# Power Control and Spreading Sequence Allocation in a CDMA Forward Link

Li Gao and Tan F. Wong, *Senior Member, IEEE*

***Abstract*—In this paper, we consider power control and sequence allocation to meet signal-to-interference ratio (SIR) targets for users in a code-division multiple-access (CDMA) forward link. Particular attention is given to the case when the number of users is larger than the spreading gain. Users in the system are classified into two classes, namely, overfaded users and nonoverfaded users, according to their effective noise densities and SIR targets. Overfaded users are allocated orthogonal *channels*, and nonoverfaded users share the remaining *channels*. The spreading sequences allocated belong to the class of sequences which minimize the extended total squared correlation (ETSC). Power efficiency of this allocation scheme is examined through comparison to Lagrangian-based searching results. The proposed allocation scheme is shown to be effective under most channel conditions. Two adaptive algorithms are presented to construct the spreading sequences iteratively. One assumes fixed power allocation and adapts the sequences only, and the other adapts both transmission powers and sequences simultaneously. Convergence of the two adaptive algorithms and the performance of joint transmitter–receiver adaptation with imperfect real-time channel parameters estimation are examined via computer simulations.**

***Index Terms*—Code-division multiple access (CDMA), forward link, minimum mean-squared-error (MMSE) receiver, optimal sequence, power control, total squared correlation (TSC), Welch's bound.**

## I. Introduction

In direct-sequence code-division multiple-access (DS-CDMA) systems, the information-bearing signal for each user is spread over a wide bandwidth by means of a spreading waveform unique to that user. Multiple-access interference (MAI), which refers to the interference between direct-sequence users due to many simultaneous users in the same frequency band, is a factor that limits the capacity and performance of DS-CDMA systems. Two general approaches have been extensively studied to improve the performance of CDMA systems: multiuser detection and power control. After Verdú [1] proposed the optimum multiuser receiver, much research has focused on devising suboptimal multiuser receivers [2]–[7], which are more feasible to implement in practice. Performance of several important linear receivers, including the matched filter, minimum mean-squared error (MMSE), and decorrelator receivers, is analyzed in a reverse-link single-cell "large" system with random sequences [8], [9]. Not surprisingly, the MMSE receiver is the most efficient one in terms of maximizing the *user capacity*.[1]

Transmitter power control is currently considered indispensable for successful signal transmission/reception in the CDMA reverse link. The objective of the reverse-link power control is to deal with the *near–far* problem, i.e., to ensure that signals from all users arrive at the receiver with about the same power, and therefore no user is substantially disadvantaged relative to other users. A number of power control schemes have been devised [10]–[12] to achieve this objective. While adaptive power control is useful to guarantee a certain signal-to-interference ratio (SIR) at the receiver, it does not exploit the full potential of transmitter adaptation. It is conceivable that joint power control and spreading sequence allocation offers a performance advantage over power control alone. It is argued [13] that the SIR achieved by the MMSE receiver depends on the choice of the spreading sequences. If the spreading sequences are chosen or adapted suitably together with adaptive power control, the MAI level as seen by the MMSE receiver can be further suppressed and hence the performance of the system is improved.

The optimization problem of power control and spreading sequence allocation in conjunction with MMSE signal reception so that the SIR targets of all users are met with the minimum total received power in a single-cell reverse-link CDMA system is solved in [14]. A distributed algorithm is also suggested [15] to obtain the optimal spreading sequences iteratively for the case when a uniform SIR target is desired. The convergence behavior of this algorithm is addressed in [16]. Variants of this optimization problem in a multicarrier setting [17] and a multicode setting [18] are also solved. User capacity of a single-cell symbol-asynchronous CDMA system with a matched filter receiver is addressed in [19]. It is claimed that the user capacity of a single-cell symbol-asynchronous CDMA system is the same as that of a single-cell synchronous system when the users' spreading sequences belong to the class of sequences that minimizes the total squared correlation (TSC). User capacity and admissibility of imperfect power-controlled CDMA systems with linear receivers in fading environment, assuming known received power distributions, are studied in [20].

Most of the cited work concern either the additive white Gaussian noise (AWGN) channel or the reverse link. Results about joint power control and spreading sequence allocation

Manuscript received September 7, 2001; revised July 16, 2003. This work was supported in part by the Office of Naval Research under Grant N000140210554. The material in this paper was presented in part at the IEEE Global Telecommunications Conference, San Antonio, TX, November 2001.

L. Gao is with Silicon Laboratories, Austin, TX 78735 USA (e-mail li.gao@silabs.com).

T. F. Wong is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: twong@ece.ufl.edu).

Communicated by V. V. Veeravalli, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2003.821990

[1] *User capacity* refers to the maximum number of users that can be supported with the desired quality of service (QoS) requirement.

in the forward link are not abundant. Similar considerations in the context of CDMA systems with multiple antennas and space-division multiple-access systems have been addressed in [21] and [22], respectively. In traditional DS-CDMA systems, power control in the forward link is not essential because the transmitted signals are time synchronized and orthogonal spreading sequences are generally employed to avoid MAI. The number of orthogonal sequences is limited by the value of spreading gain, thus creating a bottleneck that limits the number of admissible users. One approach to avoid this capacity restriction is to employ additional sequences which are not orthogonal to each other. Introduction of nonorthogonal spreading sequences causes MAI, and makes power control and sequence allocation essential to guarantee the QoS of the system.

In this paper, we consider power control and spreading sequence allocation to satisfy the SIR targets of users in a single-cell forward-link CDMA system. Each user's signal is demodulated separately and simultaneously using an MMSE receiver. We classify the users into two classes, namely, *overfaded users* and *nonoverfaded users*, according to their effective noise densities and SIR targets. Overfaded users are allocated orthogonal *channels* and nonoverfaded users share the remaining *channels*. To minimize cochannel interference to other cells, we seek the minimum total transmission power to support all users numerically using a Lagrangian-based searching method. The results obtained from the search are employed to examine the power efficiency of the proposed allocation. The Lagrangian searching results indicate that the proposed allocation scheme is effective under most channel conditions.

One advantage of the proposed allocation scheme is that spreading sequences can be easily constructed using algorithms that are simple and amenable to adaptive implementation. Two adaptive algorithms are proposed. One assumes fixed power allocation and updates the spreading sequences with the MMSE receiver weight vectors, while the other updates both the transmission powers and spreading sequences simultaneously. An adaptive CDMA system including joint channel parameter estimation, transmitter power and sequence adaptation, and adaptive MMSE signal reception is also presented to facilitate the implementation of this allocation scheme in real-life communication systems. Different from the approaches in [23] and [24], which update the spreading sequences at the receivers and send the new spreading sequences back to the base station, the approach we present here updates the spreading sequences at the base station. This reduces the bandwidth requirement in the feedback channel by requiring only a minimum amount of information to be fed back from the mobile receivers, thus making the adaptive structure more feasible for practical systems.

The rest of the paper is organized as follows. The signal and system model is defined in Section II. In Section III, we present the power control and spreading sequence allocation scheme that supports all users with their target SIRs. In Section IV, performance of this allocation scheme is evaluated in terms of its power efficiency. In Section V, adaptive algorithms are introduced to construct the spreading sequences iteratively. Practical issues of joint transmitter–receiver adaptation with channel parameter estimation are also investigated. Numerical examples are presented in Section VI to examine the convergence of the adaptive algorithms as well as the performance of the proposed joint adaptation structure. Conclusions are drawn in Section VII.

## II. System Model for CDMA Forward Link

We consider a forward-link DS-CDMA system. We assume that there are $K$ users in the system, and each user generates one stream of data symbols. The data stream of the $k$th user, for $1 \leq k \leq K$, is given by $(\ldots, b_0^{(k)}, b_1^{(k)}, b_2^{(k)}, \ldots)$. We assume that the data symbols $\{b_j^{(k)}\}$ are independent random variables with zero mean and unit variance. For binary communication, $b_j^{(k)} \in \{+1, -1\}$. The data stream of the $k$th user is spread with the short periodic spreading sequence

$$\boldsymbol{a}_k = [a_0^{(k)}, a_1^{(k)}, \ldots, a_{N-1}^{(k)}]^T$$

and then modulated to the carrier frequency $\omega_c$ to give the transmitted signal

$$s_k(t) = \mathbf{Re}\left[\sum_{i=-\infty}^{\infty} b_{\lfloor i/N \rfloor}^{(k)} a_i^{(k)} \psi(t - iT_c) e^{\mathrm{j}\omega_c t}\right]$$

where $T_c$ is the chip interval, and $\psi(t)$ is the chip waveform. We assume that $\psi(t)$ satisfies the Nyquist criterion for zero interchip interference, and $\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$.

We consider multiuser forward-link transmission and assume that the signals from the base station to the receivers of different users undergo independent slow flat fading. Data streams of users are simultaneously demodulated at the corresponding mobile receivers using the adaptive MMSE receiver structure proposed in [4]–[6]. Without loss of generality, we consider the detection of the first user's data stream. The complex baseband representation of the received signal at the first user's receiver is given by

$$r_1(t) = c_1 e^{\mathrm{j}\theta_1} \sum_{k=1}^{K} \sum_{i=-\infty}^{\infty} b_{\lfloor i/N \rfloor}^{(k)} a_i^{(k)} \psi(t - iT_c - \Delta_1) + n_1(t) \tag{1}$$

where $c_1$ and $\Delta_1$ denote the channel gain and the transmission delay from the base station to the first user's receiver, respectively, $\theta_1$ accounts for the overall phase shift of the signals, and $n_1(t)$ represents the AWGN with power spectral density (PSD) $\eta_1$ at the first user's receiver. With carrier and timing synchronization achieved, the phase shift $\theta_1$ and the transmission delay $\Delta_1$ can be set to zero. The received signal is passed through a chip-matched filter and the filter output is sampled every chip interval. To detect the zeroth symbol of the first user, we arrange the $N$ samples observed in the interval $[0, T)$ into an $N$-dimensional column vector $\boldsymbol{z}_1$, which can be expressed as

$$\boldsymbol{z}_1 = c_1 b_0^{(1)} \boldsymbol{a}_1 + \underbrace{c_1 \sum_{k=2}^{K} b_0^{(k)} \boldsymbol{a}_k + \boldsymbol{n}_1}_{\tilde{\boldsymbol{n}}_1}$$

where $c_1 b_0^{(k)} \boldsymbol{a}_k$ is the component due to the $k$th user's signal for $1 \leq k \leq K$, and $\boldsymbol{n}_1$ denotes the contribution from the AWGN.

The vector $\boldsymbol{z}_1$ is fed into a linear finite impulse response (FIR) filter and the receiver filter weights are chosen to minimize the mean-squared error (MSE) of the FIR filter

$$\boldsymbol{w}_1 = \frac{1}{c_1}\tilde{\boldsymbol{R}}_{T_1}^{-1}\boldsymbol{a}_1 \tag{2}$$

where $\tilde{\boldsymbol{R}}_{T_1}$ is the effective total correlation matrix observed by the first user given by

$$\tilde{\boldsymbol{R}}_{T_1} = \frac{1}{c_1^2}\mathrm{E}[\boldsymbol{z}_1\boldsymbol{z}_1^H] = \boldsymbol{A}_T\boldsymbol{A}_T^H + \tilde{\eta}_1\boldsymbol{I}.$$

In above, $\boldsymbol{A}_T = [\boldsymbol{a}_1\,\boldsymbol{a}_2\,\ldots\boldsymbol{a}_K]$ is an $N \times K$ matrix formed by grouping all spreading sequence vectors and $\tilde{\eta}_1 = \frac{\eta_1}{c_1^2}$ is the effective noise density at the first user's receiver. The decision statistic $Z_1 = \boldsymbol{w}_1^H\boldsymbol{z}_1$ is hard-limited to obtain the estimate of the zeroth symbol of the first user. A simple analysis shows that the optimal MMSE receiver weight vector also maximizes the receiver output SIR [18], which is given by

$$\mathrm{SIR}_1 = \boldsymbol{a}_1^H\tilde{\boldsymbol{R}}_1^{-1}\boldsymbol{a}_1 \tag{3}$$

where $\tilde{\boldsymbol{R}}_1$ is the effective noise-plus-interference correlation matrix defined by

$$\tilde{\boldsymbol{R}}_1 = \frac{1}{c_1^2}\mathrm{E}[\tilde{\boldsymbol{n}}_1\tilde{\boldsymbol{n}}_1^H] = \tilde{\boldsymbol{R}}_{T_1} - \boldsymbol{a}_1\boldsymbol{a}_1^H.$$

For ease of analysis, we derive another useful expression of the SIR using the matrix inversion formula:

$$\frac{\mathrm{SIR}_1}{1+\mathrm{SIR}_1} = \boldsymbol{a}_1^H\tilde{\boldsymbol{R}}_{T_1}^{-1}\boldsymbol{a}_1.$$

The preceding discussion applies to any user by simply replacing the subscript index 1 with the corresponding subscript index of that user.

Equation (3) indicates that the SIR achieved by the MMSE receiver depends on the choice of the spreading sequences as well as the transmission powers of the users. In the forward link, when the system resource (characterized in terms of the bandwidth and transmission power) is limited, a good power and sequence allocation scheme is vital to satisfy the QoS requirements of all users.

## III. Power Control and Spreading Sequence Allocation

In this section, we propose a transmission power and spreading sequence allocation scheme to support all users with their target SIRs in the forward-link CDMA system described in Section II.

### A. User Capacity

We assume that the $k$th user is to be supported with the target SIR $\gamma_k$, for $1 \le k \le K$. This implies

$$\boldsymbol{a}_k^H\tilde{\boldsymbol{R}}_k^{-1}\boldsymbol{a}_k \ge \gamma_k$$

or, equivalently

$$\boldsymbol{a}_k^H\tilde{\boldsymbol{R}}_{T_k}^{-1}\boldsymbol{a}_k \ge \frac{\gamma_k}{1+\gamma_k}.$$

Due to limited system capacity, there may not be a feasible allocation if there are too many users in the system or if the SIR targets of the users are too high. A necessary and sufficient condition is derived in [14] to characterizes the user capacity in the forward-link CDMA system. The maximum number of users $K$ that the system can support with spreading gain $N$ and SIR targets $\gamma_1, \gamma_2, \ldots, \gamma_K$ is limited by

$$\sum_{k=1}^{K} e(\gamma_k) < N. \tag{4}$$

Here, we follow the notation in [8] to write $e(\gamma_k) = \frac{\gamma_k}{1+\gamma_k}$. The quantity $e(\gamma_k)$ represents the *effective bandwidth* [8] that the $k$th user takes when achieving the MMSE receiver output SIR target $\gamma_k$. There is a simple interpretation for this constraint. All $K$ users are admissible if and only if the sum of their effective bandwidth is less than the spreading gain of the system.

