Research Article

Sparsity-Aware Estimation of CDMA System Parameters

Daniele Angelosante,[1] Emanuele Grossi,[2] Georgios B. Giannakis,[1] and Marco Lops[2]

[1] *Department of ECE, University of Minnesota, Minneapolis, MN 55414, USA*
[2] *DAEIMI, Università degli Studi di Cassino, 03043 Cassino, Italy*

Correspondence should be addressed to Georgios B. Giannakis, georgios@umn.edu

The number of active users, their timing offsets, and their (possibly dispersive) channels with the access point are performance-critical parameters for wireless code division multiple access (CDMA). Estimating them as accurately as possible using as short as possible training sequences can markedly improve error performance as well as the capacity of CDMA systems. The fresh look advocated here permeates benefits from recent advances in variable selection and compressive sampling approaches to multiuser communications by casting estimation of these parameters as a sparse linear regression problem. Novel estimators are developed by exploiting two forms of sparsity present: the first emerging from user (in)activity, and the second from the uncertainty on user delays and channel taps. Simulated tests demonstrate a large gain in performance when sparsity-aware estimators of CDMA parameters are compared to sparsity-agnostic standard least-squares-based alternatives.

1. Introduction

It is known that the capacity of code division multiple access (CDMA) systems can be enhanced through multiuser detection, which is capable of mitigating the multiaccess interference. A key issue in the design of a multiuser receiver is the development of high-performance parameter estimation algorithms. These parameters include the power, timing offset (possibly dispersive), channel response, and, in some cases, the number of active users accessing the network. For this reason, the problem of CDMA parameter estimation has received considerable attention over the past twenty years.

Maximum-likelihood (ML) and subspace-based approaches have been developed to jointly estimate the timing offsets of multiple users [1]. Single-user ML synchronization schemes have been also developed in [2, 3]. Blind adaptive demodulation algorithms that implicitly estimate timing offsets have been also reported [4]. A blind channel estimator based on the received data correlation has been introduced in [5]; while in [6], assuming that reliable estimates of the user delays are available, several algorithms have been pursued for estimating the remaining parameters.

While ML estimation can be prohibitively complex, subspace-based approaches require long observation records and are thus applicable to slowly varying channels only. Moreover, most CDMA-based wireless networks employ long codes to guarantee that all users achieve "on the average" comparable performance [7–9]. Long codes destroy the cyclostationarity of CDMA waveforms that subspace-based algorithms rely on. For systems employing long codes, a least-squares (LS) estimator of the channel parameters has been proposed in [10], where the users transmit a training sequence and the receiver has perfect knowledge of the number and identity of active users.

In this paper, the results of [10] are considerably broadened through signal processing tools drawn from the recently established theories of variable selection (VS) [11] and compressive sampling (CS) [12, 13]. The common theme underlying VS and CS in the presence of noise is *sparse linear regression*, that is, estimation of a sparse vector from noisy linear observations. Sparsity present in communication systems has been exploited for single-user channel estimation [14–17] and cognitive radio sensing [18]; some preliminary results in the CDMA context are reported also in [19]. In this paper, it is first argued that in most

CDMA parameter estimation tasks the vector to be estimated is sparse due to user inactivity and uncertainty on the users' timing offsets and propagation channels. Subsequently, a sparsity-aware algorithm is developed for estimating channel parameters based on the *least-absolute shrinkage and selection operator* (Lasso) [11]. The novel algorithm can operate with surprisingly short training sequences. This implies enhancement of bandwidth efficiency over the standard LS-based approach. Moreover, the developed algorithm offers considerable gains in error performance even when the number of active users and their identities are unknown. This scenario can emerge in future generation wireless networks based on cognitive radios, where secondary unlicensed users may be sporadically present over the licensed bands [20].

The rest of the paper is organized as follows. In Section 2, the signal model for a CDMA system operating under very general conditions is presented, and sparsity in the vector to be estimated is demonstrated. In Section 3, the problem of CDMA parameter estimation is stated, and a two-stage procedure exploiting the intrinsic sparsity of CDMA signals is developed. Numerical simulations testing the estimation capabilities of the developed algorithm against traditional approaches are presented in Section 4. Finally, concluding remarks are drawn in Section 5.

Notation. In what follows, all random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathbb{E}$ denotes the expectation operator. $\mathbb{R}$ and $\mathbb{C}$ are the sets of real and complex numbers, respectively. $\mathfrak{R}\{\cdot\}$ and $\mathfrak{I}\{\cdot\}$ denote real and imaginary parts. $\mathbb{R}^{m \times n}$ and $\mathbb{C}^{m \times n}$ are the set of $m \times n$ matrices with entries from $\mathbb{R}$ and $\mathbb{C}$, respectively. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote conjugate, transpose, and conjugate transpose, respectively. Column vectors and matrices are indicated through boldface lowercase and uppercase letters, respectively. $(\mathbf{x})_i$ is the ith entry of $\mathbf{x} \in \mathbb{C}^{n \times 1}$, and $\mathbf{A}^\dagger$ is the Moore-Penrose generalized inverse of $\mathbf{A}$. $\|\mathbf{x}\|_1 := \sum_{i=1}^n |\mathbf{x}_i|$ and $\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^n |\mathbf{x}_i|^2}$ are the ℓ_1 and ℓ_2 norms of $\mathbf{x} \in \mathbb{C}^{n \times 1}$, respectively, while $\|\mathbf{x}\|_0$ denotes its (pseudo) ℓ_0 norm, that is, the number of its nonzero entries. Finally, $\star$ denotes the convolution product, while $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ stand for the ceiling and floor functions, respectively.

2. Signal Model

Consider the uplink of an asynchronous direct sequence CDMA system comprising K users, not necessarily all active. Each user experiences an unknown linear channel, possibly frequency selective, but time-invariant during the transmission of a packet. Each packet contains a preamble of P known (training) symbols to be used at the receiver for parameter estimation. Letting T_s denote the symbol period, and $s_{k,p}(t)$ the signature waveform of the kth user during the pth signaling interval, the complex envelope of the received continuous-time waveform can be written as

$$\widetilde{r}(t) := \sum_{k=0}^{K-1} \sum_{p=0}^{P-1} b_k(p) \widetilde{s}_{k,p}(t - \tau_k - pT_s) \star \psi_{\text{ch},k}(t) + \widetilde{w}(t),$$

$$\tag{1}$$

where

   (i) $b_k(p)$ denotes the known pth training symbol transmitted by the kth user;

  (ii) τ_k is the *unknown* kth user's transmission delay (a.k.a. timing offset);

 (iii) $\psi_{\text{ch},k}(t)$ denotes the *unknown* channel impulse response of the kth user's link with the access point, assumed invariant during the transmission of each packet, and with delay spread bounded by Δ_{ch}; user inactivity is accounted for by allowing $\psi_{\text{ch},k}(t) = 0$ if user k is idle;

  (iv) $\widetilde{w}(t)$ stands for the zero-mean, complex, additive, white, Gaussian noise (AWGN) at the receive antenna, with power spectral density $\mathcal{N}_0$.

