Transmitter Adaptation in Multicode DS-CDMA Systems

Tan F. Wong and Tat M. Lok

***Abstract*—The problem of transmitter adaptation in the form of adapting the spreading sequences and the transmission powers of different users for a multicode direct-sequence code division multiple access (DS-CDMA) system is considered. Particular attention is given to a distributed algorithm, which updates each pair of transmitter and receiver without information from other users. The transmitter adaptation problem and the algorithm are studied from the viewpoint of a single user, as well as the viewpoint of the whole system. The algorithm is shown to give either the optimal sequences or a choice of sequences that is close to the optimal one. Simulation results show that major improvement in performance can be obtained with the proposed transmission adaptation scheme. The effect of restricting the choice of sequences to polyphase sequences is also considered.**

***Index Terms*—Multicode CDMA, polyphase sequences, power control, sequence optimization, transmitter adaptation.**

I. Introduction

To EXPLOIT the full potential of code division multiple access (CDMA) systems, different reception techniques have been developed. One of the most notable techniques is multiuser detection [1]–[3], where multiple access interference (MAI) is explicitly taken into account in the receiver design rather than just treated as background noises. Another approach is the use of adaptive antenna arrays. Multiple antennas are used at the receiver and digital beamforming techniques are used to enhance the desired signal and suppress the interference [4], [5]. All these approaches focus on the receivers.

Recently, more attention has been paid to the optimization of the transmitters. Transmitter adaptation in the form of power control has been applied in current CDMA systems [6]. The objective of power control is to guarantee that certain SNR performance targets are satisfied at the receivers of all the users by limiting the transmitted powers and, hence, the level of MAI. Distributed power control schemes have been devised [7], [8] to achieve this objective using the minimum amount of total transmission power. The application of power control to multiuser receivers [in particular, the minimum mean square error (mmse) receiver] is also proven to be useful in increasing the capacity of the system [9], [10].

Manuscript received August 1, 1999; revised June 1, 2000. This work was supported in part by the Oak Ridge Associated Universities Ralph E. Powe Junior Faculty Enhancement Award and the Research Grants Council of Hong Kong. This paper was presented in part at the IEEE International Conference on Communications, New Orleans, LA, June 2000.

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).

T. M. Lok is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong (e-mail: tmlok@ie.cuhk.edu.hk).

Publisher Item Identifier S 0733-8716(01)00958-1.

While power control is a vital part of adapting the transmitters, it does not exploit the full potential of transmitter adaptation. If the transmitted signals are chosen or adapted suitably, interference between different user signals can be minimized and hence, the performance of the system improves. Two transmitter adaptation approaches have been suggested based on the assumption that linear optimal (mmse) receivers are employed at the receiving end. The first approach [11] is to precode all the transmitted signals by a linear transformation before transmitting to minimize the MAI. The precoding transform and the linear receivers are chosen jointly to minimize the total mean squared error (MSE) at the receivers. Because of the centralized nature of this approach, it can be applied to the forward link only. The second approach is to choose or adapt the signature sequences of the users so that the MAI levels as seen by the linear optimal receivers are reduced. An early work [12] on this approach suggests iterative replacement of the transmission sequence of a user by the weight vector obtained at the mmse receiver. It is shown that the resulting MSE at the receiver can be reduced, provided that other users' transmissions are fixed. More recently, the optimization problem of choosing a set of signature sequences with minimum total power so that the SNR targets of the users are met over an additive white Gaussian noise (AWGN) channel is solved [13]. A distributed algorithm is also suggested [14] to obtain the optimal WBE sequences for the case where uniform SNR targets are desired. A variant of this optimization problem in a multicarrier setting is solved independently in [15], where centralized and decentralized algorithms based on the method of Lagrange multiplier are also devised to solve the more general problem in a fading channel.

With the increasing demand for multimedia communications, systems should be able to support multirate communications. A simple way to support multirate communications in future wideband CDMA systems is based on the multicode approach, in which multiple spreading sequences for multiple transmission streams are assigned to a user when his/her rate requirement exceeds the basic level. In this paper, we consider transmitter adaptation in the form of adapting the spreading sequences as well as the powers of different transmissions of the users in a synchronous multicode CDMA system.

In Section II, we describe the system model and give a brief discussion on blind linear multiuser detection [3], which is assumed to be employed at the receivers. In Section III, we consider the optimization problem of choosing the spreading sequences with minimum power for transmission. We approach the optimization from both the single-user and multiuser viewpoints. In the single-user view point, only the sequences of one of the users are allowed to adapt, and the transmissions of the other users are fixed. The objective is to select a set of sequences for the transmission streams of the user so that a uniform target SNR is achieved for each stream. In the multiuser viewpoint, all the sequences of all the users are selected jointly to achieve a uniform SNR target with minimum total power. We note that this problem has been considered in [13] and [15]. However, our emphasis is on a novel distributed algorithm. With a suitable interpretation, the results for the single-user multistream problem can also be used in multiple-user scenarios with general noises and interference and are beyond those in [13] and [15].

As will be discussed in Section III, the two viewpoints, with different optimal solutions, represent different scenarios in a wireless communication system. Although the two viewpoints are different, we show that a single simple distributed algorithm can be employed to iteratively select the sequences for transmission. We show that the algorithm gives optimal solutions for the single-user, single-stream, and multiuser adaptation problems. For the single-user multistream problem, the algorithm provides a set of sequences which lie in the subspace spanned by the optimal sequences. Compared to the results in [12], [13], and [15], our approach provides a unified treatment on the sequence optimization problem from the two equally important viewpoints. We discuss some practical considerations, such as restricting the choice of sequences to the set of polyphase sequences, in applying the proposed algorithm in Section IV. Numerical examples obtained from computer simulation are provided in Section V to illustrate the theoretical developments in the previous sections.