### B. Power and Spreading Sequence Allocation

For the case of $K \le N$, the constraint in (4) is always satisfied due to the fact that $0 < e(\gamma_k) < 1$ for any positive $\gamma_k$. In this case, we can allocate orthogonal sequences to the users. With this allocation, each user does not suffer interference from the other users. The transmission power of the $k$th user, for $1 \le k \le K$, can be set as

$$p_k = \|\boldsymbol{a}_k\|^2 = \gamma_k\tilde{\eta}_k \tag{5}$$

to satisfy the SIR target $\gamma_k$. The total transmission power is $P_T = \sum_{k=1}^{K}\gamma_k\tilde{\eta}_k$.

Our main interest lies in the case of $K > N$. A simple observation of the structure of effective total correlation matrices

$$\tilde{\boldsymbol{R}}_{T_k} = \boldsymbol{A}_T\boldsymbol{A}_T^H + \tilde{\eta}_k\boldsymbol{I}, \qquad \text{for } k = 1, 2, \ldots, K$$

reveals that they share the same contribution, namely, $\boldsymbol{A}_T\boldsymbol{A}_T^H$, from the spreading sequences. If we interpret the space spanned by the eigenvectors of the correlation matrix of a user as the *channel space* seen by this user, then all the users share the same channel space as the effective correlation matrices have the same eigenvectors. The use of this channel space by the users, as specified by a proper power and sequence allocation scheme, should be determined by the effective noise densities and target SIRs, which may vary widely from user to user.

Following the idea of user classification in [14], we classify the users into two different classes, and employ different power and sequence allocation strategies to these two classes of users. Intuitively, it is reasonable to allocate orthogonal *channels* to users whose channel conditions are much worse than those of the other users and whose target SIRs are much higher than those of the other users, and let the other users share the rest of the channel space. In this way, the performance of the whole system would not be ruined by those users with poor channel conditions and high SIR targets. Based on this observation, we present in what follows a power control and spreading sequence allocation scheme to support all users with their target SIRs in the forward link of the CDMA system modeled by (1) when the user capacity condition in (4) is satisfied.

First, we need the following result to formalize the classification of users.

*Proposition 1:* Suppose that $K > N$ and the user capacity condition in (4) is satisfied. Without loss of generality, assume a descending order of the products of effective noise densities and target SIRs, i.e., $\gamma_1\tilde{\eta}_1 \geq \gamma_2\tilde{\eta}_2 \geq \cdots \geq \gamma_K\tilde{\eta}_K$. Define $\gamma_0\tilde{\eta}_0 = \infty$. Let

$$K^* = \max\left\{0 \leq k < N : k < N - \sum_{l=k+1}^{K} e(\gamma_l)\right\}.$$

There exists a unique number $k^* \in \{0, 1, \ldots, K^*\}$ such that

$$\gamma_{k^*+1}\tilde{\eta}_{k^*+1} \leq \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)} < \gamma_{k^*}\tilde{\eta}_{k^*}. \quad (6)$$

*Proof:* See Appendix I. □

We denote the first $k^*$ users, i.e., users $1, 2, \ldots, k^*$ as *overfaded users*, and users $k^*+1, k^*+2, \ldots, K$ as *nonoverfaded users*. We note that there can be at most $N-1$ overfaded users. Also when $k^* = 0$, it means that there is no overfaded users.

Now we introduce a power and sequence allocation scheme based on this classification of users. Our goal is to allocate the sequence set $\boldsymbol{A}_T$ to the users with transmission powers

$$p_k = \begin{cases} \gamma_k\tilde{\eta}_k, & \text{for } k = 1, 2, \ldots, k^* \\ e(\gamma_k)\left[\dfrac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)} + \tilde{\eta}_k\right], & \text{for } k = k^*+1, k^*+2, \ldots, K \end{cases} \quad (7)$$

such that the matrix $\boldsymbol{A}_T\boldsymbol{A}_T^H$ has eigenvalues

$$\lambda_n^* = \begin{cases} \gamma_n\tilde{\eta}_n, & \text{for } n = 1, 2, \ldots, k^* \\ \dfrac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}, & \text{for } n = k^*+1, k^*+2, \ldots, N. \end{cases} \quad (8)$$

Before describing the reason behind such an allocation, we need to show that such a construction is possible. One such sequence set is described below. First, we allocate $k^*$ orthogonal channels to the $k^*$ overfaded users in order to satisfy their SIR requirements. This can be done as described below. Generate an arbitrary $N \times N$ unitary matrix $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N]$ with column vectors $\boldsymbol{u}_n$ for $n = 1, 2, \ldots, N$. The spreading sequence of the $k$th user, for $1 \leq k \leq k^*$, is assigned as

$$\boldsymbol{a}_k = \sqrt{p_k}\boldsymbol{u}_k$$

where $p_k$ is the transmission power in (7). The remaining $K - k^*$ nonoverfaded users share the remaining $N - k^*$ channels, i.e., the $(N - k^*)$-dimensional subspace spanned by $\boldsymbol{u}_{k^*+1}, \boldsymbol{u}_{k^*+2}, \ldots, \boldsymbol{u}_N$. Proposition 2 stated below justifies the existence of such a set of spreading sequences for the nonoverfaded users, satisfying the power allocation scheme described by (7).

*Proposition 2:* Under the same assumptions in Proposition 1, we can construct an $(N - k^*) \times (K - k^*)$ real matrix

$$\check{\boldsymbol{A}} = [\check{\boldsymbol{a}}_{k^*+1}, \check{\boldsymbol{a}}_{k^*+2}, \ldots, \check{\boldsymbol{a}}_K]$$

that is characterized by [2]

$$\check{\boldsymbol{A}}\check{\boldsymbol{A}}^T = \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}\boldsymbol{I}$$
$$\mathrm{diag}(\check{\boldsymbol{A}}^T\check{\boldsymbol{A}}) = [p_{k^*+1}, \ldots, p_K]^T \quad (9)$$

where $p_{k^*+1}, p_{k^*+2}, \ldots, p_K$ are defined in (7).

*Proof:* See Appendix II. □

We note that $\check{\boldsymbol{A}}$ can be constructed recursively using the algorithm given in [17]. The spreading sequence $\boldsymbol{a}_k$ of the $k$th user, for $k = k^*+1, k^*+2, \ldots, K$, can be constructed by padding $k^*$ zeros at the beginning of the column vector $\check{\boldsymbol{a}}_k$ and then left-multiplying by the unitary matrix $\boldsymbol{U}$, i.e.,

$$\boldsymbol{a}_k = \boldsymbol{U}[\underbrace{0\cdots0}_{k^*\text{zeros}}\ \check{\boldsymbol{a}}_k^T]^T.$$

In summary, the total transmission power required by this allocation scheme is

$$P_T = \sum_{k=1}^{K} p_k = \sum_{l=1}^{k^*}\gamma_l\tilde{\eta}_l + \frac{(N-k^*)\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}. \quad (10)$$

The spreading sequence matrix is

$$\boldsymbol{A}_T = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k^*}, \boldsymbol{a}_{k^*+1}, \ldots, \boldsymbol{a}_K]$$
$$= \boldsymbol{U}\begin{bmatrix} \sqrt{p_1} & & & \\ & \ddots & & \mathbf{0} \\ & & \sqrt{p_{k^*}} & \\ & \mathbf{0} & & \check{\boldsymbol{A}} \end{bmatrix}$$

and the effective total correlation matrix $\tilde{\boldsymbol{R}}_{T_k}$ observed at the $k$th user's receiver is

$$\tilde{\boldsymbol{R}}_{T_k} = \boldsymbol{A}_T\boldsymbol{A}_T^H + \tilde{\eta}_k\boldsymbol{I}$$
$$= \boldsymbol{U}\begin{bmatrix} \gamma_1\tilde{\eta}_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \gamma_{k^*}\tilde{\eta}_{k^*} & \\ & \mathbf{0} & & \dfrac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}\boldsymbol{I} \end{bmatrix}\boldsymbol{U}^H + \tilde{\eta}_k\boldsymbol{I}.$$

We proceed to show that all users satisfy their SIR constraints with the proposed allocation scheme. Recall that the SIR constraint of the $k$th user is equivalent to $\boldsymbol{a}_k^H\tilde{\boldsymbol{R}}_{T_k}^{-1}\boldsymbol{a}_k \geq e(\gamma_k)$. Indeed, for $k = 1, 2, \ldots, k^*$

$$\boldsymbol{a}_k^H\tilde{\boldsymbol{R}}_{T_k}^{-1}\boldsymbol{a}_k = \frac{\gamma_k\tilde{\eta}_k}{\gamma_k\tilde{\eta}_k + \tilde{\eta}_k} = e(\gamma_k)$$

[2]We use the notation $\mathrm{diag}(\check{\boldsymbol{A}}^T\check{\boldsymbol{A}})$ to indicate the main diagonal of square matrix $\check{\boldsymbol{A}}^T\check{\boldsymbol{A}}$.

and, for $k = k^* + 1, k^* + 2, \ldots, K$

$$\begin{aligned} \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_{T_k}^{-1} \boldsymbol{a}_k &= \check{\boldsymbol{a}}_k^T (\check{\boldsymbol{A}}\check{\boldsymbol{A}}^T + \tilde{\eta}_k \boldsymbol{I})^{-1} \check{\boldsymbol{a}}_k \\ &= \frac{p_k}{\frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)} + \tilde{\eta}_k} \\ &= e(\gamma_k). \end{aligned}$$

We list some observations and comments pertaining to the proposed power and sequence allocation scheme as follows.

- When the effective noise densities of all users are equal, i.e., $\tilde{\eta}_1 = \cdots = \tilde{\eta}_K = \tilde{\eta}$, the channel model reduces to an AWGN channel and it is easy to show that the set of spreading sequences constructed above reduces to the set of optimal spreading sequences for an AWGN channel given in [14] and [17].
- Proposition 3 below indicates that the sequences constructed above minimize the extended total squared correlation (ETSC) of spreading sequences with unequal powers defined as

$$\text{ETSC} = \sum_{k=1}^{K} \sum_{l=1}^{K} |\boldsymbol{a}_k^H \boldsymbol{a}_l|^2. \tag{11}$$

*Proposition 3:* When $K \leq N$, given the transmission powers $p_k$, for $1 \leq k \leq K$, as defined in (5), orthogonal spreading sequences give the minimum ETSC. When $K > N$, given the transmission powers $p_k$, for $1 \leq k \leq K$, as defined in (7), the class of spreading sequence sets characterized by (8) minimizes the ETSC.

*Proof:* See Appendix III. □

This property will be used in Section V to develop iterative algorithms to construct the proposed sequence sets.

- When the effective noise densities and target SIRs of users are not widely spread, there will not be any overfaded user, i.e., $k^* = 0$. In this case, all $K$ users share the $N$ channels and the spreading sequence matrix $\boldsymbol{A}_T$ is characterized by

$$\boldsymbol{A}_T \boldsymbol{A}_T^H = \frac{P_T}{N} \boldsymbol{I}$$
$$\text{diag}(\boldsymbol{A}_T^H \boldsymbol{A}_T) = [p_1, p_2, \ldots, p_K]^T$$

where

$$p_k = e(\gamma_k) \left( \frac{\sum_{l=1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - \sum_{l=1}^{K} e(\gamma_l)} + \tilde{\eta}_k \right)$$

for $k = 1, 2, \ldots, K$, and

$$P_T = \sum_{k=1}^{K} p_k = \frac{N \sum_{l=1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - \sum_{l=1}^{K} e(\gamma_l)}.$$

Due to the fact that $\boldsymbol{A}_T \boldsymbol{A}_T^H = \frac{P_T}{N} \boldsymbol{I}$, the ETSC of the spreading sequences satisfies Welch's bound [25] with equality (WBE), i.e.,

$$\sum_{k=1}^{K} \sum_{l=1}^{K} |\boldsymbol{a}_k^H \boldsymbol{a}_l|^2 = \frac{\left( \sum_{k=1}^{K} p_k \right)^2}{N}.$$

Following [14], we call this set of spreading sequences as generalized WBE (GWBE) sequences with different powers, to distinguish from the WBE sequence set (defined in [26]), in which all sequences are of unit power. With the GWBE sequences, all the $N$ eigenvalues of the matrix $\boldsymbol{A}_T \boldsymbol{A}_T^H$ are the same. This means that all the channels in the CDMA system are evenly occupied and equally crowded.

- For any given system, the choice of the spreading sequences is not unique (although the power allocation scheme is fixed). For example, any unitary rotation of the spreading sequence set has the same eigenvalue distribution and is still a valid allocation to support all users. If the unitary matrix $\boldsymbol{U}$ is chosen to be an identity matrix, the system becomes a hybrid TDMA–CDMA system. Elementary vectors are assigned as spreading sequences of the overfaded users. As a result of this allocation, the signal of the $k$th overfaded users is transmitted only during the $k$th-chip interval. Nonoverfaded users do not transmit during the first $k^*$ chip intervals, and the actual spreading gain of the nonoverfaded users reduces to $N - k^*$.
- Using this well-structured power and sequence allocation scheme, the MMSE receiver weight vectors can be easily obtained.

For $k = 1, 2, \ldots, k^*$

$$\begin{aligned} \boldsymbol{W}_k &= \frac{1}{c_k} \tilde{\boldsymbol{R}}_{T_k}^{-1} \boldsymbol{a}_k \\ &= \frac{1}{c_k} \boldsymbol{U}(\boldsymbol{\Lambda} + \tilde{\eta}_k \boldsymbol{I})^{-1} \boldsymbol{U}^H \sqrt{p_k} \boldsymbol{u}_k \\ &= \frac{1}{c_k (1 + \gamma_k) \tilde{\eta}_k} \boldsymbol{a}_k. \end{aligned}$$

For $k = k^* + 1, k^* + 2, \ldots, K$

$$\begin{aligned} \boldsymbol{W}_k &= \frac{1}{c_k} \tilde{\boldsymbol{R}}_{T_k}^{-1} \boldsymbol{a}_k \\ &= \frac{1}{c_k} \boldsymbol{U}(\boldsymbol{\Lambda} + \tilde{\eta}_k \boldsymbol{I})^{-1} \boldsymbol{U}^H \boldsymbol{U} [0 \cdots 0 \check{\boldsymbol{a}}_k^T]^T \\ &= \frac{1}{c_k \left( \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)} + \tilde{\eta}_k \right)} \boldsymbol{a}_k. \end{aligned}$$

We note that the MMSE receivers are the same as the matched filter receivers (up to a scaling factor). This property will also be used in Section V to develop iterative algorithms to construct the proposed sequence set.