If N denotes the processing gain, $T_c := T_s/N$ the chip period, and $\{c_{k,p}(n)\}_{n=0}^{N-1}$ the spreading sequence of the kth user, the corresponding signature waveform is

$$\widetilde{s}_{k,p}(t) := \sqrt{\mathcal{E}_k} \sum_{n=0}^{N-1} c_{k,p}(n) \psi_{\text{tx}}(t - nT_c),$$

$$\tag{2}$$

where $\psi_{\text{tx}}(t)$ is the chip waveform with support $[0, \Delta_{\text{tx}}]$, and $\mathcal{E}_k$ is the baseband energy of the kth user. The dependence of $c_{k,p}(n)$ on the symbol index p accounts for the possible adoption of long aperiodic spreading sequences.

At the receiver side, $r(t)$ is fed to a lowpass filter whose impulse response $\psi_{\text{rx}}(t)$ is a unit-energy waveform with support $[0, \Delta_{\text{rx}}]$. The output of this filter is subsequently (over) sampled at rate M/T_c to obtain the samples

$$r(i) := (\widetilde{r} \star \psi_{\text{rx}})\left(\frac{iT_c}{M}\right)$$

$$= \sum_{k=0}^{K-1} \sum_{p=0}^{P-1} b_k(p) \sum_{n=0}^{N-1} c_{k,p}(n) g_k\left(\frac{iT_c}{M} - nT_c - pT_s\right) + w(i),$$

$$\tag{3}$$

where $w(i) := (\widetilde{w} \star \psi_{\text{rx}})(iT_c/M)$ and $g_k(t) := \sqrt{\mathcal{E}_k} \psi_{\text{tx}}(t) \star \psi_{\text{ch},k}(t - \tau_k) \star \psi_{\text{rx}}(t)$. The samples corresponding to the interval $(pT_s, (p+1)T_s]$ can be concatenated to form the $NM \times 1$ complex vector:

$$\mathbf{r}(p) := [r(pNM + 1) \cdots r((p+1)NM)]^T.$$

$$\tag{4}$$

Notice that all unknown quantities (user delays and channel responses) in (3) have been lumped into the samples of the functions $\{g_k(t)\}_{k=0}^{K-1}$. In addition, the support of each $g_k(t)$ is always included in the interval $[\tau_k, \tau_k + \Delta_{\text{tx}} + \Delta_{\text{rx}} + \Delta_{\text{ch}}]$ so that the number of its nonzero samples, call it L_k, is upper bounded by

$$L := \left\lfloor \frac{\Delta_{\text{tx}} + \Delta_{\text{rx}} + \Delta_{\text{ch}}}{T_c/M} \right\rfloor \geq L_k.$$

$$\tag{5}$$

With $\tau_{\max}$ denoting the maximum value that the user delays can assume, and letting

$$Q := \left\lceil \frac{\tau_{\max} + \Delta_{\text{tx}} + \Delta_{\text{rx}} + \Delta_{\text{ch}}}{T_c/M} \right\rceil - 1,$$

$$\tag{6}$$

the nonzero samples of $g_k(t)$ are always contained in the $Q \times 1$ vector:

$$\mathbf{g}_k := \left[g_k\left(\frac{T_c}{M}\right) \cdots g_k\left(\frac{QT_c}{M}\right) \right]^T \tag{7}$$

From the definition of $g_k(t)$, it follows that $g_k(\tau_k) = g_k(\tau_k + \Delta_{\text{tx}} + \Delta_{\text{rx}} + \Delta_{\text{ch}}) = 0$, and $g_k(0) = g_k(\tau_{\max} + \Delta_{\text{tx}} + \Delta_{\text{rx}} + \Delta_{\text{ch}}) = 0$. Observe that the position of the first nonzero entry of $\mathbf{g}_k$ is $d_k + 1$, where $d_k := \lfloor \tau_k M/T_c \rfloor \leq d_{\max}$, and $d_{\max} := \lfloor \tau_{\max} M/T_c \rfloor$.

Upon inspecting (3), it follows that the actual sampled signature of the kth user is a linear combination of the spreading sequence $\{c_{k,p}(n)\}_{n=0}^{N-1}$ weighted through the entries of $\mathbf{g}_k$. In order to exemplify this linear dependence, rewrite the ℓth entry of $r(p)$ as follows:

$r(pNM + \ell)$

$$= \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} \sum_{j=0}^{P-1} b_k(j) c_{k,j}(n) g_k\left(\frac{\ell T_c}{M} - nT_c + (p-j)T_s \right)$$

$$+ w(pNM + \ell) \tag{8}$$

$$= \sum_{k=0}^{K-1} \sum_{q=1}^{Q} \sum_{j=0}^{P-1} b_k(j) c_{k,j}\left[(p-j)N + \frac{(\ell-q)}{M} \right] g_k\left(\frac{qT_c}{M} \right)$$

$$+ w(pNM + \ell),$$

where $c_{k,j}[q/M] := c_{k,j}(q/M)$ if $q = 0, M, \ldots, (N-1)M$, and zero otherwise. Since $c_{k,j}[(p-j)N + (\ell-q)/M)]$ is zero if $p - j \notin \{0, 1, \ldots, \overline{m}\}$, (8) simplifies to

$r(pNM + \ell)$

$$= \sum_{k=0}^{K-1} \sum_{q=1}^{Q} \sum_{i=0}^{\overline{m}} b_k(p-i) c_{k,p-i}\left[iN + \frac{(\ell-q)}{M} \right] g_k\left(\frac{qT_c}{M} \right) \tag{9}$$

$$+ w(pNM + \ell),$$

where $\overline{m} := \lfloor (Q-M)/NM + 1 - 1/NM \rfloor = \lceil (Q-M)/NM \rceil$ represents the number of past interfering symbols. Thus, upon defining the $NM \times Q$ matrices $\{\mathbf{C}_{k,p}\}$ with (ℓ, q)th entries

$$\left[\mathbf{C}_{k,p} \right]_{q=1,\ldots,Q}^{\ell=1,\ldots,NM} := \sum_{i=0}^{\overline{m}} b_k(p-i) c_{k,p-i}\left[iN + \frac{(\ell-q)}{M} \right], \tag{10}$$

the received vector in (4) can be expressed as

$$\mathbf{r}(p) = \sum_{k=0}^{K-1} \mathbf{C}_{k,p} \mathbf{g}_k + \mathbf{w}(p), \tag{11}$$

where $\mathbf{w}(p) := [w(pNM + 1) \cdots w((p+1)NM)]^T$ is a zero-mean circularly symmetric Gaussian vector with $\mathbb{E}[w_i(p) w_j^*(p)] = \mathcal{N}_0 r_{\psi_{\text{rx}}}((i-j)T_c/M)$, and $r_{\psi_{\text{rx}}}$ denotes the autocorrelation of the receive filter. All the unknowns pertaining to the kth user are contained in $\mathbf{g}_k$, often referred to as the *composite channel* in the multiuser detection

parlance [10]. In particular, if user k is idle, then $\mathbf{g}_k$ is the null vector; while if user k is active, then $\mathbf{g}_k$ contains a block of at most L nonzero entries, whose position depends on τ_k, and whose values depend on τ_k and $\psi_{\text{ch},k}$.