II. System Model

In this section, we describe the model of the multicode DS-CDMA system. We assume that there are K simultaneous users in the system. The kth user, for $1 \le k \le K$, generates M_k streams of data symbols. Altogether, there are $M = \sum_{k=1}^{K} M_k$ streams of data symbols to be transmitted. The mth stream, for $1 \le m \le M$, is given by

$$\left(\cdots, b_0^{(m)}, b_1^{(m)}, b_2^{(m)}, \cdots\right). \tag{1}$$

The data symbols $b_j^{(m)}$ are independent random variables with zero mean and unit variance.

The kth user generates M_k periodic spreading sequences of period N. The spreading sequence to spread the mth data stream, for $1 \le m \le M$, is given by

$$\left(\ldots a_0^{(m)}, a_1^{(m)}, \ldots, a_{N-1}^{(m)} \ldots\right). \tag{2}$$

We will use the notation $\mathbf{a}_m$ to denote the vector $[a_0^{(m)}, a_1^{(m)}, \cdots, a_{N-1}^{(m)}]^T$ containing one period of the sequence.

The mth data stream is spread with the mth spreading sequence and is then modulated to give the following signal:

$$s_m(t) = \mathrm{Re}\left[\sum_{i=-\infty}^{\infty} b_{\lfloor i/N \rfloor}^{(m)} a_i^{(m)} \psi(t - iT_c) e^{\mathrm{j}\omega t}\right] \tag{3}$$

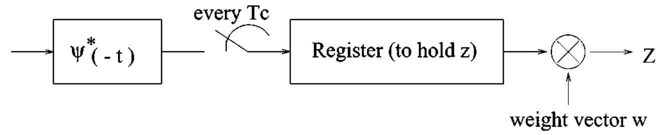

Fig. 1. Distributed linear multiuser receiver for a data stream.

where

T_c delay between consecutive chips;
ω carrier frequency;
$\psi(t)$ chip waveform.

We assume that $\psi(t)$ satisfies the Nyquist criterion for zero-interchip interference, and $\int_{-\infty}^{\infty} |\psi(t)|^2 \, dt = T_c$. For convenience, we set T_c to one hereafter. The transmitted signal for the kth user consists of a sum of M_k signals of the form in (3). For example, the transmitted signal for the first user is given by $\sum_{m=1}^{M_1} s_m(t)$.

We now describe the channel model. We consider a synchronous CDMA system in an additive white Gaussian noise (AWGN) channel. The received signal in complex baseband representation is given by

$$r(t) = \sum_{m=1}^{M} \sum_{i=-\infty}^{\infty} b_{\lfloor i/N \rfloor}^{(m)} a_i^{(m)} \psi(t - i) e^{\mathrm{j}\theta_m} + n(t) \tag{4}$$

where θ_m accounts for the overall phase shift of the mth signal, and $n(t)$ represents the AWGN.

We assume that each data stream is demodulated separately. For example, the receiver shown in Fig. 1 is employed to detect the first data stream. We assume that carrier synchronization has been achieved with the first signal. Therefore, the phase shift θ_1 of the first signal can be taken to be zero. The received signal is passed through a chip-matched filter. The output of the filter is sampled every chip interval. To detect the 0th symbol of the first data stream, we observe the N samples at the output of the chip-matched filter in the interval $[0, T)$, where $T = N$. We arrange the N samples into an N-dimensional vector $\mathbf{z}$. The component of $\mathbf{z}$ due to the first signal is given by $b_0^{(1)}\mathbf{a}_1$. The component of $\mathbf{z}$ due to the mth signal is given by

$$b_0^{(m)} e^{\mathrm{j}\theta_m} \mathbf{a}_m \tag{5}$$

for $m > 1$. Therefore, $\mathbf{z}$ can be written as

$$\mathbf{z} = b_0^{(1)}\mathbf{a}_1 + \sum_{m=2}^{M} e^{\mathrm{j}\theta_m} b_0^{(m)} \mathbf{a}_m + \mathbf{n} \tag{6}$$

where $\mathbf{n}$ denotes the contribution due to the AWGN.

The decision statistic Z for the symbol $b_0^{(1)}$ is obtained by an appropriate linear combination of the N samples, i.e.,

$$Z = \mathbf{w}_1^H \mathbf{z}. \tag{7}$$

The weight vector $\mathbf{w}_1$ is chosen to maximize the SNR defined by

$$\mathrm{SNR} = \frac{|\mathbf{w}_1^H \mathbf{a}_1|^2}{\mathbf{w}_1^H \mathbf{R}_1 \mathbf{w}_1} \tag{8}$$

where the noise-plus-interference correlation matrix $\mathbf{R}_1$ is given by

$$\mathbf{R}_1 = \mathrm{E}\left[\left(\sum_{m=2}^{M} e^{j\theta_m} b_0^{(m)} \mathbf{a}_m + \mathbf{n}\right) \cdot \left(\sum_{m=2}^{M} e^{j\theta_m} b_0^{(m)} \mathbf{a}_m + \mathbf{n}\right)^H\right] = \sum_{m=2}^{M} \mathbf{a}_m \mathbf{a}_m^H + \eta \mathbf{I} \tag{9}$$

with η as the power spectral density of the AWGN. It can be shown that the weight vector that maximizes the SNR is given by

$$\mathbf{w}_1 = \mathbf{R}_1^{-1} \mathbf{a}_1. \tag{10}$$

With this optimal weight vector, the resulting SNR is given by

$$\mathrm{SNR}_1 = \mathbf{a}_1^H \mathbf{R}_1^{-1} \mathbf{a}_1. \tag{11}$$

Equivalently, we can also select the weight vector $\mathbf{w}_1$ to be

$$\mathbf{w}_1 = \mathbf{R}_T^{-1} \mathbf{a}_1 \tag{12}$$

where the total correlation matrix $\mathbf{R}_T$ is given by

$$\mathbf{R}_T = \mathrm{E}[\mathbf{z}\mathbf{z}^H] = \mathbf{R}_1 + \mathbf{a}_1 \mathbf{a}_1^H. \tag{13}$$