## IV. Performance Evaluation

High power efficiency is one main design concern in wireless communication systems. In the forward link, minimizing the transmission power in one cell can reduce the cochannel interference to users in adjacent cells. In this section, we try to gauge the performance of the power and sequence allocation scheme described in the previous section in terms of the total transmission power that it requires to support the SIR targets of the users.

### A. Lower Bound on Transmission Power

In the absence of MAI, the minimum transmission power of the $k$th user to meet the SIR requirement is $p_k = \gamma_k \tilde{\eta}_k$. It is straightforward that the user needs more power in the existence of MAI, i.e., $p_k \geq \gamma_k \tilde{\eta}_k$. Summing up the lower bounds of transmission powers of all users, we obtain a lower bound of the total transmission power

$$P_T \geq \sum_{k=1}^{K} \gamma_k \tilde{\eta}_k = P_B. \tag{12}$$

Because of the possible large spread of the effective noise densities of all users and the *crowdedness* of the system, this trivial lower bound may be loose.

### B. Lagrangian Global Optimal Search

To better gauge the effectiveness of the proposed power and sequence allocation scheme, we seek the minimum total transmission power numerically. We consider the Lagrangian searching method similar to the one in [17] to minimize the transmission power and incorporate the SIR requirements as a penalty function. The Lagrangian function $L$ is formed as follows:

$$L = \sum_{l=1}^{K} \|\boldsymbol{a}_l\|^2 + \sum_{l=1}^{K} \lambda_l (\gamma_l - \boldsymbol{a}_l^H \tilde{\boldsymbol{R}}_l^{-1} \boldsymbol{a}_l)^2$$

where $\lambda_l$, for $1 \leq l \leq K$, are the Lagrange multipliers. Given the system and channel parameters including the spreading gain $N$, number of users $K$, target SIR $\gamma_l$, and effective noise densities $\tilde{\eta}_l$, for $1 \leq l \leq K$, the Lagrangian function $L$ is a function of $\lambda_l$ and the components of $\boldsymbol{a}_l$ for $1 \leq l \leq K$. We consider a gradient search approach to seek a stationary point of the Lagrangian function. The derivative of $L$ with respect to $\lambda_k$ is

$$\frac{\partial L}{\partial \lambda_k} = (\gamma_k - \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_k^{-1} \boldsymbol{a}_k)^2$$

for $k = 1, 2, \ldots, K$. The derivative of $L$ with respect to $\boldsymbol{a}_k$ can be obtained as follows. First we notice that $\tilde{\boldsymbol{R}}_k^{-1}$ is not a function of $\boldsymbol{a}_k$, while $\tilde{\boldsymbol{R}}_l^{-1}$ for $l \neq k$ can be expressed explicitly as a function of $\boldsymbol{a}_k$ via the matrix inversion formula

$$\begin{aligned} \tilde{\boldsymbol{R}}_l^{-1} &= (\boldsymbol{A}_T \boldsymbol{A}_T^H - \boldsymbol{a}_l \boldsymbol{a}_l^H + \tilde{\eta}_l \boldsymbol{I})^{-1} \\ &= (\underbrace{\boldsymbol{A}_T \boldsymbol{A}_T^H - \boldsymbol{a}_l \boldsymbol{a}_l^H - \boldsymbol{a}_k \boldsymbol{a}_k^H + \tilde{\eta}_l \boldsymbol{I}}_{\tilde{\boldsymbol{R}}_{lk}} + \boldsymbol{a}_k \boldsymbol{a}_k^H)^{-1} \\ &= \tilde{\boldsymbol{R}}_{lk}^{-1} - \frac{\tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_{lk}^{-1}}{1 + \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k}. \end{aligned}$$

Therefore, the derivative of $L$ with respect to $\boldsymbol{a}_k$, for $k = 1, 2, \ldots, K$, is given by

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{a}_k} &= 2\boldsymbol{a}_k - 4\lambda_k (\gamma_k - \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_k^{-1} \boldsymbol{a}_k) \tilde{\boldsymbol{R}}_k^{-1} \boldsymbol{a}_k \\ &\quad + 4 \sum_{l \neq k} \lambda_l (\gamma_l - \boldsymbol{a}_l^H \tilde{\boldsymbol{R}}_l^{-1} \boldsymbol{a}_l) \\ &\quad \cdot \left\{ \frac{\boldsymbol{a}_l^H \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k}{1 + \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k} \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_l - \frac{|\boldsymbol{a}_l^H \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k|^2}{(1 + \boldsymbol{a}_k^H \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k)^2} \tilde{\boldsymbol{R}}_{lk}^{-1} \boldsymbol{a}_k \right\}. \end{aligned}$$

At each iteration of the gradient search, $\boldsymbol{a}_k$ is updated using a gradient descent algorithm and $\lambda_k$ is updated using a gradient ascent algorithm.

For $k = 1, 2, \ldots, K$

$$\begin{aligned} \boldsymbol{a}_k &\leftarrow \boldsymbol{a}_k - u_{\boldsymbol{a}} \frac{\partial L}{\partial \boldsymbol{a}_k} \\ \lambda_k &\leftarrow \lambda_k + u_\lambda \frac{\partial L}{\partial \lambda_k} \end{aligned}$$

where $u_{\boldsymbol{a}}$ is the step size of updating the spreading sequences and $u_\lambda$ is the step size of updating the Lagrange multipliers. The gradient descent algorithm may converge to a local minimum of the searching surface. Only the local minima which meet the target SIRs of the users are treated as valid searching results. After convergence, the total transmission power can be calculated as $P_L = \sum_{k=1}^{K} \|\boldsymbol{a}_k\|^2$.

We present three numerical examples to examine the power efficiency of the allocation scheme proposed in Section III. We assume a sample system with spreading gain $N = 7$. The inverse of effective noise densities of the users $\frac{1}{\tilde{\eta}_k}$, for $1 \leq k \leq K$, are generated from independent exponential distributions (corresponding to Rayleigh fading) with $E[\frac{1}{\tilde{\eta}_k}] = 10$ dB. We conducted three simulation examples with $K = 7, 8$, and 9, which represented three cases when the number of users was smaller than or equal to the spreading gain, the number of users was larger than the spreading gain, and the system was heavily loaded, respectively. For all three cases, we conducted 100 independent realizations and compared the best Lagrangian searching results to the total transmission power $P_T$ given in (10) and the trivial lower bound $P_B$ given in (12). For the case of $K = 7$, the target SIRs of the seven users were set to [9 9 7 7 7 4 4] dB. For the case of $K = 8$, the target SIRs of the eight users were set to [9 7 7 4 4 4 4 4] dB. For the case of $K = 9$, the target SIRs of the users were set to [9 9 7 7 4 4 4 2 2] dB.

As shown in Fig. 1, the Lagrangian search always gives the orthogonal sequences which are the optimal sequences for the case of $K = N$. For the two cases of $K > N$, as shown in Figs. 2 and 3, the Lagrangian searching results indicate the existence of spreading sequences which satisfy the SIR requirements with less total transmission power than the power required by the proposed power and sequence allocation scheme. It appears from the search results that the optimal sequence set partitions the channel space in a manner similar to the allocation proposed in Section III. In particular, users with large SIR targets and effective noise densities are assigned orthogonal sequences. However, the distributions of powers and eigenvalues of the matrix $\boldsymbol{A}_T \boldsymbol{A}_T^H$ are different from (8). In addition, we observe that the

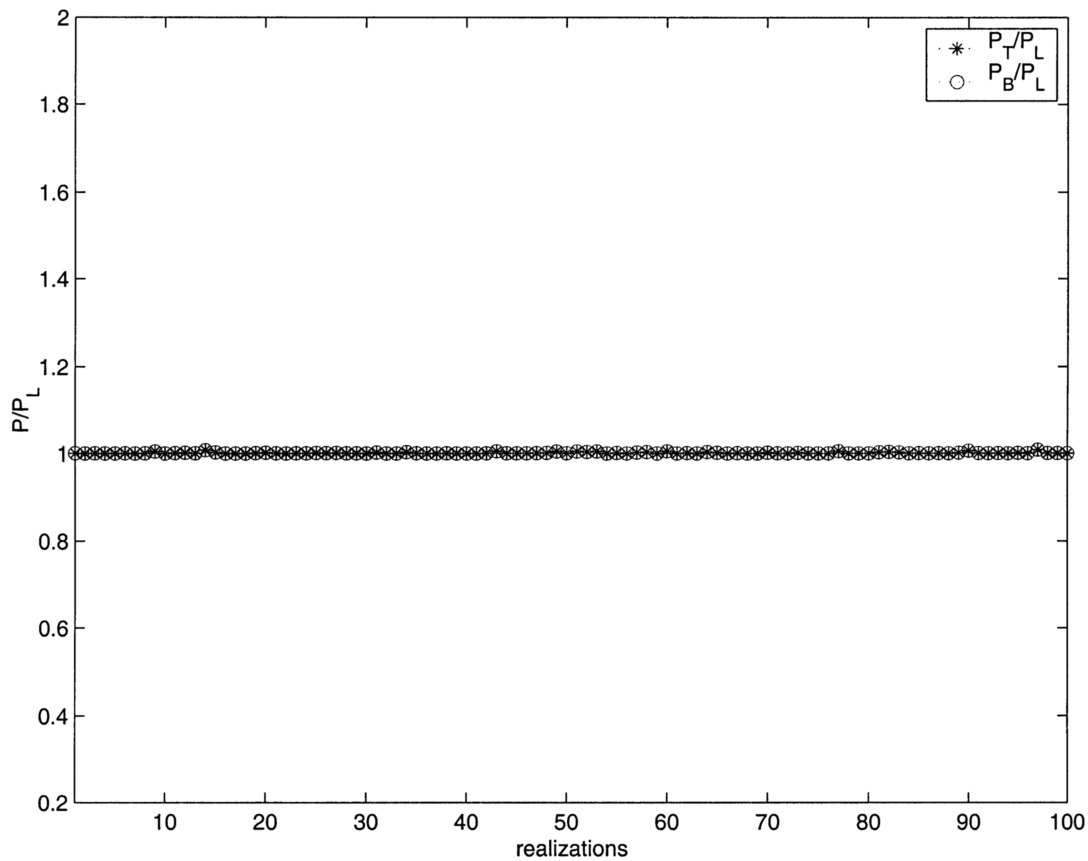


Fig. 1. Lagrangian search results: $N = 7, K = 7$, and $\gamma = [9\ 9\ 7\ \ 7\ \ 7\ \ 4\ \ 4]$ dB.

optimal sequence sets obtained from searching do not have the property that the MMSE filter weights are just scaled versions of the sequences. It turns out that if we impose this condition, i.e., by using the matched filter receiver, then the Lagrangian search results seem to indicate that the proposed sequence allocation is optimal.

In spite of the observation that the sequences constructed in Section III are not optimal in terms of minimizing the total transmission power when $K > N$, results from experiments under different channel conditions indicate that the sequences proposed in Section III only consume slightly more power than the best Lagrangian searching results except for the case when the system is very heavily loaded. As we will show in the next section, an important advantage of the class of sequence sets proposed in Section III over that searched by the Lagrangian method is that the proposed sequence set can be easily constructed using algorithms that are simple and amenable to adaptive implementation. The trival lower bound is also presented in Figs. 1–3. The lower bound becomes very loose when the system is heavily loaded.

## V. Iterative Sequence Construction

Given the effective noise densities and SIR targets of the users in the system, the matrix $\boldsymbol{A}_T = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_K]$ can be constructed following the procedure in Section III. However, this construction is not feasible for practical implementation, especially when the channel conditions change with time. From the preceding analysis, we know that the sequences constructed in Section III are the set of sequences which minimizes the ETSC. Making use of this property and the observation that for this class of sequences, the MMSE receiver weight vectors are the same as those of the matched filter up to a scaling factor, we introduce adaptive algorithms to construct the spreading sequences iteratively.

### A. An Iterative ETSC-Reduction Algorithm

First we extend the TSC-reduction algorithm in [15] to include the case in which each user has a different effective noise density and is allocated a different transmission power. Starting with random spreading sequences

$$\boldsymbol{A}_T[0] = [\boldsymbol{a}_1[0], \boldsymbol{a}_2[0], \ldots, \boldsymbol{a}_K[0]]$$

with the predefined power allocation in (7), the algorithm replaces the spreading sequences by their corresponding MMSE receiver weight vectors (up to a scaling factor to guarantee fixed transmission powers) to obtain new spreading sequences at each iteration. Within each iteration, the spreading sequences of all users are updated sequentially starting from the first user as follows.