Finally, the observations $\{\mathbf{r}(p)\}$ collected during the training phase can be stacked in a $\overline{P}NM \times 1$ vector:

$$\mathbf{r} := \left[\mathbf{r}^T(0) \cdots \mathbf{r}^T(\overline{P} - 1) \right]^T = \mathbf{C}\mathbf{g} + \mathbf{w} \tag{12}$$

with $\overline{P} := P + \overline{m}$, $\mathbf{w} := [\mathbf{w}^T(0) \cdots \mathbf{w}^T(\overline{P} - 1)]^T$, $\mathbf{g} := [\mathbf{g}_0^T \cdots \mathbf{g}_{K-1}^T]^T$, and

$$\mathbf{C} := \begin{bmatrix} \mathbf{C}_{0,0} & \cdots & \mathbf{C}_{K-1,0} \\ \vdots & & \vdots \\ \mathbf{C}_{0,\overline{P}-1} & \cdots & \mathbf{C}_{K-1,\overline{P}-1} \end{bmatrix}. \tag{13}$$

After prewhitening, one arrives at the linear regression model:

$$\mathbf{y} := \mathbf{L}^{-1}\mathbf{r} = \mathbf{S}\mathbf{g} + \mathbf{n}, \tag{14}$$

where $\mathbf{L}$ is the Cholesky factor of the noise covariance matrix, $\mathbf{S} := \mathbf{L}^{-1}\mathbf{C}$, and $\mathbf{n} := \mathbf{L}^{-1}\mathbf{w}$ denotes the zero-mean AWGN.

3. CDMA Parameter Estimation

Based on the data $\mathbf{y}$ in (14) and knowledge of the matrix $\mathbf{S}$, the receiver has to decipher the active users and estimate their composite channels $\{\mathbf{g}_k\}_{k=0}^{K-1}$. These estimates are instrumental for the subsequent multiuser detection phase. Since the kth user delay and channel response have been included in $\mathbf{g}_k$, acquiring this composite channel amounts to estimating the following:

(i) the user activity state, represented by the variables $\{u_k\}_{k=0}^{K-1} \in \{0, 1\}$;

(ii) the user delays in terms of multiples of T_c/M, that is, $d_k \in \{0, 1, \ldots, d_{\max}\}$;

(iii) the channel delay spread, that is, the number of nonzero entries of $\mathbf{g}_k$, $L_k \in \{1, \ldots, L\}$;

(iv) the channel amplitude profile $\widetilde{\mathbf{g}}_k$, which is contained in the values of the nonzero entries of $\mathbf{g}_k$.

The joint user activity detection and channel estimation for all K users can be solved by maximizing the likelihood $f(\mathbf{y} \mid \{u_k\}_{k=0}^{K-1}, \{d_k\}_{k=0}^{K-1}, \{L_k\}_{k=0}^{K-1}, \{\widetilde{\mathbf{g}}_k\}_{k=0}^{K-1})$. However, this procedure incurs exponential complexity in the maximum number of users K. Moreover, if user transmissions undergo multipath propagation, an exhaustive search over all possible subsets of active paths must be carried out for each value of the delay spread; hence, the complexity is exponential also in the multipath delay spread, L. The computational complexity of the ML approach renders this procedure generally infeasible.

For this reason, the problem of CDMA parameter estimation is solved here in two stages.

(1) Estimate the whole vector $\mathbf{g}$ based on $\mathbf{y}$, by minimizing a suitably chosen cost function.

(2) Given estimates $\{\hat{\mathbf{g}}_k\}_{k=0}^{K-1}$, detect the active users and estimate their channels and timing offsets based on a single-user procedure.

While related two-stage approaches have been used in the past (e.g., [10]), the contribution here is twofold: (i) we exploit the intrinsic sparsity of the CDMA signal at the first stage in order to improve system performance as well as bandwidth efficiency; (ii) we develop a number of sparsity-embracing solutions to parameter estimation problems tailored for CDMA systems. In the following sections, these two stages are elaborated further.

3.1. First Stage: Estimation of Sparse Signals. Standard pilot-assisted estimation in CDMA systems provides an LS estimate of $\mathbf{g}$ [10], which amounts to solving

$$\arg \min_{\mathbf{x} \in \mathbb{C}^{KQ}} \|\mathbf{y} - \mathbf{Sx}\|_2^2. \qquad (15)$$

If $\mathbf{S}$ has full column rank, the unique solution of (15) is $\mathbf{S}^\dagger \mathbf{y}$. This approach, however, does not take into account the structure of $\mathbf{g}$ in (14). In most cases of interest $\mathbf{g}$ is sparse, meaning that the number of its nonzero entries is much smaller than its dimension. In the present context, this condition translates to $\|\mathbf{g}\|_0 \ll KQ$. With $\mathcal{K}_a$ denoting the set of active users, $\|\mathbf{g}\|_0 = \sum_{k \in \mathcal{K}_a} \|\mathbf{g}_k\|_0 = \sum_{k \in \mathcal{K}_a} L_k$, which reveals that two facts concur in the sparsity of $\mathbf{g}$: the intrinsic sparsity of the sampled signals $\mathbf{g}_k$ and the possibility of several users being idle.

3.1.1. Sparsity of $\mathbf{g}_k$. The uncertainty regarding the kth user's delay τ_k and the corresponding channel response $\psi_{\mathrm{ch},k}$ forces the receiver to consider a large number of samples of $g_k(t)$, most of which are actually zero. From (5) and (6), the condition for sparsity translates to $L_k \ll Q$. This is always guaranteed in asynchronous systems with a large τ_k uncertainty so that $\tau_{\max} \gg \Delta_{\mathrm{ch}} + \Delta_{\mathrm{tx}} + \Delta_{\mathrm{rx}}$. However, $L_k \ll Q$ holds true also if $L_k \ll L$. This condition occurs, for example, in multipath channels when there are only a few paths or, in general, when the measure of the support of $\psi_{\mathrm{ch},k}$ is much smaller than Δ_{ch}.

3.1.2. User Inactivity. Recall that when user k is idle, vector $\mathbf{g}_k$ is null. In fact, if the number of active users is on the average much smaller than K, then $\mathbf{g}$ is sparse. This can be a typical situation encountered in wireless cognitive radio networks, where secondary users sporadically transmit over licensed bands [20].