This alternative choice of the weight vector also maximizes the SNR since

$$\mathbf{R}_T^{-1} \mathbf{a}_1 = \frac{1}{1 + \mathbf{a}_1^H \mathbf{R}_1^{-1} \mathbf{a}_1} \mathbf{R}_1^{-1} \mathbf{a}_1 \tag{14}$$

i.e., $\mathbf{R}_T^{-1}\mathbf{a}_1$ is just a scalar multiple of $\mathbf{R}_1^{-1}\mathbf{a}_1$ [3]. We note that a blind multiuser detector [3] is obtained by this choice of the weight vector since $\mathbf{R}_T$ can be readily estimated from the received signal. A useful alternative expression for the maximum SNR in (11) is

$$\mathrm{SNR}_1 = \frac{\mathbf{a}_1^H \mathbf{R}_T^{-1} \mathbf{a}_1}{1 - \mathbf{a}_1^H \mathbf{R}_T^{-1} \mathbf{a}_1} \tag{15}$$

where $0 \leq \mathbf{a}_1^H \mathbf{R}_T^{-1} \mathbf{a}_1 < 1$.

The above discussion applies to any data stream by simply replacing the subscript index 1 with the index of the data stream. We also note [13], [15] that

$$\sum_{m=1}^{M} \mathbf{a}_m^H \mathbf{R}_T^{-1} \mathbf{a}_m < N \tag{16}$$

where N is the spreading gain. Because of the above equation, we can associate the important physical interpretation to the quantity $\mathbf{a}_m^H \mathbf{R}_T^{-1} \mathbf{a}_m$ that it represents the "effective" amount of bandwidth [13], [15] out of the total spectrum used by the mth data stream. We will refer to the quantity $\mathbf{a}_m^H \mathbf{R}_T^{-1} \mathbf{a}_m$ as the *effective bandwidth usage* of the mth data stream.

III. Transmitter Adaptation

From the discussion in the previous section, we conclude that the SNR achieved by the optimal linear receiver depends on the spreading sequences. Hence, the performance of the system can be optimized by suitably choosing the spreading sequences.

A. Single-User Single-Stream Adaptation

First, let us consider the problem from the point of view of a single user and assume that all other users do not adapt their transmitters. A possible practical scenario described by this problem is that a new user with high priority is admitted into the system and is about to adapt his/her transmission sequences. The system needs to sacrifice the performance of the existing users to guarantee performance level of this user. Without loss of generality, we focus on the first user. To obtain some insights on the general problem, we start with the simple case where the first user is transmitting only one stream of data symbols, i.e., $M_1 = 1$. The goal of the first user is to achieve a target SNR γ with the minimum amount of power. With the optimal linear receiver, this goal can be formulated as the following the optimization problem:

$$\min \|\mathbf{a}_1\|^2$$

subject to

$$\mathbf{a}_1^H \mathbf{R}_1^{-1} \mathbf{a}_1 = \gamma. \tag{17}$$

Equivalently, the constraint can be rewritten as

$$\mathbf{a}_1^H \mathbf{R}_T^{-1} \mathbf{a}_1 = \zeta \tag{18}$$

where $\zeta = \gamma/(1+\gamma)$, and the total correlation matrix $\mathbf{R}_T$ is given by (13). A closed-form solution for this optimization problem can be readily obtained by the method of Lagrange multiplier. The solution of the optimization problem must satisfy the following equation:

$$\mathbf{R}_1^{-1} \mathbf{a}_1 = \alpha \mathbf{a}_1 \tag{19}$$

where α is the Lagrange multiplier chosen to satisfy the constraint. Hence, $\mathbf{a}_1$ should be chosen as an eigenvector of $\mathbf{R}_1^{-1}$ or, equivalently, an eigenvector of $\mathbf{R}_1$. With this choice, the power is given by $\|\mathbf{a}_1\|^2 = \gamma/\alpha$. Hence, to minimize the power, $\mathbf{a}_1$ should be chosen as the eigenvector associated with the smallest eigenvalue $\lambda_{1,N}$ of $\mathbf{R}_1$, and $\|\mathbf{a}_1\|^2 = \lambda_{1,N}\gamma$ (assuming the eigenvalues of $\mathbf{R}_1$ are distinct). This optimal choice of the sequence has a simple physical interpretation. Each eigenvector of $\mathbf{R}_1$ represents a "channel" in the CDMA system, and the corresponding eigenvalue indicates the amount of interference in that channel. To minimize the transmission power, a new data stream should, of course, choose the channel (the sequence) with the least amount of interference.

The addition of the data stream of the first user to the system causes the SNR performance of the existing data streams to degrade. Equivalently, the effective bandwidth usages of the existing data streams are reduced. The degree of this reduction in the effective bandwidth usages of the existing data streams provides us another way to characterize the optimal choice of $\mathbf{a}_1$ described above. First, the total effective bandwidth usage of all the existing data streams after the addition of the first user's data stream is given by

$$\begin{aligned}\sum_{m=2}^{M} &\mathbf{a}_m^H \mathbf{R}_T^{-1} \mathbf{a}_m \\ &= \sum_{m=2}^{M} \mathbf{a}_m^H (\mathbf{R}_1 + \mathbf{a}_1 \mathbf{a}_1^H)^{-1} \mathbf{a}_m \\ &= \sum_{m=2}^{M} \mathbf{a}_m^H \mathbf{R}_1^{-1} \mathbf{a}_m - \sum_{m=2}^{M} \frac{\mathbf{a}_1^H \mathbf{R}_1^{-1} \mathbf{a}_m \mathbf{a}_m^H \mathbf{R}_1^{-1} \mathbf{a}_1}{1 + \mathbf{a}_1^H \mathbf{R}_1^{-1} \mathbf{a}_1} \\ &= \sum_{m=2}^{M} \mathbf{a}_m^H \mathbf{R}_1^{-1} \mathbf{a}_m - \frac{\mathbf{a}_1^H (\mathbf{R}_1^{-1} - \eta \mathbf{R}_1^{-2}) \mathbf{a}_1}{1 + \mathbf{a}_1^H \mathbf{R}_1^{-1} \mathbf{a}_1}. \end{aligned} \tag{20}$$

