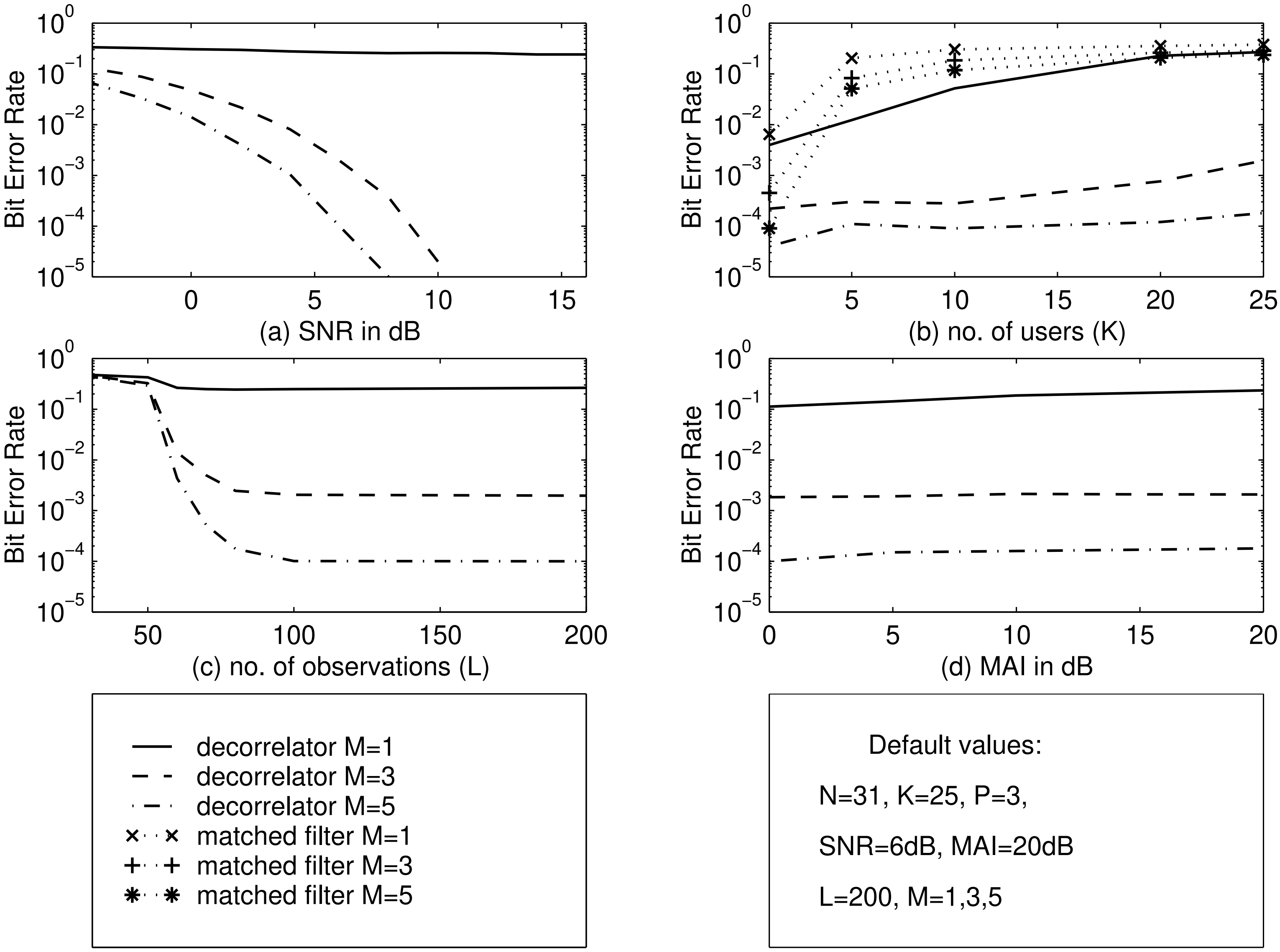

Figure 6: Multipath, multiple sensors case: BER from a matched filter receiver and a *single-shot* decorrelating detector using estimated composite channel impulse response ($\hat{\mathbf{z}_k}$) and multiple sensors (M). $\beta = 0.7$.