While sparsity due to user inactivity has been recently taken into account for multiuser detection [21, 22], sparsity emerging from the uncertainty of timing offsets and propagation channels has not been exploited for multiuser parameter estimation. The ordinary LS estimate adopted in [10] does not provide a sparse solution, and this may lead to disappointing performance. This is especially true if short training sequences are used. In fact, if $\overline{P}NM < KQ$, the matrix $\mathbf{S}$ becomes fat, and (15) does not admit a unique solution. In this case, $\mathbf{S}^\dagger \mathbf{y}$ does not yield consistent estimates

in the small-noise limit (i.e., as $\mathcal{N}_0 \rightarrow 0$). Motivated by these facts, an estimator that takes into account the sparsity of $\mathbf{g}$ is developed herein. Specifically, the estimate of $\mathbf{g}$ is chosen as the solution of the following convex optimization problem:

$$\hat{\mathbf{g}} := \arg \min_{\mathbf{x} \in \mathbb{C}^{KQ}} \left\{ \|\mathbf{y} - \mathbf{Sx}\|_2^2 + \lambda \|\mathbf{x}\|_1 \right\}. \qquad (16)$$

The scaled ℓ_1 norm, $\lambda \|\mathbf{x}\|_1$, is essentially a regularization term controlling the degree of sparsity in the estimated vector $\hat{\mathbf{g}}$. The convex problem in (16) can be solved via quadratic programming for which a number of efficient solvers with complexity comparable to LS are available [23–25].

The most widely applicable solver of (16) is the Lasso [11]. Lasso's original formulation deals with real variables, which for complex variables suggests concatenating the real and imaginary parts of $\mathbf{g}$. The limitation of this approach is that real and imaginary parts per entry are treated separately. In the present context however, real and imaginary parts should be treated as a "group" (in the sense of [25]) that should be selected (or not) by the Lasso. This requirement motivates employment of what is known as group-Lasso [25]. In particular, the CDMA parameter estimator in (16) is actually a group-Lasso estimator with regularization term given by $\lambda \|\mathbf{x}\|_1 = \lambda \sum_{i=1}^{KQ} \sqrt{|\Re\{(\mathbf{x})_i\}|^2 + |\Im\{(\mathbf{x})_i\}|^2}$, even if in the sequel, with a slight abuse of terminology, it will be referred to as Lasso for complex sparse vectors.

The choice of λ affects the performance of the Lasso estimator critically. Indeed, the larger the λ, the more coefficients are shrunk to zero. While asymptotically (as $P \rightarrow \infty$) optimal choices are available for the Lasso [26], for finite sample sizes the value suggested in [13] is $\lambda = \sqrt{2\mathcal{N}_0 \ln(KQ)}$. Alternatively, the problem can be solved for various λ's, and the best one can be selected via cross-validation, or using a model selection criterion (see [23, 25] for further details). As far as performance analysis of Lasso, this is nontrivial especially for finite samples and represents a topic of contemporary research. The so-termed restricted isometry property (RIP) is popular in the field of CS [12] and provides a sufficient condition for consistent estimation of sparse vectors. However, checking the RIP for the problem at hand is NP-hard, since the matrix $\mathbf{S}$ has complicated structure, and is out of the scope of this work.

Finally, it is known that Lasso generally outperforms greedy techniques, such as the matching pursuit (MP) and orthogonal matching pursuit (OMP) alternatives [27, 28]. These select the nonzero components *sequentially* and thus suffer from error propagation especially at relatively low SNR, where decisions at early stages are prone to be erroneous [13, 28, 29].

3.2. Second Stage: Parameter Estimation. The second stage improves the rough estimate at the output of the first stage and provides a decision as to the user activity. In order to maintain an affordable computational complexity, a single-user procedure is employed; namely, for each user k the refined estimate of its composite channel, call it $\hat{\mathbf{g}}_k''$, and the decision on its activity state are taken based on $\hat{\mathbf{g}}_k$ only. Any a priori information regarding the structure of $\mathbf{g}_k$ can be

included at this point, for example, the presence of a cluster of length bounded by L or any other information arising from the particular problem at hand. This stage is more application dependent and will be specifically designed for different parameter estimation problems. In any case, it is impossible to find the probability distribution of the Lasso estimator based on finite samples [30]. This is also the case of the LS estimator in (15) due to multiaccess interference [10]. For this reason, the design of the second stage will be carried by resorting to suboptimal procedures.

3.3. Chip-Synchronous CDMA over AWGN Channels.

A notable simplification of the signal model developed in Section 2 occurs in AWGN channels when chip-synchronism is ensured by higher level network protocols. The following assumptions are adopted:

(i) the user delays are integer multiples of T_c, that is, $\tau_k = d_k T_c$;

(ii) the channel impulse responses are $\psi_{\mathrm{ch},k}(t) = \delta(t)$;

(iii) $\psi_{\mathrm{tx}} \star \psi_{\mathrm{rx}}$ and $\psi_{\mathrm{rx}} \star \psi_{\mathrm{rx}}$ are Nyquist waveforms, and no oversampling is performed ($M = 1$);

(iv) the number of active users is known, and without loss of generality (w.l.o.g.), equal to K.

In this case, $\mathbf{y} = \mathbf{r}$ since there is no need for whitening; $L_k = 1$, that is, there is only one nonzero entry in each $\mathbf{g}_k$; $Q = d_{\max} + 1$ so that if $d_{\max} \gg 1$, the vector $\mathbf{g}$ is sparse and the first stage suffices to recover an estimate $\hat{\mathbf{g}}$.

The second stage, on the other hand, has to provide an estimate of the set of user delays $\{d_k\}_{k=0}^{K-1}$ starting from $\{\hat{\mathbf{g}}_k\}_{k=0}^{K-1}$. Observe that no channel estimation or user activity detection is needed here. The proposed procedure consists in selecting the candidate delay in order to maximize the energy of the composite channel, that is,

$$\hat{d}_k = \arg\max_{i \in \{0,1,\dots,d_{\max}\}} \left| (\hat{\mathbf{g}}_k)_{i+1} \right|, \quad k = 0, 1, \dots, K-1. \quad (17)$$

3.4. Asynchronous CDMA over Flat Fading Channels.

A slightly more complex scenario is considered here. The following assumptions are made:

(i) the channel impulse responses are $\psi_{\mathrm{ch},k}(t) = A_k \delta(t)$;

(ii) the transmit and receive filters are rectangular waveforms over $[0, T_c]$ and no oversampling is performed ($M = 1$);

(iii) the number of active users is known, and w.l.o.g. set equal to K.

In this case, $\mathbf{y} = \mathbf{r}$, and the $L_k = 2$ nonzero entries of $\mathbf{g}_k$ appear in two consecutive positions, namely,

$$(\mathbf{g}_k)_{d_k+1} = \frac{\sqrt{\mathcal{E}_k} A_k (1 - \delta_k)}{T_c}, \qquad (\mathbf{g}_k)_{d_k+2} = \frac{\sqrt{\mathcal{E}_k} A_k \delta_k}{T_c}, \quad (18)$$

where $\delta_k := \tau_k - d_k T_c$. Further, $Q = \lceil \tau_{\max}/T_c \rceil + 1$, so that if $\tau_{\max} \gg T_c$, the vector $\mathbf{g}$ is sparse.