*Algorithm 1:* At the $j$th iteration, for the $k$th user, update the effective total correlation matrix

$$\begin{aligned}\boldsymbol{A}_{T_k}[j] &= [\boldsymbol{a}_1[j+1], \ldots, \boldsymbol{a}_{k-1}[j+1], \boldsymbol{a}_k[j], \ldots, \boldsymbol{a}_K[j]]\\ \tilde{\boldsymbol{R}}_{T_k}[j] &= \boldsymbol{A}_{T_k}[j]\boldsymbol{A}_{T_k}^H[j] + \tilde{\eta}_k \boldsymbol{I}.\end{aligned}$$

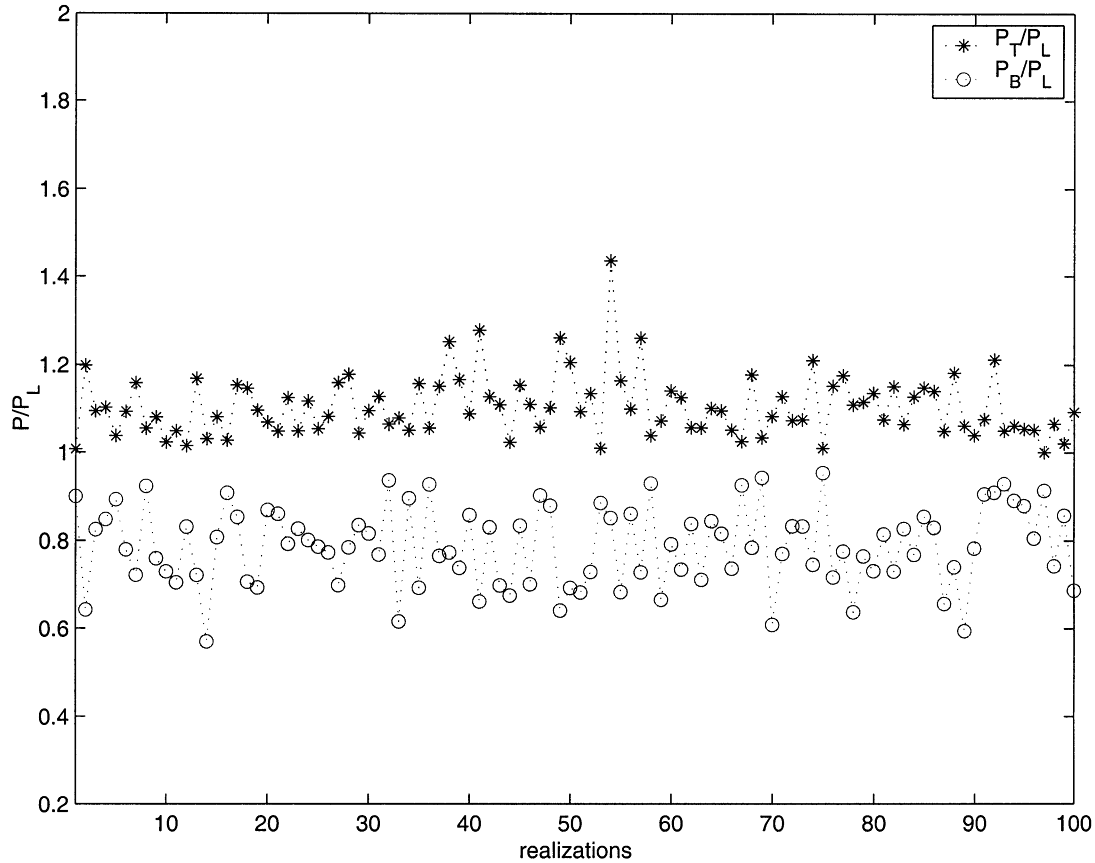


Fig. 2. Lagrangian search results: $N = 7$, $K = 8$, and $\gamma = [9\ 7\ 7\ 4\ 4\ 4\ 4\ 4]$ dB.

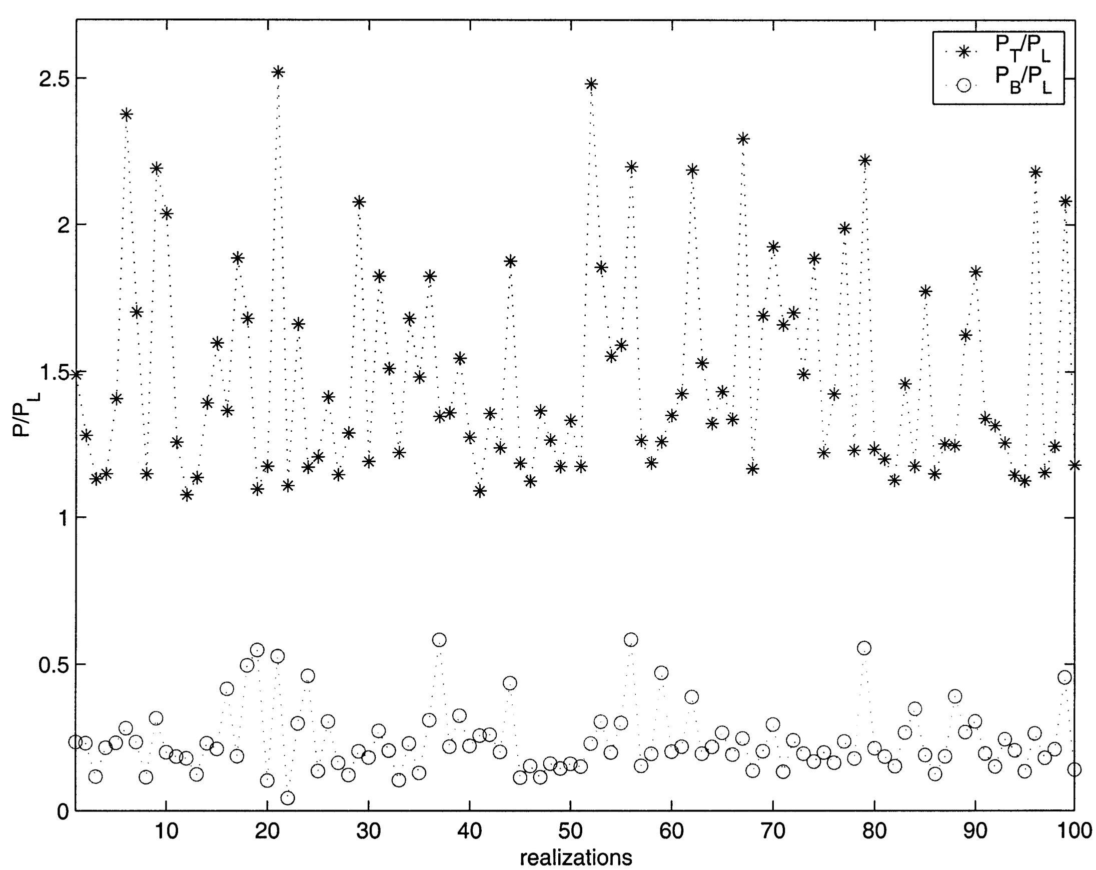


Fig. 3. Lagrangian search results: $N = 7$, $K = 9$, and $\gamma = [9\ 9\ 7\ 7\ 4\ 4\ 4\ 2\ 2]$ dB.

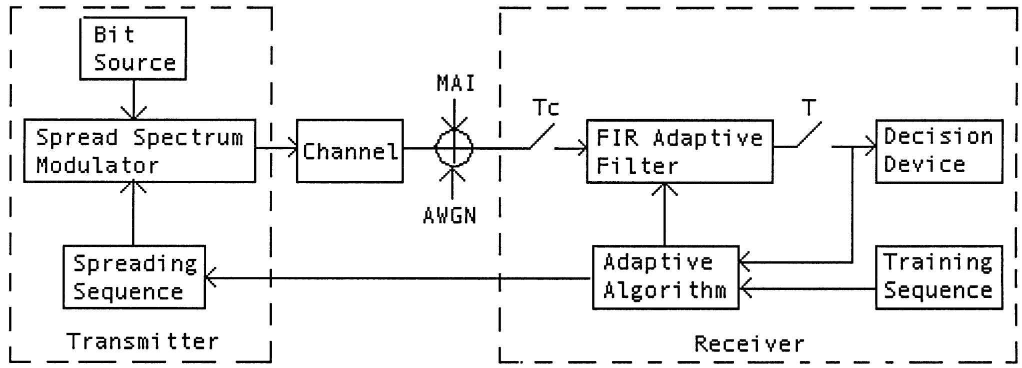


Fig. 4. Adaptive CDMA system block diagram.

Update the sequence by

$$\boldsymbol{a}_k[j+1] = \sqrt{\frac{p_k}{\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_{T_k}^{-2}[j]\boldsymbol{a}_k[j]}}\tilde{\boldsymbol{R}}_{T_k}^{-1}[j]\boldsymbol{a}_k[j]. \quad (13)$$

An equivalent expression of (13), which makes the analysis of convergence of the algorithm easier is

$$\boldsymbol{a}_k[j+1] = \sqrt{\frac{p_k}{\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-2}[j]\boldsymbol{a}_k[j]}}\tilde{\boldsymbol{R}}_k^{-1}[j]\boldsymbol{a}_k[j]$$

where $\tilde{\boldsymbol{R}}_k[j] = \tilde{\boldsymbol{R}}_{T_k}[j] - \boldsymbol{a}_k[j]\boldsymbol{a}_k^H[j]$. A discussion of the convergence of Algorithm 1 is provided in Appendix IV. During the adaptation process, the ETSC decreases monotonically until it reaches a fixed point. We note that the calculation of the transmission powers requires exact knowledge of the effective noise densities. Errors in the estimates of these parameters in a practical system may cause inaccuracy in power allocation and thus performance degradation.

### B. Modified Algorithm With Power Adaptation

Another approach is to adjust both the transmission powers and sequences of the users at each iteration to meet the SIR requirement.

*Algorithm 2:* At the $j$th iteration, for the $k$th user, update the spreading sequence by

$$\boldsymbol{a}_k[j+1] = g_k[j]\tilde{\boldsymbol{R}}_{T_k}^{-1}[j]\boldsymbol{a}_k[j]$$

where the coefficient $g_k[j]$ is chosen so that

$$\boldsymbol{a}_k^H[j+1]\tilde{\boldsymbol{R}}_{T_k}^{-1}[j]\boldsymbol{a}_k[j+1] = e(\gamma_k).$$

Equivalently, the update equation can be written as

$$\boldsymbol{a}_k[j+1] = \sqrt{\frac{e(\gamma_k)}{\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_{T_k}^{-3}[j]\boldsymbol{a}_k[j]}}\tilde{\boldsymbol{R}}_{T_k}^{-1}[j]\boldsymbol{a}_k[j].$$

This algorithm is heuristically obtained from the previous algorithm by forcing the transmission power of each user to a level at which the target SIR is satisfied at each iteration. The class of sequence sets that minimize the ETSC are fixed points of this algorithm. It is not hard to see that this algorithm behaves asymptotically close to Algorithm 1 near the ETSC-minimizing fixed points. Thus, convergence to such a fixed point (in the same sense as the convergence of Algorithm 1) is guaranteed if Algorithm 2 is started sufficiently close to that fixed point. Simulation results show that Algorithm 2 always converges to one of these fixed points starting from random sequences. In addition, this algorithm is more robust against errors in estimating the channel parameters (see Section VI). Nevertheless, it remains an open problem to analytically establish the convergence of Algorithm 2. One main difficulty is that the ETSC does not converge monotonically as in Algorithm 1.

### C. Adaptive Structure With Joint Channel Estimation and Transmitter–Receiver Adaptation

Algorithms 1 and 2 assume explicit knowledge of the effective noise densities to calculate the correlation matrices, and assume that the receivers always work with the optimal weight vectors. In a practical communication system, the channel conditions may change with time. It is impractical to estimate the channel parameters and construct a new set of spreading sequences every time the channel conditions change. It is also difficult for the MMSE receivers to adjust their filter weights when a completely different set of spreading sequences is assigned to the users. In order to solve the problem practically, we develop an adaptive structure with joint channel estimation, transmitter power, and sequence adaptation, and adaptive MMSE signal reception in the forward-link CDMA system. The joint adaptive structure is illustrated in Fig. 4.

With this structure, the base station transmits training symbols in blocks with the current spreading sequences. The receiver weight vectors of all users are adapted independently and simultaneously based on the MMSE criterion, and the estimated MSE at the receivers are fed back to the base station. The base station collects the feedback information from the receivers and estimates the effective noise densities. The spreading sequences of all users are updated at the base station in a centralized manner and then employed to transmit the data of the next block. The iterative procedure consists of the following steps.

1) The base station transmits the $j$th block of training data symbols (contain $L$ data bits) of the $k$th user with the current spreading sequence $\boldsymbol{a}_k[j]$ for $1 \le k \le K$. The initial spreading sequence $\boldsymbol{A}_T[0]$ can be assigned arbitrarily.

2) The weight vector $\boldsymbol{w}_{k,j,l}$ of the receiver of the $k$th user is adapted using exponentially weighted recursive least

squares (RLS) algorithm [27], [23] for $1 \leq k \leq K$. At the end of reception of the $j$th block of data, the MSE for the $j$th block of transmission is estimated as

$$\hat{\varepsilon}_{k,j} = \frac{1}{L}\sum_{l=1}^{L}\left|b_{k,j,l} - \boldsymbol{w}_{k,j,L}^H \boldsymbol{z}_{k,j,l}\right|^2$$

and this estimates is sent back to the base station. Here $\boldsymbol{z}_{k,j,l}$ is the $l$th chip-matched filter output vector, and $\boldsymbol{w}_{k,j,L}$ is the $L$th (last) updated weight vector at the $k$th receiver within the $j$th block of transmission.

3) The base station knows the current spreading sequences $\boldsymbol{A}_T[j]$ and applies eigenvalue decomposition to $\boldsymbol{A}_T[j]\boldsymbol{A}_T^H[j]$

$$\boldsymbol{A}_T[j]\boldsymbol{A}_T^H[j] = \boldsymbol{U}[j]\boldsymbol{\Lambda}[j]\boldsymbol{U}^H[j].$$

Define

$$\tilde{\boldsymbol{a}}_k[j] = \boldsymbol{U}^H[j]\boldsymbol{a}_k[j] = [\tilde{a}_{0,j}^{(k)}, \tilde{a}_{1,j}^{(k)}, \dots, \tilde{a}_{N-1,j}^{(k)}]^T.$$

Then estimates of the effective noise densities of the users can be obtained by solving the following set of equations:

$$1 - \hat{\varepsilon}_{k,j} = \sum_{n=0}^{N-1}\frac{|\tilde{a}_{n,j}^{(k)}|^2}{\lambda_{n,j} + \hat{\eta}_k}$$

for $k = 1, 2, \dots, K$, where $\lambda_{0,j}, \lambda_{1,j}, \dots, \lambda_{N-1,j}$ are the eigenvalues in the diagonal matrix $\boldsymbol{\Lambda}[j]$.