We note that the first term on the right-hand side of (20) is the total effective bandwidth usage of all the existing data streams before the addition of the first user's data stream, and the second term represents the reduction in the total effective bandwidth usage of all the existing data streams caused by the addition of the new data stream. It is easy to see (e.g., by using the method of Langrange multiplier) that when the effective bandwidth usage of the new data stream is constrained to be ζ, the reduction term is minimized by choosing $\mathbf{a}_1$ to be the eigenvector associated with the smallest eigenvalue of $\mathbf{R}_1$ as before. In summary, the optimal choice of $\mathbf{a}_1$ described above minimizes the penalty on the total effective bandwidth of the existing data streams for the addition of a new one.

To select this optimal sequence, the transmitter of the first user needs to stay idle until the receiver finishes estimating $\mathbf{R}_1$, solves the eigen-problem, and feeds back the optimal sequence. To prevent the transmitter from staying idle, iterative algorithms to calculate the optimal sequence can be developed based on the well-known power method [16]. Proposition 1, whose proof is given in Appendix A, gives an example of such algorithms.

Propostion 1: Assume that $\mathbf{R}_1$ has distinct eigenvalues. Let $\lambda_{1,n}$ be the eigenvalues arranged in a descending order. Given any initial vector $\mathbf{a}_1[0]$,[1] which is not orthogonal to the eigenvector associated with $\lambda_{1,N}$, and if the constant gain g is greater than $\lambda_{1,N}$, then the iterative procedure

$$\mathbf{a}_1[j+1] = g\mathbf{R}_T^{-1}[j]\mathbf{a}_1[j] \tag{21}$$

converges. Further, if g is chosen to be

$$g = \lambda_{1,N}(1+\gamma) \tag{22}$$

the target SNR γ is achieved with the minimum transmission power $\lambda_{1,N}\gamma$.

We note that since the first user's transmitter is not idle during the iterative process, it is more convenient to employ $\mathbf{R}_T$ in the algorithm as it can be estimated directly from the observation vectors.

[1]We use the notation $[j]$ to emphasize that a quantity is varying from iteration to iteration. For example, $\mathbf{R}_T^{-1}[j]$ means the matrix value of $\mathbf{R}_T^{-1}$ at the jth iteration.

B. Single-User Multistream Adaptation

Now, we try to generalize the results in the previous section to the case in which the first user transmits more than one stream of data symbols. We still assume that all other users do not adapt their transmitters. In this case, the goal of the first user is to achieve the target SNRs for all his/her data streams with the minimum amount of total power. For simplicity, we assume a uniform target SNR γ for all the data streams of the first user. Mathematically, the optimization problem is given by

$$\min \sum_{m=1}^{M_1} \|\mathbf{a}_m\|^2$$

subject to

$$\mathbf{a}_m^H \mathbf{R}_m^{-1} \mathbf{a}_m = \gamma \tag{23}$$

for $1 \le m \le M_1$, where

$$\mathbf{R}_m = \sum_{p \neq m} \mathbf{a}_p \mathbf{a}_p^H + \eta \mathbf{I}. \tag{24}$$

Equivalently, the constraints in (23) can be rewritten as

$$\mathbf{a}_m^H \mathbf{R}_T^{-1} \mathbf{a}_m = \zeta \tag{25}$$

for $1 \le m \le M_1$, where $\zeta = \gamma/(1+\gamma)$. In the discussion below, it is more convenient to group all the sequences of the first user together to form the $N \times M_1$ matrix

$$\mathbf{A}_{1\cdots M_1} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_{M_1}]. \tag{26}$$

We note that the total correlation matrix $\mathbf{R}_T$ is now given by

$$\mathbf{R}_T = \mathbf{R}_{1\cdots M_1} + \mathbf{A}_{1\cdots M_1}\mathbf{A}_{1\cdots M_1}^H \tag{27}$$

where

$$\mathbf{R}_{1\cdots M_1} = \sum_{m=M_1+1}^{M} \mathbf{a}_m \mathbf{a}_m^H + \eta \mathbf{I} \tag{28}$$

is the noise-plus-interference correlation observed by the first user.

Using the notation developed above, we can rewrite the optimization problem in (23) in matrix form

$$\min \operatorname{tr}[\mathbf{A}_{1\cdots M_1}^H \mathbf{A}_{1\cdots M_1}]$$

subject to

$$\operatorname{diag}[\mathbf{A}_{1\cdots M_1}^H (\mathbf{A}_{1\cdots M_1}\mathbf{A}_{1\cdots M_1}^H + \mathbf{R}_{1\cdots M_1})^{-1} \mathbf{A}_{1\cdots M_1}] = \zeta \mathbf{I} \tag{29}$$

where the notation $\operatorname{tr}[\cdot]$ indicates the trace of a matrix, and the operator diag$[\cdot]$[2] takes the diagonal of a matrix to form a diagonal matrix.

The multistream optimization problem in (23) is much more complex than the single-stream problem. However, one would expect that the intuition in the single-stream case extends to the

[2]We also use the notation $\operatorname{diag}[x_1, x_2, \ldots, x_N]$ to represent a diagonal matrix with $x_1, x_2, \ldots, x_N$ as the diagonal elements.

multistream problem. Proposition 2, whose proof is given in Appendix B, makes such intuition concrete by saying that the optimal sequences lie in the subspace spanned by the eigenvectors corresponding to M_1 smallest eigenvalues of $\mathbf{R}_{1\cdots M_1}$.