The second stage has to be designed to estimate the set of user delays $\{\tau_k\}_{k=0}^{K-1}$ starting from $\{\hat{\mathbf{g}}_k\}_{k=0}^{K-1}$. This problem can be solved as in [10] to obtain

$$\hat{d}_k = \arg\max_{i \in \{0,1,\dots,d_{\max}\}} \left\{ \left| (\hat{\mathbf{g}}_k)_{i+1} \right|^2 + \left| (\hat{\mathbf{g}}_k)_{i+2} \right|^2 \right\},$$

$$\hat{\delta}_k = \frac{\left| (\hat{\mathbf{g}}_k)_{\hat{d}_k+2} \right|}{\left| (\hat{\mathbf{g}}_k)_{\hat{d}_k+1} \right| + \left| (\hat{\mathbf{g}}_k)_{\hat{d}_k+2} \right|}, \quad (19)$$

$$\hat{\tau}_k = \hat{d}_k T_c + \hat{\delta}_k.$$

3.5. Multipath Fading Channels and Known Number of Users.

An asynchronous system operating under the following assumptions is considered here:

(i) the channel impulse responses are $\psi_{\mathrm{ch},k}(t) = \sum_{w=1}^{W} A_{w,k} \delta(t - \tau_{k,w})$, with W, $A_{w,k}$, and $\tau_{w,k}$ unknown;

(ii) the number of active users is known, and w.l.o.g. set equal to K.

Given the absence of information regarding path delays, the following two cases are of interest.

3.5.1. Sparsity Due to Asynchronism.

If $\tau_{\max} \gg \Delta_{\mathrm{ch}} + \Delta_{\mathrm{tx}} + \Delta_{\mathrm{rx}}$, the vectors $\mathbf{g}_k$ are sparse and contain a cluster of at most L nonzero entries; hence, an effective procedure is

$$\hat{d}_k = \arg\max_{i \in \{0,1,\dots,d_{\max}\}} \sum_{\ell=0}^{L-1} \left| (\hat{\mathbf{g}}_k)_{i+\ell+1} \right|^2, \quad (20a)$$

$$(\hat{\mathbf{g}}_k'')_i = \begin{cases} (\hat{\mathbf{g}}_k)_i, & \text{if } i \in \left\{ \hat{d}_k + 1, \dots, \hat{d}_k + L \right\} \\ 0, & \text{otherwise,} \end{cases} \quad (20b)$$

where $\hat{\mathbf{g}}_k''$ denotes the refined estimate at the output of the second stage.

3.5.2. Sparsity due to the Channel.

If $\Delta_{\mathrm{ch}} \gg W(\Delta_{\mathrm{tx}} + \Delta_{\mathrm{rx}})$, the vectors $\mathbf{g}_k$ are sparse ($L_k \ll L$), but they have no particular structure. Therefore, the second stage cannot improve the estimate of the first stage, and $\hat{\mathbf{g}}_k$ is taken as an estimate of the composite channel.

3.6. Dispersive Channels and Unknown Number of Users.

No assumption is made here and the system considered is the general one of Section 2. Given the estimates $\{\hat{\mathbf{g}}_k\}_{k=0}^{K-1}$, user k is declared to be active based on $\hat{\mathbf{g}}_k$, in which case $\hat{\mathbf{g}}_k$ is taken as an estimate of the composite channel. The second stage performs binary hypothesis testing: user idle versus active. Given the non-Bayesian setting, a simple yet effective procedure is the following energy detector:

$$\|\hat{\mathbf{g}}_k\|_2^2 \begin{cases} > \gamma, & \text{declare user } k \text{ active with channel } \hat{\mathbf{g}}_k, \\ < \gamma, & \text{declare user } k \text{ idle,} \end{cases}$$

$$(21)$$

where the threshold γ is set to meet the required specifications in terms of false alarm probability (i.e., the probability of declaring a user to be active when it is not). As in the previous section, if sparsity comes mostly from user asynchronism, the second stage runs the energy detector starting from the refined estimates $\hat{\mathbf{g}}_k''$ of the procedure in (20b).

4. Simulated Tests

The performance of the proposed estimation strategies is assessed in this section through Monte Carlo simulations. A CDMA system employing random (long) spreading codes with rectangular chip-waveforms at both transmitter and receiver ($\Delta_{tx} = \Delta_{rx} = T_c$), $M = 1$, and $N = 15$ is simulated using binary (randomly generated) training sequences. The cases described in Sections 3.3 to 3.6 are considered next.

Test Case 1. The system model in Section 3.3, with $d_{max} = 14$, and $K = 5$, is considered first. Figure 1 shows the probability of incorrect delay estimation for the *kth* user, $P_e := \mathbb{P}(\{\hat{d}_k \neq d_k\})$, versus the signal-to-noise ratio SNR $:= \mathcal{E}_k/\mathcal{N}_0$ (expressed in dB). Perfect power control is considered among users (i.e., the interference-to-signal ratio, ISR $:= \mathcal{E}_j/\mathcal{E}_k$, is equal to 0 dB, for all $j \neq k$), and different values of preamble lengths P are tested. The penalty parameter of (16) has been set as $\lambda = \sqrt{2\mathcal{N}_0 \ln(KQ)}$. The performance of the proposed Lasso-based algorithm is compared with two classical strategies: the LS-based algorithm (i.e., a two-stage procedure employing the LS at the first stage [10]), and the conventional sliding correlator [31], that is,

$$\hat{d}_k = \arg\max_{i \in \{0, \ldots, d_{max}\}} \left| \mathbf{s}_{k+i+1}^H \mathbf{y} \right|, \quad k = 0, \ldots, K-1 \quad (22)$$

with $\mathbf{s}_j$ denoting the *j*th column of $\mathbf{S}$. From Figure 1 it can be seen that both procedures are largely outperformed by the Lasso-based algorithm. In Figure 2 a severe near-far scenario is tested, at ISR = 10 dB. As expected, the two-stage algorithms are near-far resistant, while the conventional sliding correlator, which is a single-user algorithm, is ineffective. Moreover, the LS-based algorithm is outperformed by the Lasso one, which exhibits a gain of two orders of magnitudes at 10 dB for $P = 5$.

Test Case 2. The system considered here is that of Section 3.4 with $\tau_{max} = T_s$ and $K = 5$, and the procedures are compared in terms of the mean-square error of the estimated user delay, MSE $:= \mathbb{E}[(\tau_k - \hat{\tau}_k)^2]$. Figure 3 shows the MSE versus SNR for various P values at ISR = 0 dB. Notice that the Lasso-based algorithm outperforms the LS-based one exhibiting gain as high as 10 dB at MSE = 10^{-1} and $P = 9$. The same trend is witnessed in Figure 4, where a near-far scenario with ISR = 10 dB is tested.