4) Using estimates $\hat{\eta}_k$, the base station estimates the transmission powers $\hat{p}_k$ for $1 \leq k \leq K$, of the users according to the allocation scheme proposed in Section III and updates the spreading sequences of all users sequentially starting from the first user. The spreading sequences can be updated with either Algorithm 1 or 2 proposed earlier. For the $k$th user, its spreading sequence is updated according to

$$\boldsymbol{a}_k[j+1] = \sqrt{\frac{\hat{p}_k}{\boldsymbol{a}_k^H[j]\hat{\boldsymbol{R}}_{T_k}^{-2}[j]\boldsymbol{a}_k[j]}}\hat{\boldsymbol{R}}_{T_k}^{-1}[j]\boldsymbol{a}_k[j]$$

as in Algorithm 1, or according to

$$\boldsymbol{a}_k[j+1] = \sqrt{\frac{e(\gamma_k)}{\boldsymbol{a}_k^H[j]\hat{\boldsymbol{R}}_{T_k}^{-3}[j]\boldsymbol{a}_k[j]}}\hat{\boldsymbol{R}}_{T_k}^{-1}[j]\boldsymbol{a}_k[j]$$

as in Algorithm 2. In the preceding expression

$$\hat{\boldsymbol{R}}_{T_k}[j] = \boldsymbol{A}_{T_k}[j]\boldsymbol{A}_{T_k}^H[j] + \hat{\eta}_k\,\boldsymbol{I}$$

and

$$\boldsymbol{A}_{T_k}[j] = [\boldsymbol{a}_1[j+1], \dots, \boldsymbol{a}_{k-1}[j+1], \boldsymbol{a}_k[j], \dots, \boldsymbol{a}_K[j]].$$

At the end of the $j$th iteration, the updated spreading sequences $\boldsymbol{A}_T[j+1]$ are employed to transmit the next block of data. The adaptation process repeats from Steps 1) to 4) until the end of the training period.

If continuous updating of the spreading sequences and receiver weights are desired after the training period, the system can be switched into a decision feedback mode. Symbol decisions made by the receivers are employed to replace the training symbols.

## VI. Simulation Results

In this section, we study the performance of the proposed adaptive algorithms via computer simulations. Throughout the section, we assume the same sample system and use the same target SIRs described in Section IV.

### A. *Performance of Algorithms 1 and 2*

For each algorithm, we conduct two simulation examples, with $K = 7$ and 9, respectively, to check the convergence of the adaptive algorithms. Perfect channel information is assumed to be available at both the base station and the mobile receivers. Simulation results using Algorithm 1 are shown in Figs. 5 and 6, and simulation results using Algorithm 2 are shown in Figs. 7 and 8. The simulation results indicate that the receiver output SIRs of all users converge to the target SIRs using both algorithms and the total transmission power converges to the total power $P_T$ given in (10) using Algorithm 2. Eigendecomposition of the spreading sequences after convergence indicates that the eigenvalues always converge to the optimal eigenvalue distribution given in (14) starting from random spreading sequences, i.e., both Algorithms 1 and 2 always converge to the fixed point giving the minimum ETSC. We also observe that Algorithm 1 converges faster than Algorithm 2.

### B. *Performance of Joint Channel Estimation and the Transmitter–Receiver Adaptive System*

In this subsection, we study the performance of the proposed joint channel estimation and the transmitter–receiver adaptive system. The performance of the system is measured via the receiver output SIR. More precisely, the receiver output SIR of the $l$th symbol in the $j$th block of the $k$th user's data stream is calculated as

$$\mathrm{SIR}_{k,j,l} = \frac{|\boldsymbol{w}_{k,j,l}^H\boldsymbol{a}_k[j]|^2}{\boldsymbol{w}_{k,j,l}^H\tilde{\boldsymbol{R}}_k[j]\boldsymbol{w}_{k,j,l}}$$

where

$$\tilde{\boldsymbol{R}}_k[j] = \sum_{l\neq k}\boldsymbol{a}_l[j]\boldsymbol{a}_l^H[j] + \tilde{\eta}_k\boldsymbol{I}.$$

The total transmission power during the $j$th block of transmission $P[j]$ is calculated as $P[j] = \sum_{k=1}^{K}\|\boldsymbol{a}_k[j]\|^2$.

During the training period, 3000 training symbols for each user are sent in blocks. After the training period, the system is switched to the decision feedback mode. Ten thousand data symbols are sent and symbol decisions made by the receiver are used to replace the training symbol for continuous update of the receiver weights and spreading sequences. According to the discussion in [23], we employ a block length $L = 100$. We study the performance when there are $K = 8$ users in the system, and set the target SIRs of users as [9 7 7 4 4 4 4 4] dB. Starting from random sequences, all users adapt their transmissions and receptions.

Figs. 9 and 10 show the receiver output SIRs and the total transmission power during the adaptation process for Algo-

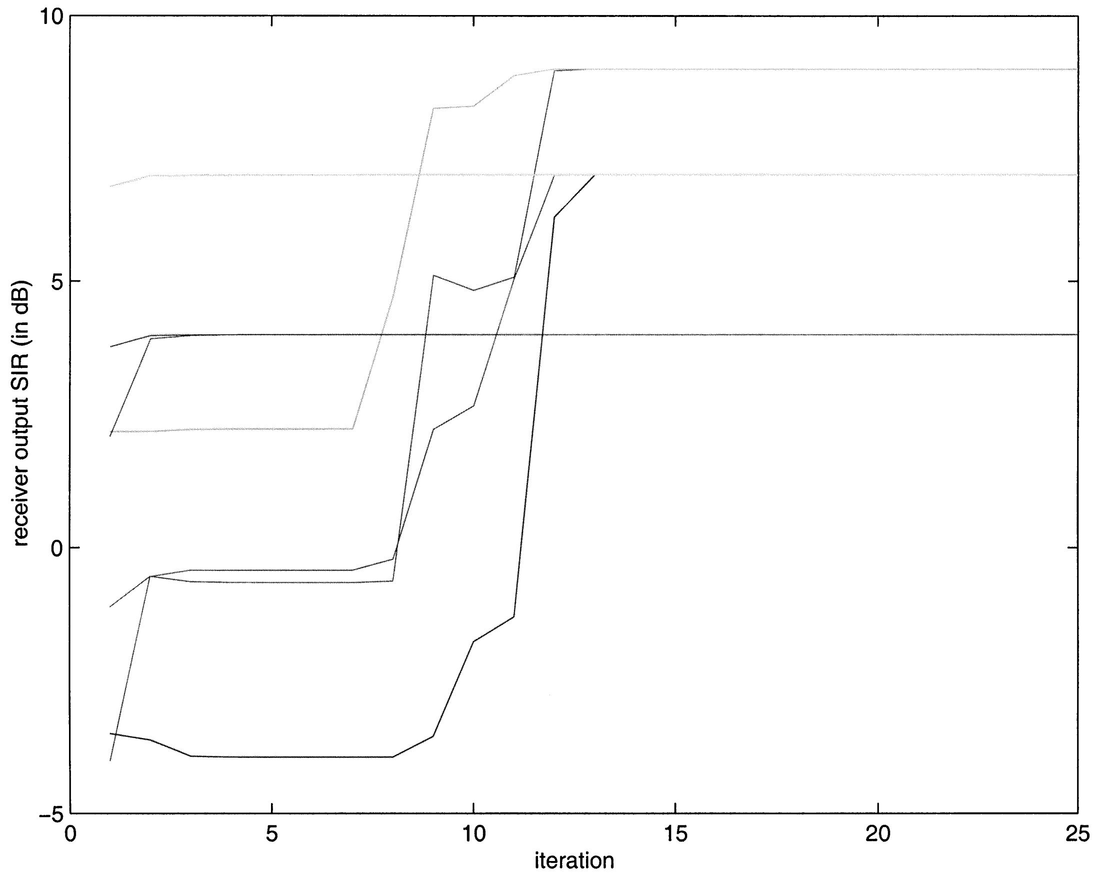


Fig. 5. Convergence of Algorithm 1: $N = 7$, $K = 7$, and $\gamma = [9\ 9\ 7\ 7\ 7\ 4\ 4]$ dB.

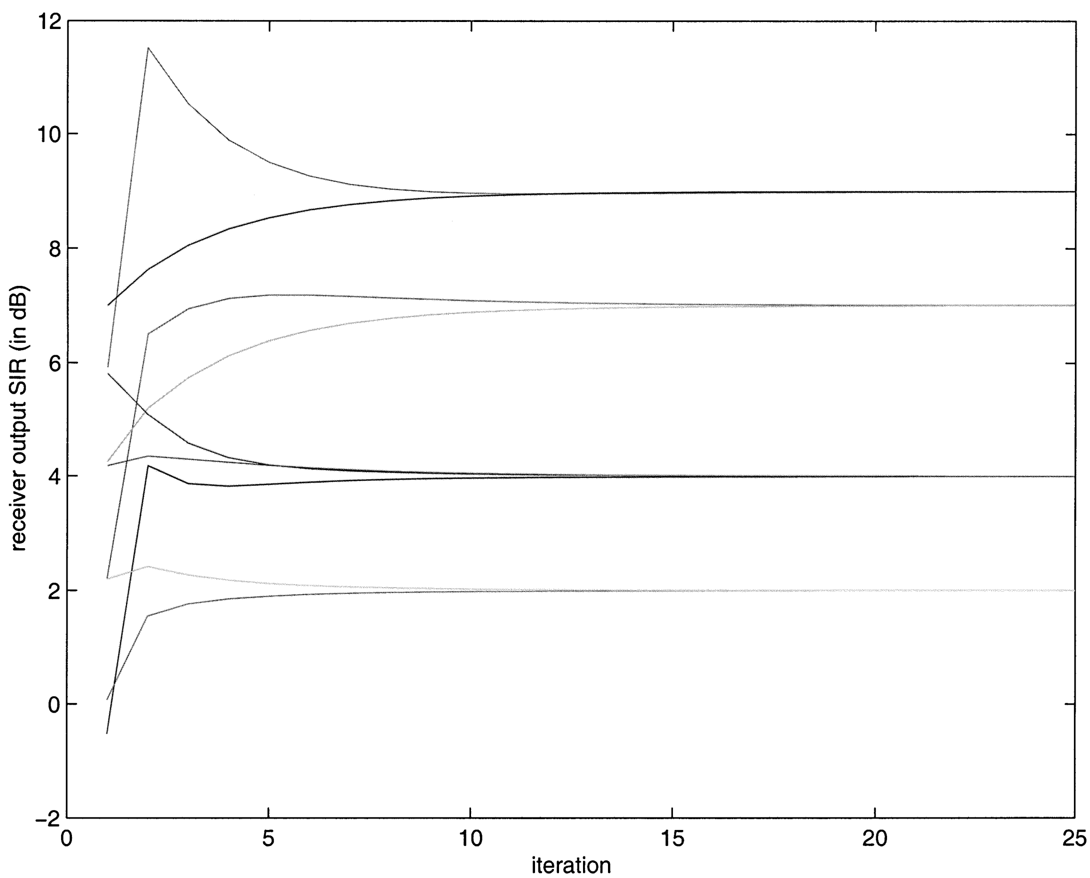


Fig. 6. Convergence of Algorithm 1: $N = 7$, $K = 9$, and $\gamma = [9\ 9\ 7\ 7\ 4\ 4\ 4\ 2\ 2]$ dB.

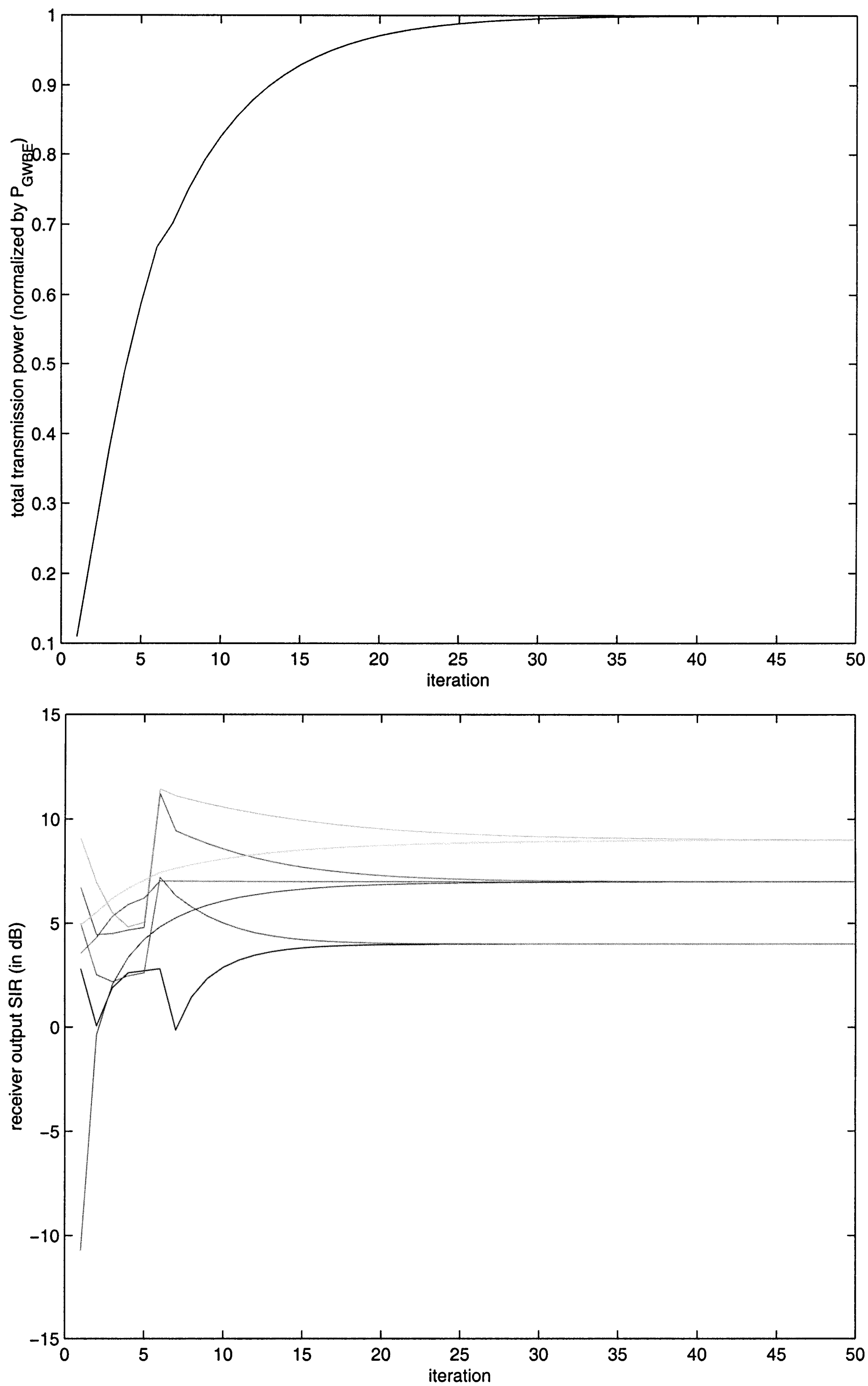


Fig. 7. Convergence of Algorithm 2: $N = 7$, $K = 7$, and $\gamma = [9\ 9\ 7\ 7\ 7\ 4\ 4]$ dB.

rithms 1 and 2, respectively. From these and other results that we have obtained from extensive simulations, it appears that the RLS algorithm is able to track the sequence set updates across block boundaries. A qualitative explanation of why the algorithms work in the decision-feedback mode is that the ETSC-reduction process converges to sequence sets that are the same as the MMSE filter weights. After the initial training period, the updated sequence set is not too different from the previous MMSE filter weights as long as the estimates of the effective noise densities are not too far off. As a consequence, it is easier for the RLS algorithm to track the changes in the sequence set.