Proposition 2: Suppose $M_1 \le N$. Let $\mathbf{R}_{1\cdots M_1} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ be the spectral factorization of the matrix $\mathbf{R}_{1\cdots M_1}$ with the eigenvalues λ_n of $\mathbf{R}_{1\cdots M_1}$ arranged in a descending order in the diagonal matrix $\mathbf{\Lambda}$. Let β_i, for $i = 1, \ldots, M_1$ be the solution of the optimization problem

$$\min \sum_{i=1}^{M_1} \lambda_{N-M_1+i}\beta_i$$

subject to

$$\sum_{i=1}^{M_1} \frac{\beta_i}{1+\beta_i} = M_1\zeta = M_1\frac{\gamma}{1+\gamma}$$
$$0 \le \beta_1 \le \beta_2 \le \cdots \le \beta_{M_1}. \qquad (30)$$

Construct the $N \times M_1$ matrix $\tilde{\mathbf{U}}$ from the last M_1 columns (keeping the original order) of $\mathbf{U}$ and the $M_1 \times M_1$ unitary matrix $\mathbf{V}$ such that the diagonal elements of

$$\mathbf{V}\mathrm{diag}\left[\frac{\beta_1}{1+\beta_1}, \frac{\beta_2}{1+\beta_2}, \ldots, \frac{\beta_{M_1}}{1+\beta_{M_1}}\right]\mathbf{V}^H$$

are all equal to ζ. Then

$$\mathbf{A}_{1\cdots M_1} = \tilde{\mathbf{U}}\mathrm{diag}\left[\sqrt{\beta_1\lambda_{N-M_1+1}}, \sqrt{\beta_2\lambda_{N-M_1+2}}, \cdots \sqrt{\beta_{M_1}\lambda_N}\right]\mathbf{V}^H \qquad (31)$$

is a solution to the optimization problem in (23), and the resulting minimum total transmission power is $\sum_{i=1}^{M_1}\lambda_{N-M_1+i}\beta_i$.

We note that the $\mathbf{V}$ described above can be obtained by an iterative procedure as described in the proof of Lemma 3 in Appendix B. The solution for the optimization problem in (30) is given in the following proposition.

Propostion 3: Let M^* denote the integer in $\{0, 1, \ldots, M_1 - 1\}$ that

$$M_1 - M^* - 1 \le M_1\zeta < M_1 - M^*. \qquad (32)$$

Then, there exists a unique integer $M_* \in \{0, 1, \ldots, M^*\}$ such that

$$\frac{\sqrt{\lambda_{N-M_1+i}}}{\sum_{m=i}^{M_1}\sqrt{\lambda_{N-M_1+m}}} \ge \frac{1}{M_1 - i + 1 - M_1\zeta} \qquad (33)$$

for $i = 1, \cdots, M_*$, and

$$\frac{\sqrt{\lambda_{N-M_1+i}}}{\sum_{m=M_*+1}^{M_1}\sqrt{\lambda_{N-M_1+m}}} < \frac{1}{M_1 - M_* - M_1\zeta} \qquad (34)$$

for $i = M_* + 1, \ldots, M_1$. Furthermore

$$\beta_i = \begin{cases} 0 & \text{if } 1 \le i \le M_* \\ \dfrac{1}{M_1 - M_* - M_1\zeta} \cdot \dfrac{\sum_{m=M_*+1}^{M_1}\sqrt{\lambda_{N-M_1+m}}}{\sqrt{\lambda_{N-M_1+i}}} - 1 & \text{if } M_*+1 \le i \le M_1 \end{cases} \qquad (35)$$

is a solution to the optimization problem in (30).

We note that since Propositions 2 and 3 hold for any noise-plus-interference matrix $\mathbf{R}_{1\cdots M_1}$, they actually characterize the general solution of the multistream optimization problem for any type of noises and interference. Furthermore, we may also assume that the streams come from multiple users, instead of coming from the same user. Then, the propositions solve the multiuser optimization problem for any type of noises and interference and are beyond those in [13] and [15], which deal with AWGN channels.

The physical interpretation of Proposition 2 is similar to that of the single-stream case, i.e., we should use the M_1 least congested "channels" to add new data streams to minimize the total transmission power required. The exact choice within the M_1 least congested channels is specified by Proposition 3. The integer M_* in Proposition 3 has the physical meaning that the first M_* channels of the M_1 channels chosen by Proposition 2 are still *overcongested*, indicated by the fact that the eigenvalues corresponding to these M_* channels are still large compared to the eigenvalues corresponding to the other $M_1 - M_*$ channels. Since these M_* channels are overcongested, it would be better, in terms of minimizing the total transmission power, to avoid using them. This is achieved by setting $\beta_1, \ldots, \beta_{M_*}$ to zero. Another important observation from Propositions 2 and 3 is that the conventional wisdom of choosing orthogonal sequences for the multiple data streams in a multicode system does not always minimize the transmission power. In fact, only in some very special case, such as when the M_1 least congested channels have the same amount of interference, orthogonal sequences are optimal.

Because of the complexity in constructing the optimal sequences, iterative algorithms to select the transmission sequences are especially desirable for the multi-stream case. Unlike the single-stream case, it is generally difficult to have a simple iterative algorithm with guaranteed convergence to a solution of the multistream optimization problem in (23). However, it can be shown that a simple extension of the iterative algorithm for the single-stream case selects, after a large number of iterations, transmission sequences from the subspace spanned by the eigenvectors corresponding to the M_1 smallest eigenvalues of $\mathbf{R}_{1\cdots M_1}$. Proposition 4, whose proof is given in Appendix D, formalizes this statement.