Test Case 3. The system of Section 3.5 is considered here. The channels have uniform power delay profile with $W = 3$ and complex, circularly symmetric, Gaussian amplitudes,
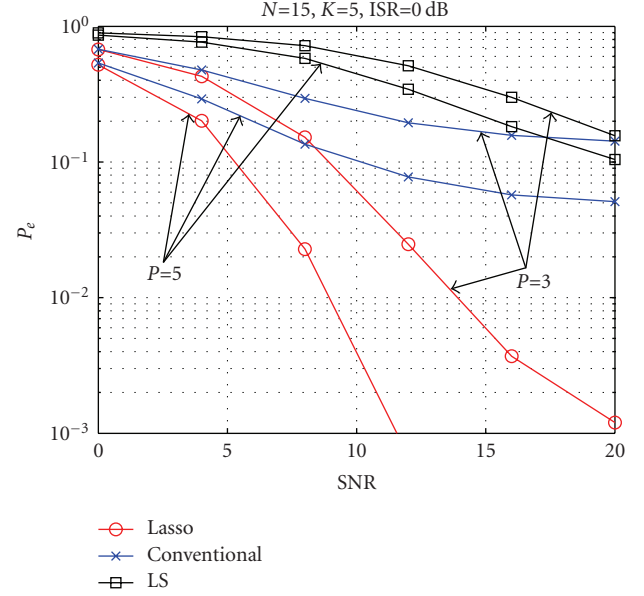

FIGURE 1: Probability of incorrect delay estimation versus SNR in a power-controlled scenario (chip-synchronous system with AWGN and known K).
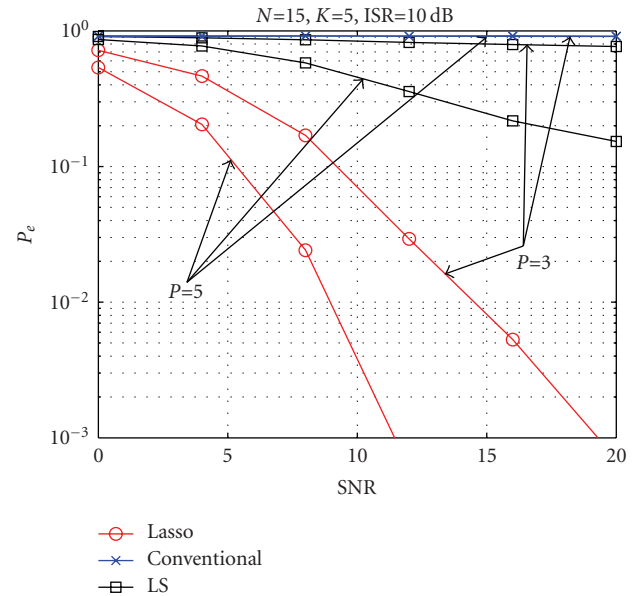

FIGURE 2: Probability of incorrect delay estimation versus SNR in a near-far scenario (chip-synchronous system with AWGN and known K).

while channel path delays are uniformly distributed in the interval $[0, 10T_c]$. The user delays, on the other hand, have been randomly picked from the interval $[0, 5T_c]$, which amounts to assuming that coarse synchronization has been achieved at the receiver side [10]. Since sparsity comes mostly from the channel structure, the second stage has not been implemented. The performance measure adopted is the normalized mean-square error for the composite channel, NMSE $:= \mathbb{E}[\|\mathbf{g} - \hat{\mathbf{g}}\|_2^2]/\mathbb{E}[\|\mathbf{g}\|_2^2]$. Figures 5 and 6 depict the
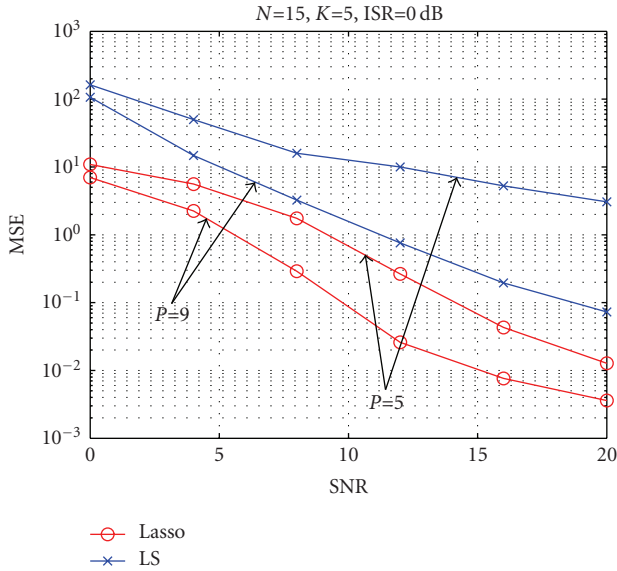
FIGURE 3: Mean-square error of the users' time delay versus SNR in a power-controlled scenario (chip-synchronous system with AWGN and known K).
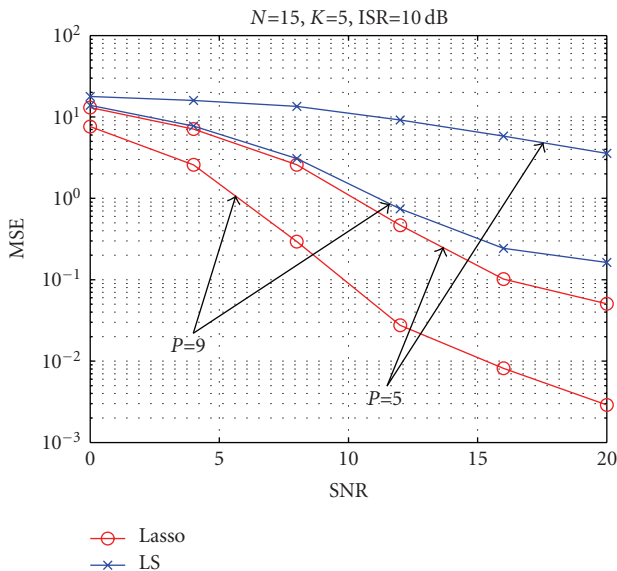

FIGURE 4: Mean-square error of the users' time delay versus SNR in a near-far scenario (chip-synchronous system with AWGN and known K).