Due to the imperfect estimation of channel parameters, there are oscillations for both the SIR and the transmission power curves during adaptation process. Comparing Algorithms 1 and

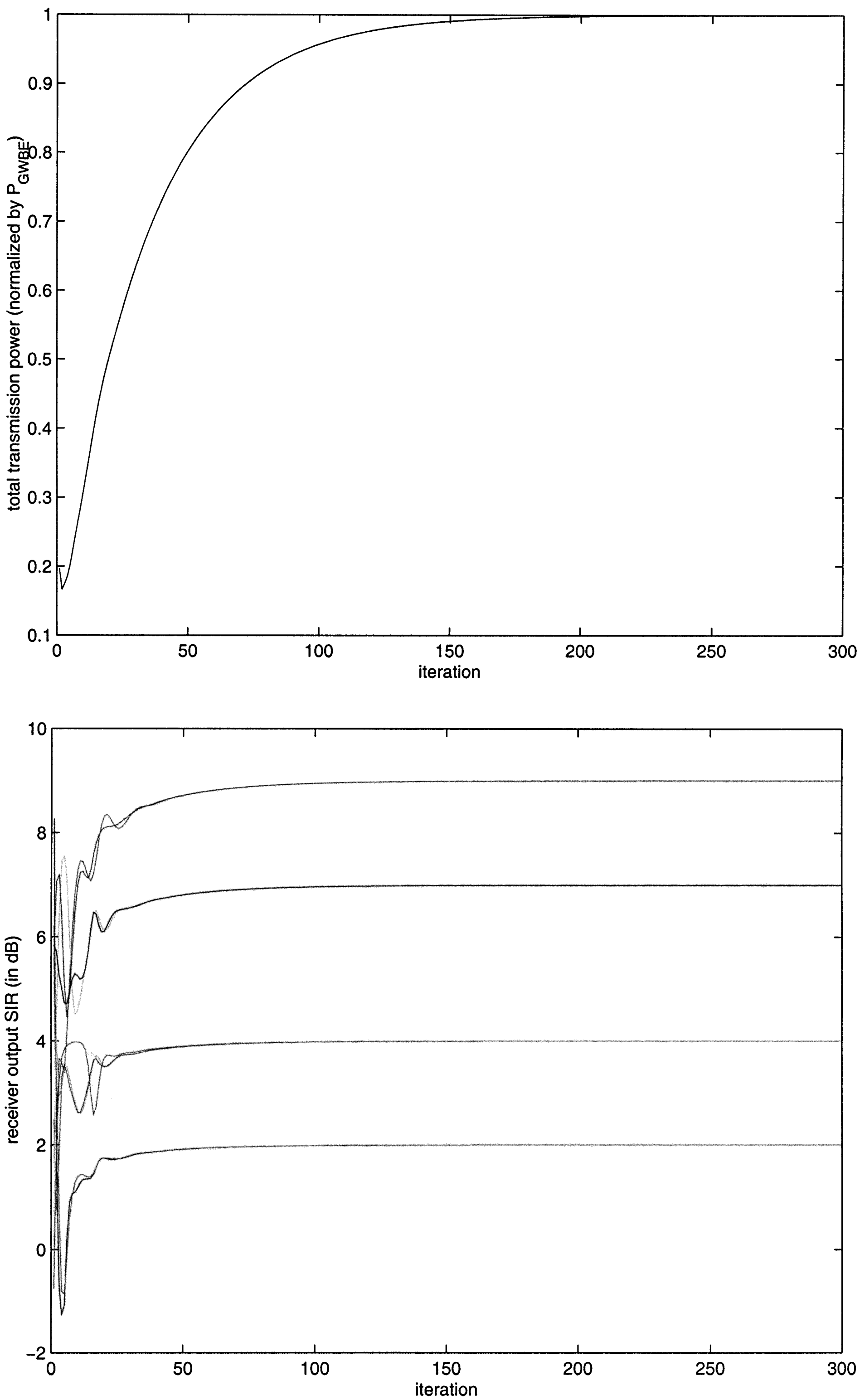


Fig. 8. Convergence of Algorithm 2: $N = 7$, $K = 9$, and $\gamma = [9\ 9\ 7\ 7\ 4\ 4\ 4\ 2\ 2]$ dB.

2, we observe less oscillations when using Algorithm 2. The oscillations resulted from the use of Algorithm 1 arise from the sudden changes in the values of the estimated effective noise densities from block to block. From the simulation results, it appears that the effect of these estimation errors is mitigated by avoiding explicit calculation of the transmission powers from the estimated effective noise densities. As a result, Algorithm 2 appears to be more robust against channel estimation errors than Algorithm 1 and hence is more suitable for practical implementation.

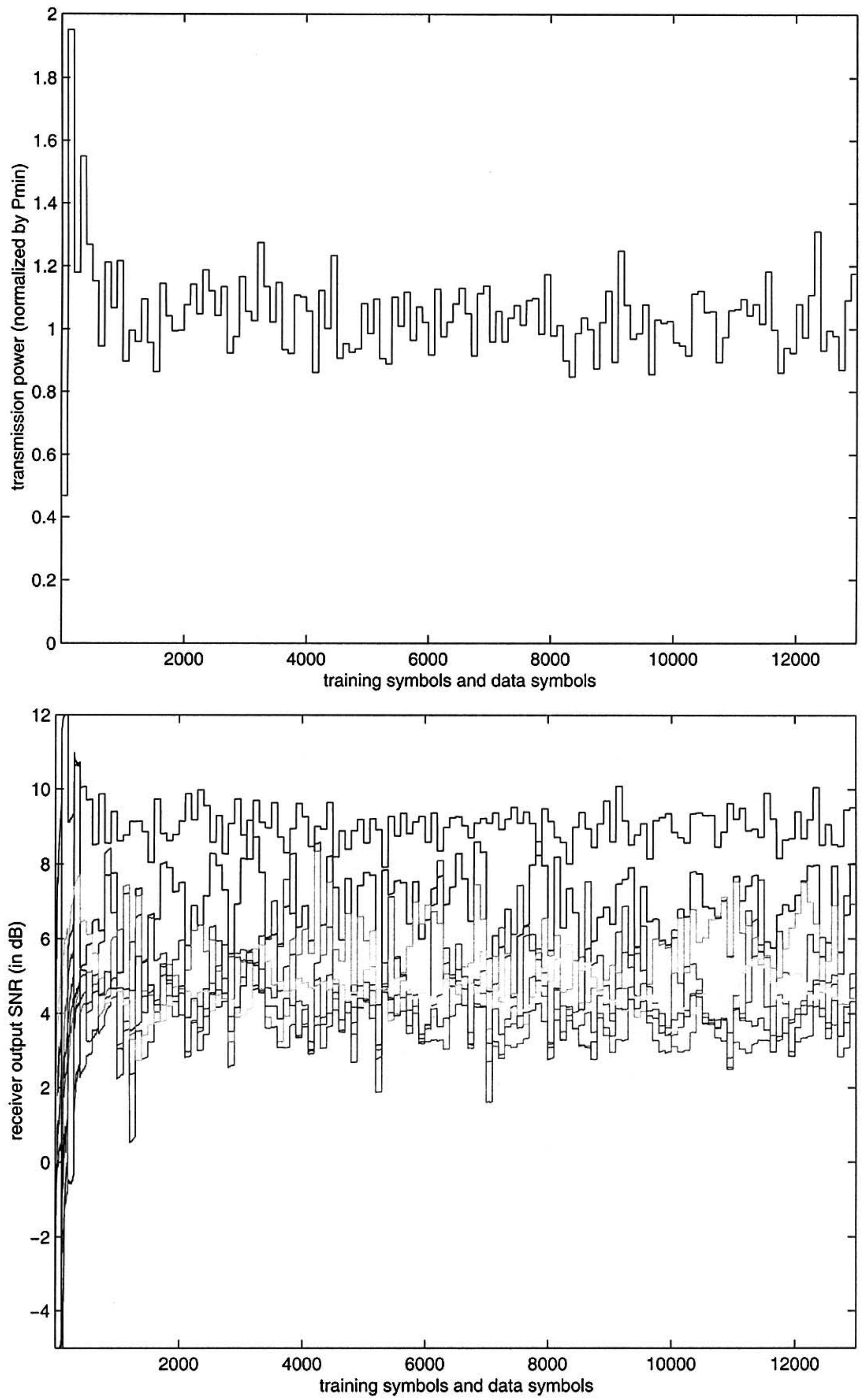


Fig. 9. Performance of joint Tx-Rx adaptation based on Algorithm 1: $N = 7$ and $K = 8$.

## VII. Conclusion

We have proposed a power and sequence allocation scheme within a single cell of a forward-link CDMA system to support users with unequal target SIRs. When the number of users is smaller than or equal to the spreading gain, orthogonal sequences are allocated. Otherwise, we classify users into two classes, namely, overfaded users and nonoverfaded users. Overfaded users are allocated orthogonal *channels* and nonoverfaded users share the remaining *channels*. In both cases, the sequences allocated are the class of sequences which minimizes the ETSC. Employing a Lagrangian search for the optimal set of sequences, we have found that the proposed

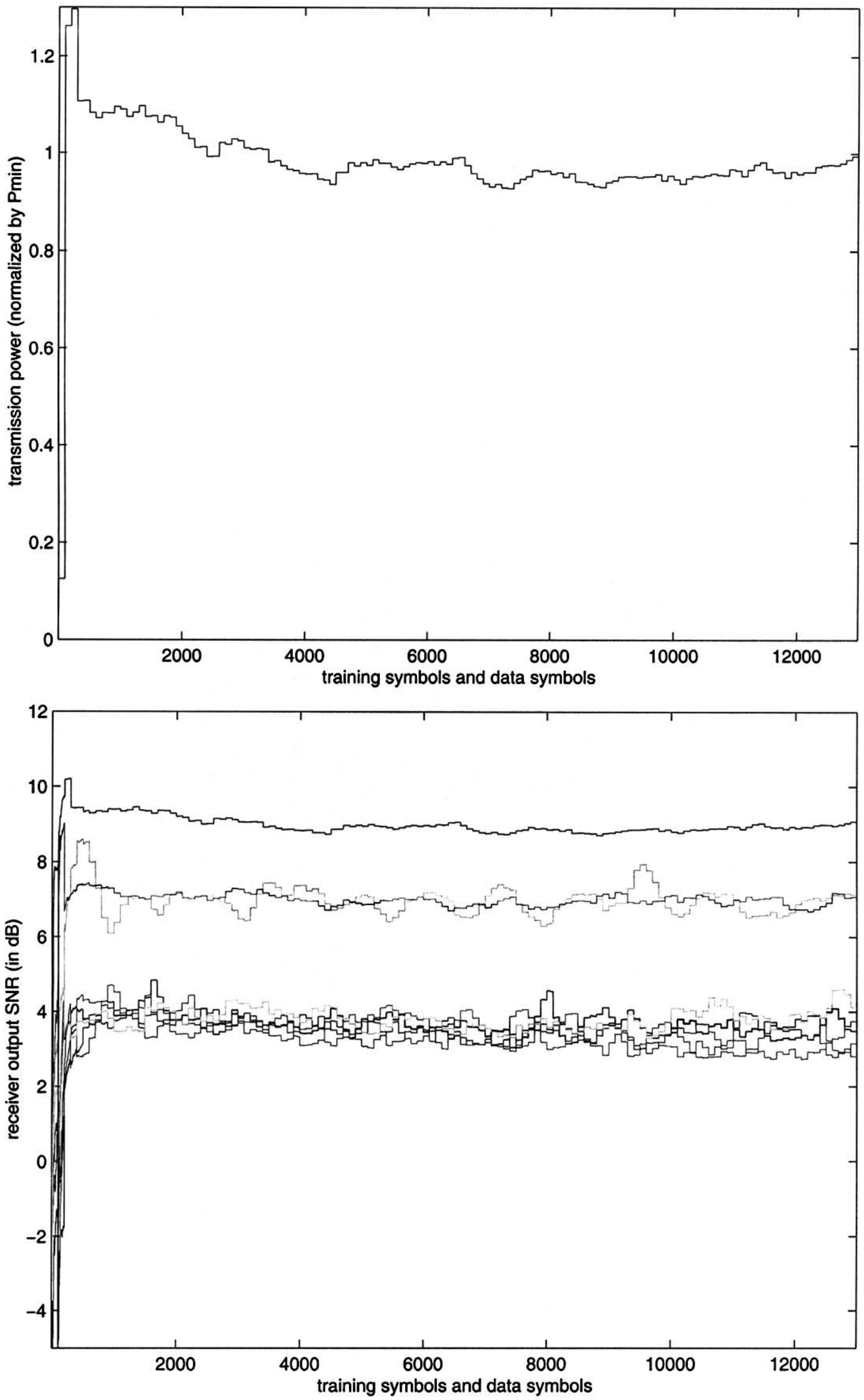


Fig. 10. Performance of joint Tx-Rx adaptation based on Algorithm 2: $N = 7$ and $K = 8$.

sequence allocation schemeconsumes only slightly more power than the best Lagrangian searching results.