Propostion 4: Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of $\mathbf{R}_{1\cdots M_1}$ arranged in a descending order. Suppose $M_1 < N$ and $\lambda_{N-M_1} > \lambda_{N-M_1+1}$. Consider the iterative procedure

$$\mathbf{A}_{1\cdots M_1}[j+1] = \mathbf{R}_T^{-1}[j]\mathbf{A}_{1\cdots M_1}[j]\mathbf{G}[j] \qquad (36)$$

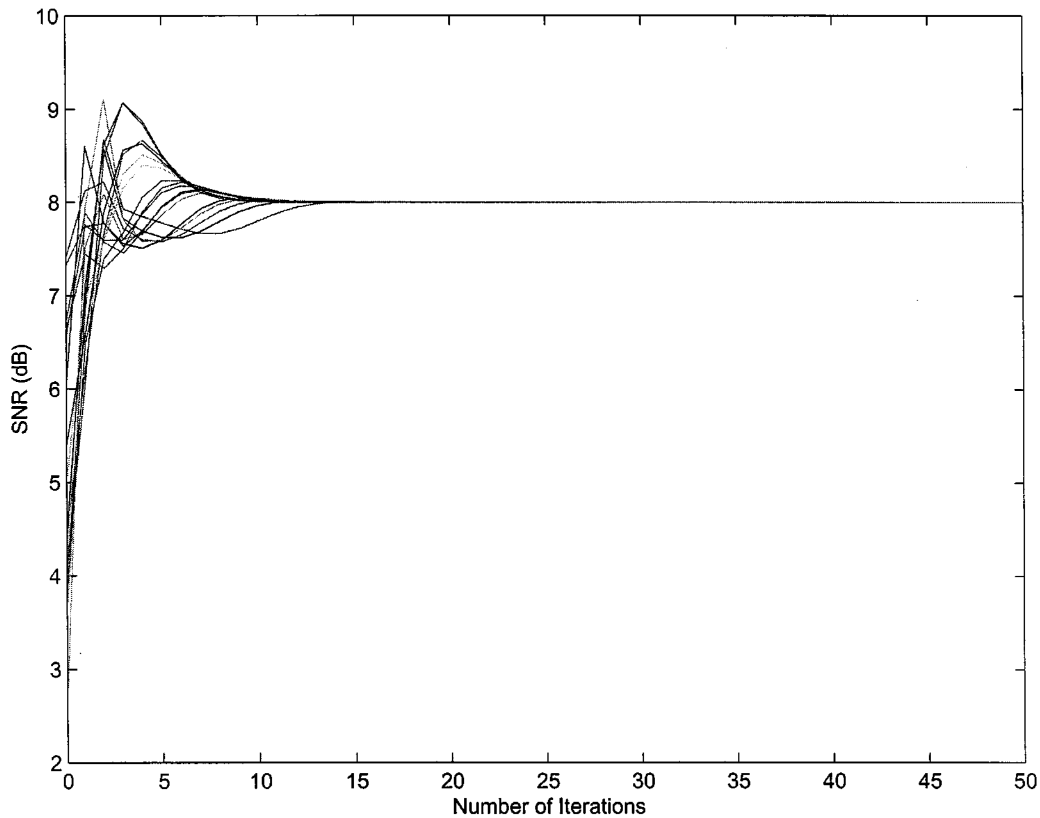

Fig. 2. Convergence of the SNRs achieved by Algorithm 1.

where $\mathbf{G}[0]$, $\mathbf{G}[1]$, $\cdots$ is a sequence of nonsingular $M_1 \times M_1$ matrices chosen to guarantee that the norm of $\mathbf{A}_{1\cdots M_1}[j]$ is bounded below by some $\varepsilon > 0$. Let $\mathcal{U}$ be the subspace spanned by the eigenvectors associated with the M_1 smallest eigenvalues of $\mathbf{R}_{1\cdots M_1}$. If the column space of initial matrix $\mathbf{A}_{1\cdots M_1}[0]$ is of dimension M_1 and it has a nonzero projection onto $\mathcal{U}$, then the column space of $\mathbf{A}_{1\cdots M_1}[j]$ converges to $\mathcal{U}$.

We note that the gain matrices $\mathbf{G}[j]$, $j = 0, 1, \ldots$ act like the constant gain g in the single-stream algorithm to ensure the SNR targets are met. Different choices of the gain matrices will require different total transmission powers. The choice of diagonal gain matrices is of particular interest since it gives rise to decoupled updates for the M_1 data streams.

C. Multiuser Adaptation

In this section, we consider the problem of finding the optimal transmission sequences from the point of view of the whole system. All users are allowed to adapt their transmitters. The sample scenario described by this problem is that after a user leaves the system, all the other users adjust their transmission sequences in order to reduce the overall transmission power needed. In this case, the goal is to achieve the target SNRs for all the data streams of all the users with the minimum amount of total power. Again, we assume a uniform target SNR γ for all the data streams for simplicity. Mathematically, the optimization problem can be expressed in the following form:

$$\min \sum_{m=1}^{M} \|\mathbf{a}_m\|^2$$

subject to

$$\mathbf{a}_m^H \mathbf{R}_m^{-1} \mathbf{a}_m = \gamma \tag{37}$$

for $1 \leq m \leq M$. Equivalently, the constraints in (37) can be rewritten as

$$\mathbf{a}_m^H \mathbf{R}_T^{-1} \mathbf{a}_m = \zeta \tag{38}$$

for $1 \leq m \leq M$, where $\zeta = \gamma/(1+\gamma)$. Again, we group all the sequences of the data streams together to form the $N \times M$ matrix

$$\mathbf{A}_T = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_M\,]. \tag{39}$$

We note that the total correlation matrix $\mathbf{R}_T$ is now given by

$$\mathbf{R}_T = \mathbf{A}_T \mathbf{A}_T^H + \eta \mathbf{I}. \tag{40}$$

The solution of this multiuser optimization problem is given in [13] and [15]. For the case of $M \leq N$, the constraint in (37) can always be satisfied, and the optimal choice of sequences is characterized by

$$\mathbf{A}_T^H \mathbf{A}_T = \eta\gamma \mathbf{I}. \tag{41}$$

The minimum transmission power needed per stream is $\eta\gamma$. On the other hand, for the case of $M > N$, the constraint in (37) can be satisfied if and only if $\zeta < N/M$, and if this condition is satisfied, the optimal choice of sequences is characterized by

$$\mathbf{A}_T \mathbf{A}_T^H = \frac{\eta\zeta}{N/M - \zeta} \mathbf{I}. \tag{42}$$