NMSE versus the SNR for different training packet lengths P and $K = 5$. Lasso, LS, and OMP solvers are simulated: the Lasso parameter λ is tuned as described in Section 3.1, while the OMP algorithm stops as soon as the number of nonzero entries of its estimate equals that of Lasso. It can be seen that for $P = 10$, the LS estimator suffers a loss of about 3 dB at NMSE $= 10^{-4}$ with respect to Lasso, and 2 dB with respect to OMP. In the considered setting, $Q = 16$ and $\overline{m} = 1$, so that the matrix $\mathbf{S}$ in (14) is tall ($\overline{P}NM = 165$ and $KQ = 80$), and $\mathbf{S}^H\mathbf{S}$ is full-rank with overwhelming probability. In this case, LS is consistent in the small noise
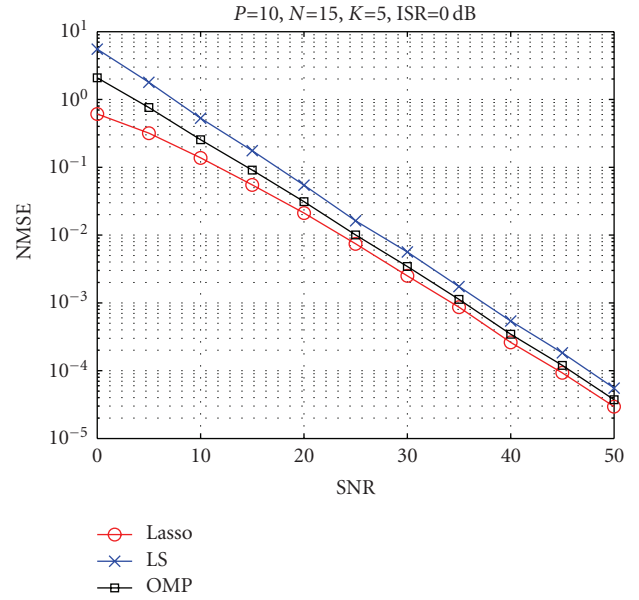

FIGURE 5: Normalized mean-square error of the composite channel versus SNR for overdeterminate case (asynchronous system with multipath fading channel and known K).

limit; nevertheless, both LS and OMP are outperformed by the Lasso-based estimator. Figure 6 considers a scenario wherein short training sequences are adopted, that is, $P = 4$. This is an underdetermined case and $\mathbf{S}$ is fat ($\overline{P}NM = 75$, so that $\overline{P}NM < KQ$). As expected, LS exhibits a floor, and it is not consistent in the small-noise limit, while Lasso and OMP are. Moreover, Lasso outperforms OMP by about 5 dB at NMSE $= 2 \times 10^{-3}$. This shows that the proposed sparsity-aware procedure can considerably reduce the number of pilot symbols without affecting the system performance; hence, throughput can be increased without error penalty.

Test Case 4. The system is that of Section 3.6 with the multipath fading channels as in Test Case 3. The number of active users, whose identity is not known, is $S = 5$, while the total number of candidate users is $K = 10$. Figure 7 depicts the receiver operating characteristic (ROC), that is, the probability of detection, $P_D := \mathbb{P}$ ("user k is declared active"|"user k is active"), versus the probability of false alarm, $P_{\text{FA}} := \mathbb{P}$ ("user k is declared active"|"user k is idle"). A training sequence of length $P = 10$ is used. It can be seen that the Lasso-based procedure largely outperforms both the OMP and the LS-based one, which turn out to be ineffective with such a short training packet. In Figure 8, the probability of missed detection $P_{\text{MD}} := 1 - P_D$ versus SNR is reported for $P_{\text{FA}} = 10^{-2}$. Again, the Lasso-based algorithm exhibits a large gain with respect to LS and OMP. Finally, Figure 9 shows the performance degradation of joint user activity detection and channel estimation (DE) with respect to the problem of channel estimation estimation only (E), case in which the receiver has knowledge of the number and identities of the users. The NMSE plotted versus the SNR
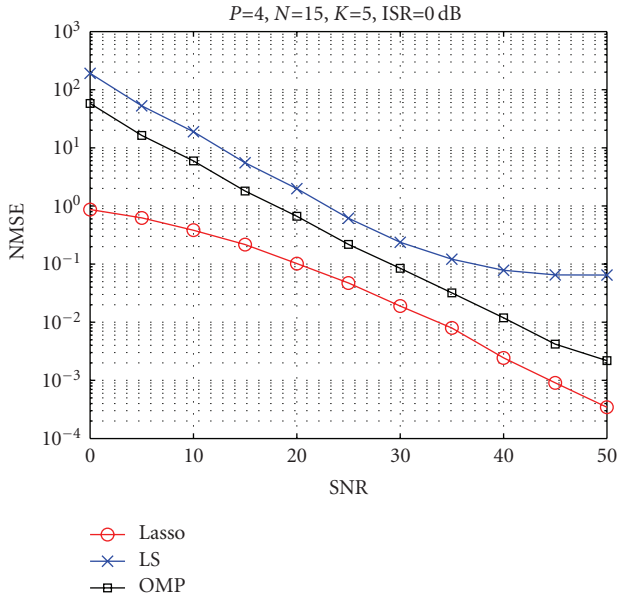
FIGURE 6: Normalized mean-square error of the composite channel versus SNR for underdeterminate case (asynchronous system with multipath fading channel and known K).

FIGURE 8: Probability of missed detection versus SNR (asynchronous system with multipath fading channel and unknown K).
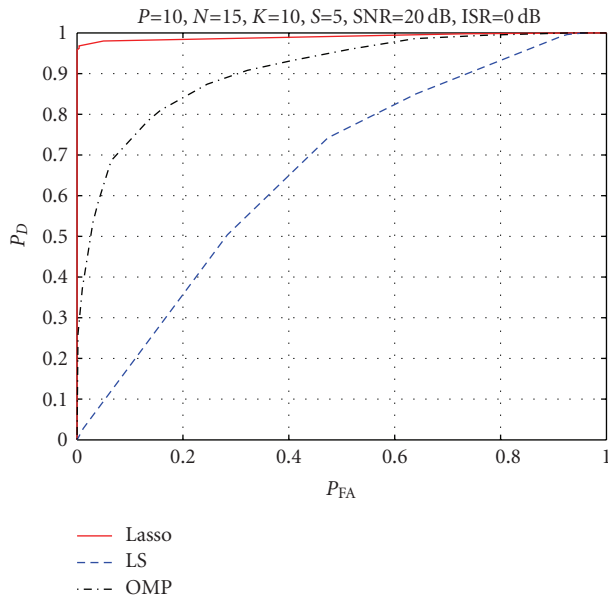

FIGURE 7: Receiver operating characteristics in a power-controlled scenario with training sequence length $P = 10$ (asynchronous system with multipath fading channel and unknown K).

FIGURE 9: Normalized mean-square error of the composite channel versus SNR (asynchronous system with multipath fading channel and unknown K).

confirms that the developed algorithm incurs minimal loss in performance even when the number and identity of active users are unknown.