Two algorithms are proposed to iteratively update the spreading sequences of the users with scaled MMSE receiver weight vectors. An important advantage, which makes the power and sequence allocation scheme and adaptive algorithms described in this paper desirable from an application viewpoint, is that it is possible to implement the adaptive algorithms based on joint transmitter–receiver adaptation combined with channel parameters estimation in real-life CDMA forward links. Simulation results show that the performance of this joint adaptation structure is robust against estimation errors of channel parameters. Furthermore, only a small amount of information needs to be fed back from the mobiles to the base station to update the spreading sequences. This reduces the bandwidth requirement in the feedback channel.

## Appendix I
## Proof of Proposition 1

We assume a descending order of the products of the effective noise densities and target SIRs, i.e., $\gamma_1\tilde{\eta}_1 \ge \gamma_2\tilde{\eta}_2 \ge \cdots \ge \gamma_K\tilde{\eta}_K$. Define $\gamma_0\tilde{\eta}_0 = \infty$. Let

$$K^* = \max\left\{0 \le k < N : k < N - \sum_{l=k+1}^{K} e(\gamma_l)\right\}.$$

Since $\sum_{l=1}^{K} e(\gamma_l) < N$, $0 \le K^* \le N-1$. Irrespective of how the effective noise densities and target SIRs of all the users are distributed, there always exists a nonempty set

$$\mathcal{F} = \left\{0 \le k \le K^* : \frac{\sum_{l=k+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k - \sum_{l=k+1}^{K} e(\gamma_l)} < \gamma_k\tilde{\eta}_k\right\}.$$

Let $k^* = \max\{k : k \in \mathcal{F}\}$. If $k^*+1 \le K^*$, from the definition of $k^*$, we know that

$$\gamma_{k^*+1}\tilde{\eta}_{k^*+1} \le \frac{\sum_{l=k^*+2}^{K} e(\gamma_l)\tilde{\eta}_l}{N - (k^*+1) - \sum_{l=k^*+2}^{K} e(\gamma_l)}.$$

This is equivalent to

$$\gamma_{k^*+1}\tilde{\eta}_{k^*+1} \le \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}.$$

If $k^*+1 > K^*$, from the definition of $K^*$

$$0 < N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l) \le 1 - e(\gamma_{k^*+1}).$$

In addition

$$[1 - e(\gamma_{k^*+1})]\gamma_{k^*+1}\tilde{\eta}_{k^*+1} = e(\gamma_{k^*+1})\tilde{\eta}_{k^*+1} \le \sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l.$$

Combining these two inequalities, we can readily obtain

$$\gamma_{k^*+1}\tilde{\eta}_{k^*+1} \le \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{1 - e(\gamma_{k^*+1})} \le \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}.$$

This proves the existence of $k^*$ in Proposition 1.

An equivalent expression of

$$\frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)} < \gamma_{k^*}\tilde{\eta}_{k^*}$$

is

$$\frac{\sum_{l=k^*}^{K} e(\gamma_l)\tilde{\eta}_l}{N - (k^*-1) - \sum_{l=k^*}^{K} e(\gamma_l)} < \gamma_{k^*}\tilde{\eta}_{k^*} \le \gamma_{k^*-1}\tilde{\eta}_{k^*-1}.$$

This means that $k^*-1 \in \mathcal{F}$. Inductively, we know that $\{1, 2, \ldots, k^*-2\} \in \mathcal{F}$. Since $\gamma_{k^*+2}\tilde{\eta}_{k^*+2} \le \gamma_{k^*+1}\tilde{\eta}_{k^*+1}$, we always have the inequality

$$\gamma_{k^*+2}\tilde{\eta}_{k^*+2} \le \gamma_{k^*+1}\tilde{\eta}_{k^*+1} \le \frac{\sum_{l=k^*+2}^{K} e(\gamma_l)\tilde{\eta}_l}{N - (k^*+1) - \sum_{l=k^*+2}^{K} e(\gamma_l)}.$$

This is equivalent to

$$\gamma_{k^*+2}\tilde{\eta}_{k^*+2} \le \frac{\sum_{l=k^*+3}^{K} e(\gamma_l)\tilde{\eta}_l}{N - (k^*+2) - \sum_{l=k^*+3}^{K} e(\gamma_l)}$$

which means that $k^*+2 \notin \mathcal{F}$. Inductively, we know that $\{k^*+3, k^*+4, \ldots, K\} \notin \mathcal{F}$. This proves the uniqueness of $k^*$ and the set of overfaded users is $\mathcal{F} = \{1, 2, \ldots, k^*\}$.

The preceding proof validates the following simple procedure that can be used to generate the set of overfaded users.

1) Set $k = 0$.
2) If $\gamma_k\tilde{\eta}_k > \frac{\sum_{l=k+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N-k-\sum_{l=k+1}^{K} e(\gamma_l)}$, then increase $k$ by 1 and repeat Step 2).
3) Otherwise, set $k^* = k-1$, and terminate.

## Appendix II
## Proof of Proposition 2

It is well known that the sum of the diagonal elements of a square matrix is equal to the sum of its eigenvalues. When a matrix is symmetric, the precise relationship between the diagonal elements and the eigenvalues can be described using the theory of majorization, which states [29] that there exists a real symmetric matrix with certain sets of eigenvalues and diagonal elements if and only if the set of eigenvalues majorizes the set of diagonal elements. In the current context, let

$$\lambda = \frac{\sum_{l=k^*+1}^{K} p_l}{N - k^*} = \frac{\sum_{l=k^*+1}^{K} e(\gamma_l)\tilde{\eta}_l}{N - k^* - \sum_{l=k^*+1}^{K} e(\gamma_l)}.$$

From the user classification constraint in (6), we have $\gamma_k\tilde{\eta}_k \le \lambda$ for $k^*+1 \le k \le K$. Thus,

$$p_k = e(\gamma_k)(\lambda + \tilde{\eta}_k) \le \frac{1}{1+\gamma_k}(\gamma_k\lambda + \lambda) = \lambda$$

for $k^*+1 \le k \le K$. Using this fact, it is easy to see that $(\underbrace{\lambda, \ldots, \lambda}_{N-k^*}, \underbrace{0, \ldots, 0}_{K-N})$ majorizes $(p_{k^*+1}, p_{k^*+2}, \ldots, p_K)$. Hence, we can construct a $(K-k^*)\times(K-k^*)$ real symmetric matrix $\boldsymbol{B}$ with $p_{k^*+1}, p_{k^*+2}, \ldots, p_K$ as diagonal elements and $\lambda, \ldots, \lambda, 0, \ldots, 0$ as eigenvalues. Let the eigendecomposition of $\boldsymbol{B}$ be

$$\boldsymbol{B} = \boldsymbol{V}\begin{bmatrix}\lambda\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^T$$

where $\boldsymbol{V}$ is a $(K-k^*)\times(K-k^*)$ orthonormal matrix. Then the required $\check{\boldsymbol{A}}$ characterized in (9) can be obtained as

$$\check{\boldsymbol{A}} = \boldsymbol{W}[\sqrt{\lambda}\boldsymbol{I}\ \ \boldsymbol{0}]\boldsymbol{V}^T$$

where $\boldsymbol{W}$ is an arbitrary $(N-k^*)\times(N-k^*)$ orthonormal matrix.

## APPENDIX III
## PROOF OF PROPOSITION 3

### A. Case 1: $K \le N$

Given the transmission powers $p_k$, for $k = 1, 2, \ldots, K$, as defined in (5), the ETSC is lower-bounded by

$$\begin{aligned}\text{ETSC} &= \sum_{k=1}^{K}\sum_{l=1}^{K}|\boldsymbol{a}_k^H\boldsymbol{a}_l|^2\\ &= \sum_{k=1}^{K}|\boldsymbol{a}_k^H\boldsymbol{a}_k|^2 + \sum_{k=1}^{K}\sum_{l\neq k}|\boldsymbol{a}_k^H\boldsymbol{a}_l|^2\\ &\geq \sum_{k=1}^{K}|\boldsymbol{a}_k^H\boldsymbol{a}_k|^2\\ &= \sum_{k=1}^{K}p_k^2.\end{aligned}$$

We note that this lower bound can be achieved by employing orthogonal sequences when $K \le N$.

### B. Case 2: $K > N$

Given an $N\times K$ sequence matrix $\boldsymbol{A}$ such that $\boldsymbol{A}^H\boldsymbol{A}$ is a $K\times K$ square matrix with diagonal elements $\{p_1, p_2, \ldots, p_K\}$, where $p_k$, for $1 \le k \le K$, are the powers defined in (7). The ETSC of this sequence set (defined in (11)) can be expressed in terms of the nonzero eigenvalues of $\boldsymbol{A}^H\boldsymbol{A}$

$$\begin{aligned}\text{ETSC} &= \operatorname{trace}(\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{A}^H)\\ &= \operatorname{trace}(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H)\\ &= \sum_{n=1}^{N}\lambda_n^2.\end{aligned}$$

For notational convenience, write $\boldsymbol{p} = (p_1, p_2, \ldots, p_K)$ and

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N, \underbrace{0, 0, \ldots, 0}_{K-N\text{ zeros}}).$$

From [29], we know that $\boldsymbol{\lambda}$ majorizes $\boldsymbol{p}$. Thus, the ETSC minimization problem can be rewritten as

$$\min_{\boldsymbol{\lambda}}\sum_{n=1}^{N}\lambda_n^2 \qquad \text{subject to } \boldsymbol{\lambda}\in\mathcal{N}$$

where

$$\mathcal{N} = \Big\{\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N, 0, \ldots, 0) : \lambda_1, \ldots, \lambda_N > 0 \text{ and } \boldsymbol{\lambda} \text{ majorizes } \boldsymbol{p}\Big\}.$$

For the class of spreading sequence sets specified in (8), $\boldsymbol{A}_T^H\boldsymbol{A}_T$ has diagonal element vector $\boldsymbol{p}$ and eigenvalue vector

$$\begin{aligned}\boldsymbol{\lambda}^* &= (\lambda_1^*, \lambda_2^*, \ldots, \lambda_N^*, \underbrace{0, 0, \ldots, 0}_{K-N\text{ zeros}})\\ &= \left(\gamma_1\tilde{\eta}_1, \ldots, \gamma_{k^*}\tilde{\eta}_{k^*}, \frac{\sum_{l=k^*+1}^{K}e(\gamma_l)\tilde{\eta}_l}{N-k^*-\sum_{l=k^*+1}^{K}e(\gamma_l)}, \ldots, \frac{\sum_{l=k^*+1}^{K}e(\gamma_l)\tilde{\eta}_l}{N-k^*-\sum_{l=k^*+1}^{K}e(\gamma_l)}, 0, \ldots, 0\right)\\ &= \left(p_1, \ldots, p_{k^*}, \frac{\sum_{l=k^*+1}^{K}p_l}{N-k^*}, \ldots, \frac{\sum_{l=k^*+1}^{K}p_l}{N-k^*}, 0, \ldots, 0\right).\end{aligned} \tag{14}$$

From the user classification constraint in Proposition 1 and proof of Proposition 2, we know that $\boldsymbol{\lambda}^* \in \mathcal{N}$ and $\lambda_1^* \ge \lambda_2^* \ge \cdots \ge \lambda_N^*$. Since $\lambda_n^2$ is a strictly convex function of $\lambda_n$, the symmetric convex map $f(\boldsymbol{\lambda}) = \sum_{n=1}^{N}\lambda_n^2$ is strictly Schur-convex [29]. A useful property of a strictly Schur-convex function is that if $\boldsymbol{\lambda}_1 \in \mathcal{N}$ majorizes $\boldsymbol{\lambda}_2 \in \mathcal{N}$ and $\boldsymbol{\lambda}_1$ is not a permutation of $\boldsymbol{\lambda}_2$, then $f(\boldsymbol{\lambda}_1) > f(\boldsymbol{\lambda}_2)$. If we can prove that

$$\boldsymbol{\lambda} \text{ majorizes } \boldsymbol{\lambda}^*, \qquad \text{for all } \boldsymbol{\lambda}\in\mathcal{N} \tag{15}$$

then the eigenvalue vector $\boldsymbol{\lambda}^*$ uniquely (up to permutations) minimizes the ETSC. Hence, in this sense the class of sequence sets specified by (8) give the minimum ETSC.

Now we proceed to justify (15). First we note that

$$\sum_{k=1}^{K}p_k = \sum_{k=1}^{N}\lambda_k = \sum_{k=1}^{N}\lambda_k^*$$

where, without loss of generality, we can assume that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$ are the elements of $\boldsymbol{\lambda}\in\mathcal{N}$. In addition, for $k = 1, 2, \ldots, k^*$

$$\sum_{i=1}^{k}\lambda_i \ge \sum_{i=1}^{k}p_i = \sum_{i=1}^{k}\lambda_i^*.$$

We complete the remaining part of the proof by induction. Suppose that for some $k^* \le k < N$

$$\sum_{i=1}^{k}\lambda_i \ge \sum_{i=1}^{k}\lambda_i^*.$$

Since

$$\sum_{i=k+1}^{N}\lambda_i = \sum_{i=1}^{N}\lambda_i^* - \sum_{i=1}^{k}\lambda_i$$

and $\lambda_{k+1} \ge \lambda_{k+2} \ge \cdots \ge \lambda_N$, we have

$$\lambda_{k+1} \ge \frac{\sum_{i=1}^{N}\lambda_i^* - \sum_{i=1}^{k}\lambda_i}{N-k}.$$

Hence,

$$\begin{aligned}\sum_{i=1}^{k+1}\lambda_i &\geq \frac{\sum_{i=1}^{N}\lambda_i^* - \sum_{i=1}^{k}\lambda_i}{N-k} + \sum_{i=1}^{k}\lambda_i \\ &= \frac{\sum_{i=1}^{k}\lambda_i^* + \sum_{i=k+1}^{N}\lambda_i^*}{N-k} + \frac{N-k-1}{N-k}\sum_{i=1}^{k}\lambda_i \\ &\geq \frac{\sum_{i=1}^{k}\lambda_i^*}{N-k} + \lambda_{k+1}^* + \frac{N-k-1}{N-k}\sum_{i=1}^{k}\lambda_i^* \\ &= \sum_{i=1}^{k+1}\lambda_i^*.\end{aligned}$$

In summary, the eigenvalue vector $\boldsymbol{\lambda}^*$ and the class of sequence sets specified by (8) gives the minimum ETSC, which can be expressed as

$$\text{ETSC} = \sum_{l=1}^{k^*} p_l^2 + \frac{\left(\sum_{l=k^*+1}^{K} p_l\right)^2}{N-k^*}.$$

## APPENDIX IV
## CONVERGENCE OF ALGORITHM 1

Algorithm 1 is a generalization of the TSC-reduction algorithm in [15] to the system with unequal transmission powers and effective noise densities. Here the convergence behavior of Algorithm 1 is examined following the idea in [15].