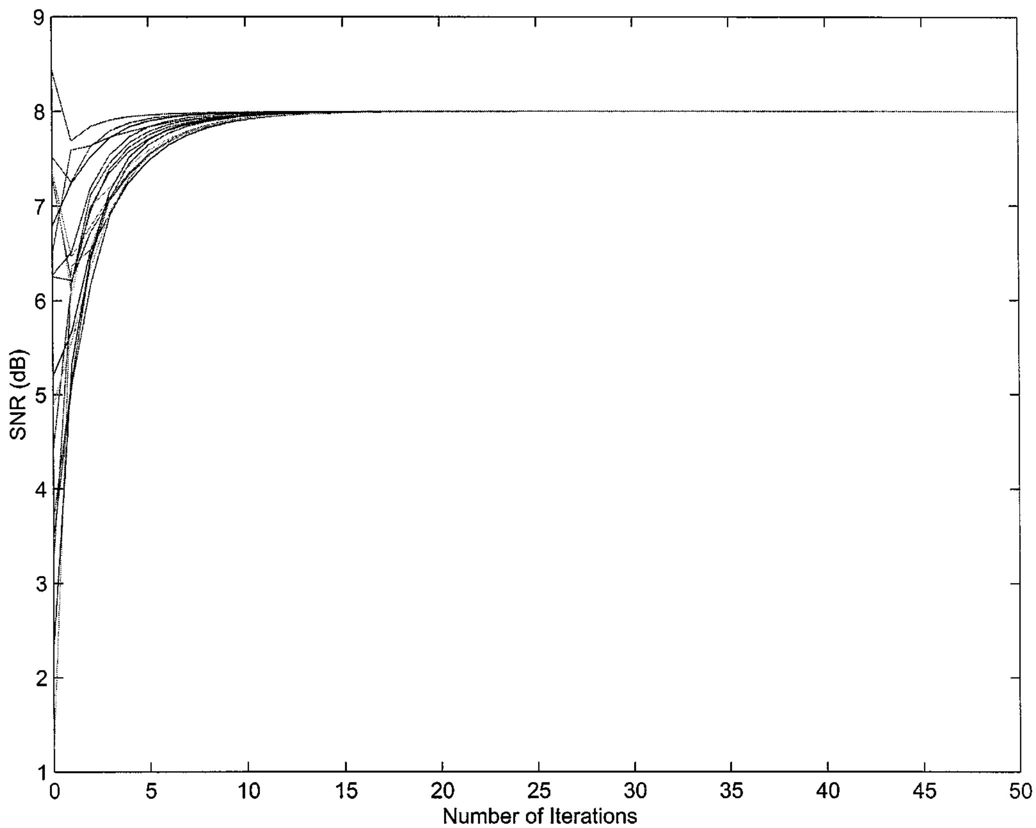

Fig. 3. Convergence of the SNRs achieved by Algorithm 2 with binary sequences.

The minimum transmission power needed per stream is then

$$\frac{\eta\zeta}{1-\dfrac{M}{N}\zeta}=\frac{\eta\gamma}{1-\left(\dfrac{M}{N}-1\right)\gamma}.$$

Proposition 5 below states that the simple extension of the single-stream algorithm in Proposition 1, which adapts all M vectors for all users, converges to a solution of the multiuser optimization problem in (37). The proof of the proposition is given in Appendix E.

Propostion 5: Given any full-ranked initial $\mathbf{A}_T[0]$, if the constant gain g is greater than η, then the iterative procedure

$$\mathbf{A}_T[j+1]=g\mathbf{R}_T^{-1}[j]\mathbf{A}_T[j] \tag{43}$$

converges. For the case of $M \le N$, if g is chosen as

$$g=\eta(1+\gamma) \tag{44}$$

the SNRs of all users' data streams converge to γ. For the case of $M > N$, if

$$\frac{\gamma}{1+\gamma}<\frac{N}{M} \tag{45}$$

g is chosen to as

$$g=\frac{\eta}{1-\dfrac{M}{N}\dfrac{\gamma}{1+\gamma}} \tag{46}$$

and the initial matrix $\mathbf{A}_T[0]$ has the SVD $\mathbf{A}_T[0]=\mathbf{U}\mathbf{S}[0]\mathbf{V}^H$, where $\mathbf{V}$ is such that

$$\operatorname{diag}\left[\mathbf{V}\begin{bmatrix}\mathbf{I}_{N\times N} & \\ & \mathbf{0}_{M-N\times M-N}\end{bmatrix}\mathbf{V}^H\right]=\frac{N}{M}\mathbf{I}_{M\times M} \tag{47}$$

then the SNRs of all users' data streams converge to γ. In either case, $\mathbf{A}_T[j]$ converges to a solution of the optimization problem in (37) requiring the minimum transmission power of $\eta\gamma$ per stream if $M \le N$ or $(\eta\gamma)/(1-(M/N-1)\gamma)$ per stream if $M > N$.

We note that the unitary matrix $\mathbf{V}$ required in the second part of the proposition can be obtained using the iterative method described in Lemma 3. Also, the algorithm above is distributed in nature since each data stream updates independently. The only form of centralized control needed is the synchronization of the updates. Finally, we point out that another iterative algorithm is suggested in [14] to obtain the optimal Welsh bound equality (WBE) sequences described in (41) and (42). The main difference between the algorithm suggested in [14] and the current one is that the former only allows one sequence to adapt at each iteration while all sequences adapt at each iteration in the latter. Moreover, the convergence to the target SNR is not addressed in [14].

IV. Practical Considerations

In this section, we discuss some practical considerations in applying the iterative algorithms in Propositions 1, 4, and 5.