5. Conclusions

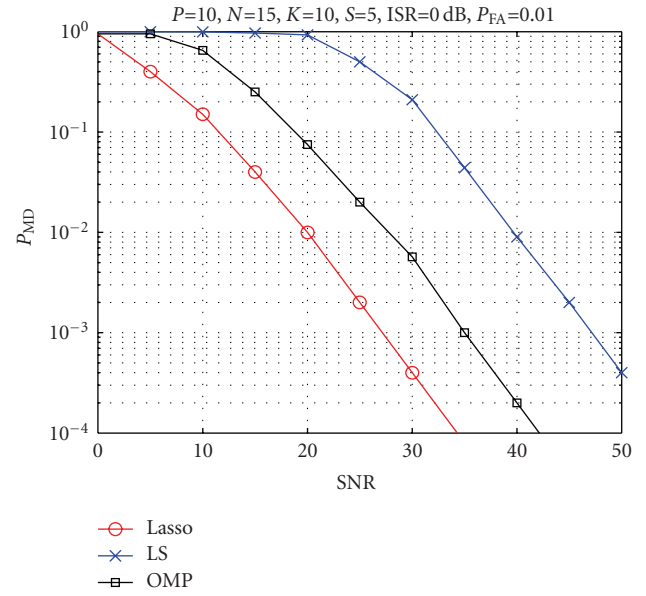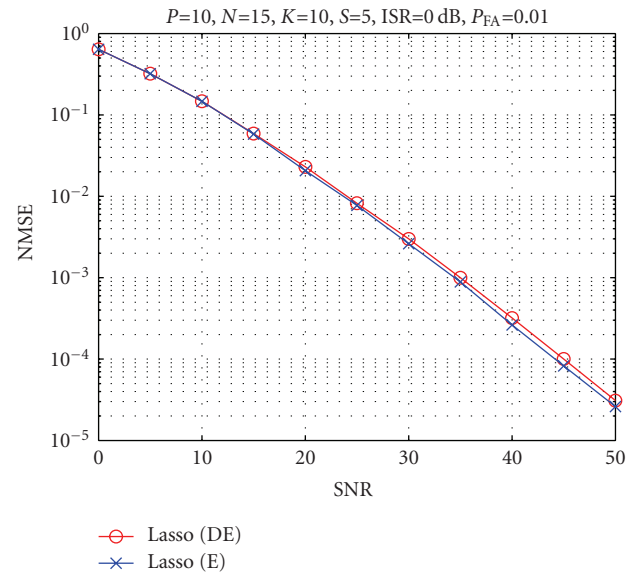The recently popular approaches to estimating sparse vectors in linear regression problems have been exploited in this paper for estimating CDMA system parameters under very general channel conditions. It has been established that the parameter vectors of interest are often sparse, and simple algorithms for parameter estimation and user activity detection based on the Lasso algorithm have been developed. Sparsity comes from the large uncertainty associated with propagation parameters, which prompts the receiver to consider many possible candidates. It turns out that sparsity-aware algorithms largely outperform their sparsity-agnostic

counterparts and guarantee consistent estimates in the small-noise limit for short pilot sequences where LS fails. From this vantage point, exploiting sparsity reduces the number of training symbols required for CDMA parameter acquisition. Interestingly, the developed algorithms are computationally efficient and remain applicable even when the user activity levels are unknown, a scenario of potential interest in wireless cognitive radio networks.

Acknowledgments

References

[1] E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Transactions on Communications*, vol. 44, no. 1, pp. 84–93, 1996.

[2] S. E. Bensley and B. Aazhang, "Maximum-likelihood synchronization of a single user for code-division multiple-access communication systems," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 392–399, 1998.

[3] D. Zheng, J. Li, S. L. Miller, and E. G. Strom, "An efficient code-timing estimator for DS-CDMA signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 82–89, 1997.

[4] U. Madhow, "Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 124–136, 1997.

[5] T. K. Moon, Z. Xie, C. K. Rushforth, and R. T. Short, "Parameter estimation in a multi-user communication system," *IEEE Transactions on Communications*, vol. 42, no. 8, pp. 2553–2560, 1994.

[6] Z. Xie, C. K. Rushforth, R. T. Short, and T. K. Moon, "Joint signal detection and parameter estimation in multiuser communications," *IEEE Transactions on Communications*, vol. 41, no. 8, pp. 1208–1216, 1993.

[7] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems," *IEEE Communications Magazine*, vol. 36, no. 9, pp. 56–69, 1998.

[8] E. H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," *IEEE Communications Magazine*, vol. 36, no. 9, pp. 48–54, 1998.

[9] L. Tong, A.-J. van der Veen, P. Dewilde, and Y. Sung, "Blind decorrelating RAKE receivers for long-code WCDMA," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1642–1655, 2003.

[10] S. Buzzi and H. V. Poor, "On parameter estimation in long-code DS/CDMA systems: Cramèr-Rao bounds and least-squares algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 545–559, 2003.

[11] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[12] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[13] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal of Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[14] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 374–377, 2002.

[15] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: channel estimation," *IEEE Journal on Selected Topics in Signal Processing*, vol. 1, no. 3, pp. 383–395, 2007.

[16] W. U. Bajwa, A. M. Sayeed, and R. Nowak, "Learning sparse doubly-selective channels," in *Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing*, pp. 575–582, Monticello, Ill, USA, September 2008.

[17] G. Taubock and F. Hlawatsch, "A compressed sensing technique for OFDM channel estimation in mobile environments: exploiting channel sparsity for reducing pilots," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 2885–2888, Las Vegas, Nev, USA, April 2008.

[18] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1847–1862, 2010.

[19] S. Kim and R. A. Iltis, "A matching-pursuit/GSIC-based algorithm for DS-CDMA sparse-channel estimation," *IEEE Signal Processing Letters*, vol. 11, no. 1, pp. 12–15, 2004.

[20] J. Mitola, "The software radio architecture," *IEEE Communications Magazine*, vol. 33, no. 5, pp. 26–38, 1995.

[21] H. Zhu and G. B. Giannakis, "Sparsity-embracing multiuser detection for CDMA systems with low activity factor," in *Proceedings of the IEEE International Symposium on Information Theory*, Seoul, Korea, June 2009.

[22] A. K. Fletcher, S. Rangan, and V. K. Goyal, "A sparsity detection framework for on-off random access channels," in *Proceedings of the IEEE International Symposium on Information Theory*, Seoul, Korea, June 2009.

[23] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[24] D. L. Donoho and Y. Tsaig, "Fast solution of l_1-norm minimization problems when the solution may be sparse," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789–4812, 2008.

[25] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society Series B*, vol. 68, no. 1, pp. 49–67, 2006.

[26] K. Knight and W. Fu, "Asymptotics for Lasso-type estimators," *Annals of Statistics*, vol. 28, no. 5, pp. 1356–1378, 2000.

[27] S. Mallat and Z. Zhang, "Matching pursuit in a time-frequency dictionary," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 4655–4666, 2007.

[28] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[29] C. R. Berger, S. Zhou, J. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: from subspace methods to compressed sensing," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1708–1721, 2010.

[30] B. M. Potscher and H. Leeb, "On the distribution of penalized maximum likelihood estimators: the LASSO, SCAD, and thresholding," *Journal of Multivariate Analysis*, vol. 100, no. 9, pp. 2065–2082, 2009.

[31] S. S. Rappaport and D. M. Grieco, "Spread-spectrum signal acquisition: methods and technology," *IEEE Communications Magazine*, vol. 22, no. 6, pp. 6–21, 1984.