### A. Convergence of ETSC in Algorithm 1

First we prove that the ETSC decreases monotonically at each iteration of Algorithm 1, using which $K$ sequences are updated sequentially starting with the first sequence. For the $k$th sequence, we note that

$$\boldsymbol{A}_{T_k}[j]\boldsymbol{A}_{T_k}^H[j] = \boldsymbol{a}_k[j]\boldsymbol{a}_k^H[j] + \boldsymbol{A}_k[j]\boldsymbol{A}_k^H[j]$$

where

$$\boldsymbol{A}_k[j] = [\boldsymbol{a}_1[j+1], \ldots, \boldsymbol{a}_{k-1}[j+1], \boldsymbol{a}_{k+1}[j], \ldots, \boldsymbol{a}_K[j]].$$

The change in the value of ETSC when updating the $k$th spreading sequence using Algorithm 1 is

$$\begin{aligned}\Delta_k[j] &= \text{trace}[(\boldsymbol{a}_k[j+1]\boldsymbol{a}_k^H[j+1] + \boldsymbol{A}_k[j]\boldsymbol{A}_k^H[j])^2] \\ &\quad - \text{trace}[(\boldsymbol{a}_k[j]\boldsymbol{a}_k^H[j] + \boldsymbol{A}_k[j]\boldsymbol{A}_k^H[j])^2] \\ &= 2\boldsymbol{a}_k^H[j+1]\tilde{\boldsymbol{R}}_k[j]\boldsymbol{a}_k[j+1] - 2\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k[j]\boldsymbol{a}_k[j] \\ &= 2p_k\frac{\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-1}[j]\boldsymbol{a}_k[j]}{\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-2}[j]\boldsymbol{a}_k[j]} - 2\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k[j]\boldsymbol{a}_k[j]. \end{aligned} \quad (16)$$

Since $\tilde{\boldsymbol{R}}_k[j]$ is of full rank and $\boldsymbol{a}_k^H[j]\boldsymbol{a}_k[j] = p_k$, we can employ the Cauchy–Schwartz inequality to write

$$\begin{aligned}p_k^2 &= \left(\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-\frac{1}{2}}[j]\tilde{\boldsymbol{R}}_k^{\frac{1}{2}}[j]\boldsymbol{a}_k[j]\right)^2 \\ &\leq \left\|\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-\frac{1}{2}}[j]\right\|^2\left\|\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{\frac{1}{2}}[j]\right\|^2 \\ &= \left(\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-1}[j]\boldsymbol{a}_k[j]\right)\left(\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k[j]\boldsymbol{a}_k[j]\right)\end{aligned} \quad (17)$$

and

$$\begin{aligned}\left(\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-1}[j]\boldsymbol{a}_k[j]\right)^2 &\leq \left\|\boldsymbol{R}_k^{-1}[j]\tilde{\boldsymbol{a}}_k[j]\right\|^2\|\boldsymbol{a}_k[j]\|^2 \\ &= \left(\boldsymbol{a}_k^H[j]\tilde{\boldsymbol{R}}_k^{-2}[j]\boldsymbol{a}_k[j]\right)p_k.\end{aligned} \quad (18)$$

Combining (16)–(18) yields $\Delta_k[j] \leq 0$, for all $j$, $k = 1, 2, \ldots, K$. Moreover, $\Delta_k[j] = 0$ for $1 \leq k \leq K$ if and only if

$$\boldsymbol{a}_k[j] = \beta_k\boldsymbol{R}_k^{-1}[j]\boldsymbol{a}_k[j] \quad (19)$$

for $1 \leq k \leq K$. An equivalent expression of (19) can be obtained using the matrix inversion formula

$$\boldsymbol{a}_k[j] = \alpha_k\boldsymbol{R}_{T_k}^{-1}[j]\boldsymbol{a}_k[j]. \quad (20)$$

Here $\beta_k$ and $\alpha_k$ are some nonzero scalar constants. This indicates that the value of ETSC keeps decreasing until each spreading sequence is the same as the MMSE receiver weight vector up to a scaling factor. Since the ETSC decreases strictly at each iteration except at the fixed point described above and the ETSC is lower-bounded given a fixed power allocation, convergence of the ETSC is guaranteed.

### B. Fixed Points of Algorithm 1

At a fixed point, the spreading sequences do not change when updated using Algorithm 1. Let us denote a fixed point of Algorithm 1 by the $N \times K$ sequence matrix $\boldsymbol{A}_T = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_K]$. For the case of $K \leq N$, inserting $\tilde{\boldsymbol{R}}_{T_k} = \boldsymbol{A}_T\boldsymbol{A}_T^H + \tilde{\eta}_k\boldsymbol{I}$ into (20), we can rewrite it as

$$\boldsymbol{A}_T\boldsymbol{A}_T^H\boldsymbol{a}_k = (\alpha_k - \tilde{\eta}_k)\boldsymbol{a}_k = \phi_k\boldsymbol{a}_k \quad (21)$$

for $k = 1, 2, \ldots, K$. Premultiplying (21) with $\boldsymbol{a}_l^H$, for $1 \leq l \leq K$, and writing the resulting equations in matrix form, we get

$$\boldsymbol{A}_T^H\boldsymbol{A}_T\boldsymbol{A}_T^H\boldsymbol{A}_T = \boldsymbol{A}_T^H\boldsymbol{A}_T\boldsymbol{\Phi} \quad (22)$$

where $\boldsymbol{\Phi} = \text{diag}[\phi_1, \phi_2, \ldots, \phi_K]$.[3] Referring to [15, Corollary 1], if Algorithm 1 starts with a full rank matrix $\boldsymbol{A}_T[0]$, $\boldsymbol{A}_T^H\boldsymbol{A}_T$ is invertible. Multiplying both sides of (22) with $(\boldsymbol{A}_T^H\boldsymbol{A}_T)^{-1}$ yields

$$\boldsymbol{A}_T^H\boldsymbol{A}_T = \boldsymbol{\Phi}.$$

[3]We use the notation $\text{diag}[\phi_1, \phi_2, \ldots, \phi_K]$ to indicate a $K \times K$ diagonal matrix with $\phi_1, \phi_2, \ldots, \phi_K$ on the main diagonal. When the argument of the same operator is a matrix, we mean the diagonal elements of the matrix.

Thus, $\boldsymbol{A}_T^H \boldsymbol{A}_T$ is diagonal. This indicates the orthogonality of the spreading sequences at the fixed point. Since the powers are fixed during the adaptation, we have

$$\mathrm{diag}[\boldsymbol{A}_T^H \boldsymbol{A}_T] = [p_1, p_2, \ldots, p_K]^T = [\phi_1, \phi_2, \ldots, \phi_K]^T.$$

For the case of $K > N$, consider the eigendecomposition $\boldsymbol{A}_T \boldsymbol{A}_T^H = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$ and let

$$\tilde{\boldsymbol{a}}_k = \boldsymbol{U}^H \boldsymbol{a}_k = [\tilde{a}_{k,1}, \tilde{a}_{k,2}, \ldots, \tilde{a}_{k,N}]^T.$$

We note that the elements $\tilde{a}_{k,n}$, for $n = 1, 2, \ldots, N$, in the sequence vector $\tilde{\boldsymbol{a}}_k$ are the projections of the spreading sequence $\boldsymbol{a}_k$ onto the space defined by the eigenvectors. Based on (21), we get

$$\boldsymbol{\Lambda}\tilde{\boldsymbol{a}}_k = \phi_k \tilde{\boldsymbol{a}}_k \tag{23}$$

for $k = 1, 2, \ldots, K$. Equation (23) indicates that $\phi_k$, for $1 \le k \le K$, are eigenvalues of $\boldsymbol{A}_T \boldsymbol{A}_T^H$. Premultiplying (23) for user $k$ with $\tilde{\boldsymbol{a}}_l^H$, and premultiplying (23) for user $l$ with $\tilde{\boldsymbol{a}}_k^H$, we obtain the following two equations:

$$\begin{aligned} \tilde{\boldsymbol{a}}_l^H \boldsymbol{\Lambda}\tilde{\boldsymbol{a}}_k &= \phi_k \tilde{\boldsymbol{a}}_l^H \tilde{\boldsymbol{a}}_k \\ \tilde{\boldsymbol{a}}_k^H \boldsymbol{\Lambda}\tilde{\boldsymbol{a}}_l &= \phi_l \tilde{\boldsymbol{a}}_k^H \tilde{\boldsymbol{a}}_l. \end{aligned}$$

Since $\boldsymbol{\Lambda}$ is diagonal, we must have

$$\phi_k \tilde{\boldsymbol{a}}_l^H \tilde{\boldsymbol{a}}_k = \phi_l \tilde{\boldsymbol{a}}_k^H \tilde{\boldsymbol{a}}_l \tag{24}$$

for all $l, k = 1, 2, \ldots, K$. This implies that $\phi_k = \phi_l$ when $\tilde{\boldsymbol{a}}_l^H \tilde{\boldsymbol{a}}_k \neq 0$. Write the column set of $\boldsymbol{A}_T$ at the fixed point as $\mathcal{A}$. Then (24) has the following implications.

i) If $\mathcal{A}$ is not split into two or more orthogonal subsets, then $\phi_k = \phi$ for all $k = 1, 2, \ldots, K$, which means that $\boldsymbol{A}_T \boldsymbol{A}_T^H$ is an identity matrix with equal eigenvalues

$$\lambda_1 = \cdots = \lambda_N = \frac{\sum_{k=1}^{K} p_k}{N} = \phi.$$

ii) If $\mathcal{A}$ is split into several orthogonal subsets, $\phi_k$ within each subset are equal.

We note that the proposed set of sequences satisfies this property of the fixed points of Algorithm 1.

### C. Global Minimum of ETSC

First, it is straightforward to check the class of sequence sets specified by (8) is the set of fixed points of Algorithm 1 that uniquely (up to permutations of the eignevalues) minimize the ETSC given the proposed power allocation. With the established monotonicity and convergence of the ETSC in Algorithm 1, it appears that there may be different fixed points with different ways of space partitioning, i.e., the ETSC may converge to local minima. However, it is shown in [16] that all local minima are unstable. As a result, any numerical error or intentional perturbation will cause Algorithm 1 to converge to the global minimum point. Interested readers are referred to [16] for the details of the proof of this property. In this sense, Algorithm 1 converges the class of sequence sets in Section III. The specific sequence set obtained at convergence is determined by the initial sequence set.

## ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and anonymous reviewers for their many thoughtful suggestions which had significantly improved this paper.

## REFERENCES

[1] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[2] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.

[3] ——, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, Apr. 1990.

[4] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.

[5] S. L. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Commun*, vol. 43, pp. 1746–1755, Feb./Mar./Apr. 1995.

[6] C. N. Pateros and G. J. Saulnier, "An adaptive correlator receiver for direct-sequence spread-spectrum communication," *IEEE Trans. Commun.*, vol. 44, pp. 1543–1552, Nov. 1996.

[7] F. Lin and L. B. Milstein, "Successive interference cancellation in multicarrier DS/CDMA," *IEEE Trans. Commun.*, vol. 48, pp. 1530–1540, Sept. 2000.

[8] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth, and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.

[9] Kiran and D. N. C. Tse, "Effective interference and effective bandwidth of linear multiuser receivers in asynchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1426–1447, July 2000.

[10] S. V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1332–1340, Sept. 1995.

[11] R. D. Yates and C. Y. Huang, "Integrated power control and base station assignment," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 638–644, Aug. 1995.

[12] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver Jr., and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303–312, May 1991.

[13] P. B. Rapajic and B. S. Vucetic, "Linear adaptive transmitter-receiver structures for asynchronous CDMA systems," *Europ. Trans. Telecommun.*, vol. 6, no. 1, pp. 21–27, Jan.-Feb. 1995.

[14] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1968–1983, Sept. 1999.

[15] S. Ulukus and R. D. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1989–1998, July 2001.

[16] P. Anigstein and V. Anantharam, "Ensuring convergence of the MMSE iteration for interference avoidance to the global optimum," *IEEE Trans. Inform. Theory*, vol. 49, pp. 873–885, Apr. 2003.

[17] T. M. Lok and T. F. Wong, "Transmitter and receiver optimization in multicarrier CDMA systems," *IEEE Trans. Commun.*, vol. 48, pp. 1197–1207, July 2000.

[18] T. F. Wong and T. M. Lok, "Transmitter adaptation in multicode DS-CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 69–82, Jan. 2001.

[19] S. Ulukus and R. D. Yates, "User capacity of asynchronous CDMA systems with optimum signature sequences," *IEEE Trans. Inform. Theory*, submitted for publication.

[20] J. S. Zhang and E. K. P. Chong, "CDMA systems in fading channels: Admissibility, network capacity, and power control," *IEEE Trans. Inform. Theory*, vol. 46, pp. 962–981, May 2000.

[21] M. Saquib, M. H. Islam, and A. Sampath, "Joint transmitter-receiver optimization in the down-link of CDMA systems," in *Proc. Conf. Information Science and Systems*, Baltimore, MD, Mar. 2001.

[22] E. Visotsky and U. Madhow, "Optimal multiuser space-time transmit filter," University of Illinois, Coordinated Science Laboratory, Urbana, IL, Tech. Rep., 1999.

[23] L. Gao and T. F. Wong, "RLS-based adaptive multicode CDMA system," in *Proc. IEEE Wireless Communications and Networking Conf.'00*, Chicago, IL, Sept. 2000.

[24] H. Y. Kwam and T. M. Lok, "Power control scheme with signature sequence adaptation for DS-CDMA systems," in *Proc. IEEE Wireless Communications and Networking Conf.'00*, Sept. 2000.

[25] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.

[26] M. Rupf and J. L. Massey, "Optimum sequences multisets for synchronous code-division multiple access channels," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1261–1266, July 1994.

[27] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[28] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1984–1991, Sept. 1999.

[29] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.