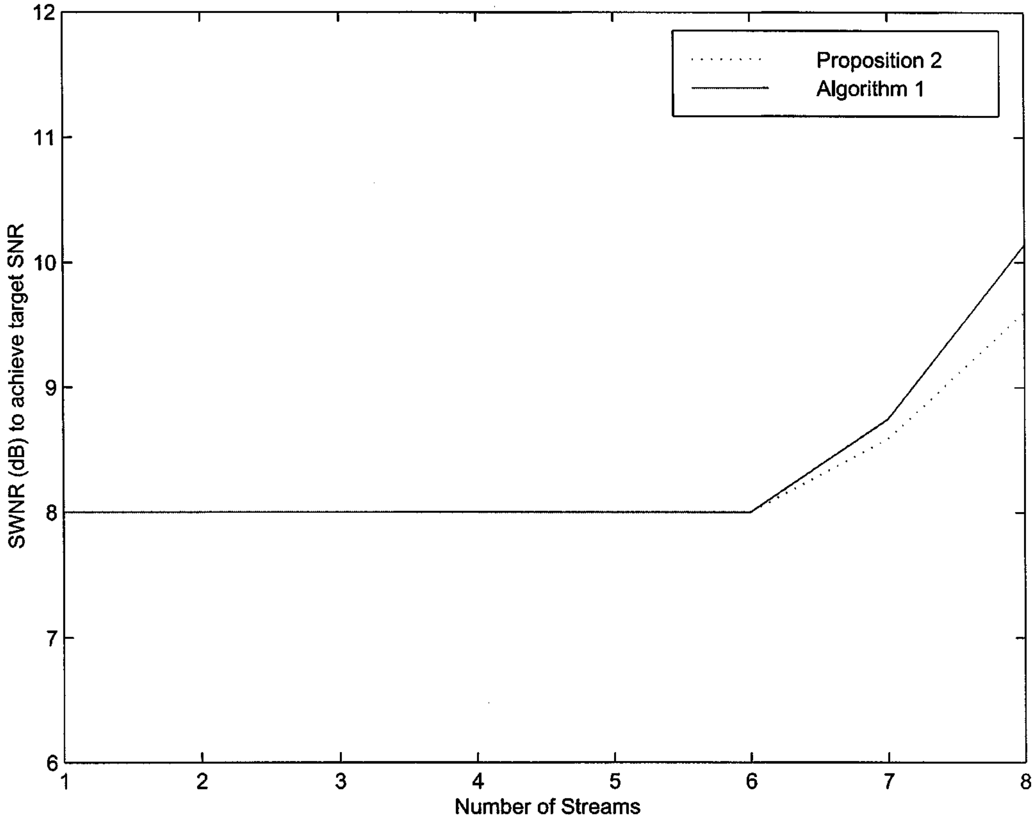

Fig. 4. Average power per stream needed for the first user's data streams in the single-user adaptation setting.

A. Practical Iterative Algorithm

In practice, there should be a central controller which admits/releases users into/from the system. The central controller also determines whether single-user adaptation or multiuser adaptation should be performed. Once this decision is made and the users are informed, all a user needs to do is to update (if he/she is allowed to) his/her transmission sequence(s) according to one of the update rules given in Proposition 1, 4, and 5 to achieve the target SNR. The major drawback of these algorithms lies in the determination of the gain in each iteration step. In order to determine the gain, a user needs to know which kind of transmission adaptation is being performed and channel parameters like $\lambda_{1,N}$ and η. Some of these parameters are often hard to obtain, making the algorithms rather impractical. To alleviate this difficulty, we modify the above algorithms to obtain the following simple algorithm, which does not use any channel parameter and performs update in the same way regardless of which kind of adaptation is being chosen.

Algorithm 1: For the mth data stream, update the sequence vector by

$$\mathbf{a}_m[j+1] = g_m[j]\mathbf{R}_T^{-1}[j]\mathbf{a}_m[j] \tag{48}$$

where the constant gain $g_m[j]$ is chosen so that

$$\mathbf{a}_m^H[j+1]\mathbf{R}_T^{-1}[j]\mathbf{a}_m[j+1] = \gamma(1-\mathbf{a}_m^H[j]\mathbf{R}_T^{-1}[j]\mathbf{a}_m[j]). \tag{49}$$

Simulation results show that this modified algorithm converges under a wide range of situations.

B. Restriction to Polyphase Sequences

In general, the optimized sequences can take on arbitrary values, which may not be desirable for implementation. One usual practical constraint on the sequences is that the sequence elements should have a constant amplitude. The common way to satisfy this constraint is to limit our choices to polyphase sequences, i.e., each sequence element is chosen from a polyphase constellation of P phases $\{c\exp(\mathrm{j}2\pi p/P)\colon\ c > 0, p = 1, \ldots, P\}$. Although, in some cases, the optimal sequences are polyphase sequences, imposing the polyphase constraint on the original optimization problems could make them intractable in general.

Here, we consider a suboptimal approach of forcing the iterative algorithm to choose polyphase sequences. This is done by approximating the resultant sequences in each iteration of Algorithm 1 by polyphase sequences. The modified algorithm is given below

Algorithm 2: For the mth data stream, obtain the vector

$$\tilde{\mathbf{a}}_m[j] = \mathbf{R}_T^{-1}[j]\mathbf{a}_m[j]. \tag{50}$$

Denote the nth element of $\tilde{\mathbf{a}}_m[j]$ by $\tilde{a}_n$. Update the sequence by

$$\mathbf{a}_m[j+1] = g_m[j][e^{\mathrm{j}2\pi p_1/P}, e^{\mathrm{j}2\pi p_1/P}, \ldots, e^{\mathrm{j}2\pi p_N/P}]^T \tag{51}$$

where, for $n = 1, \ldots, N$

$$p_n = \arg\min_p |\arg(\tilde{a}_n) - 2\pi p/P| \tag{52}$$

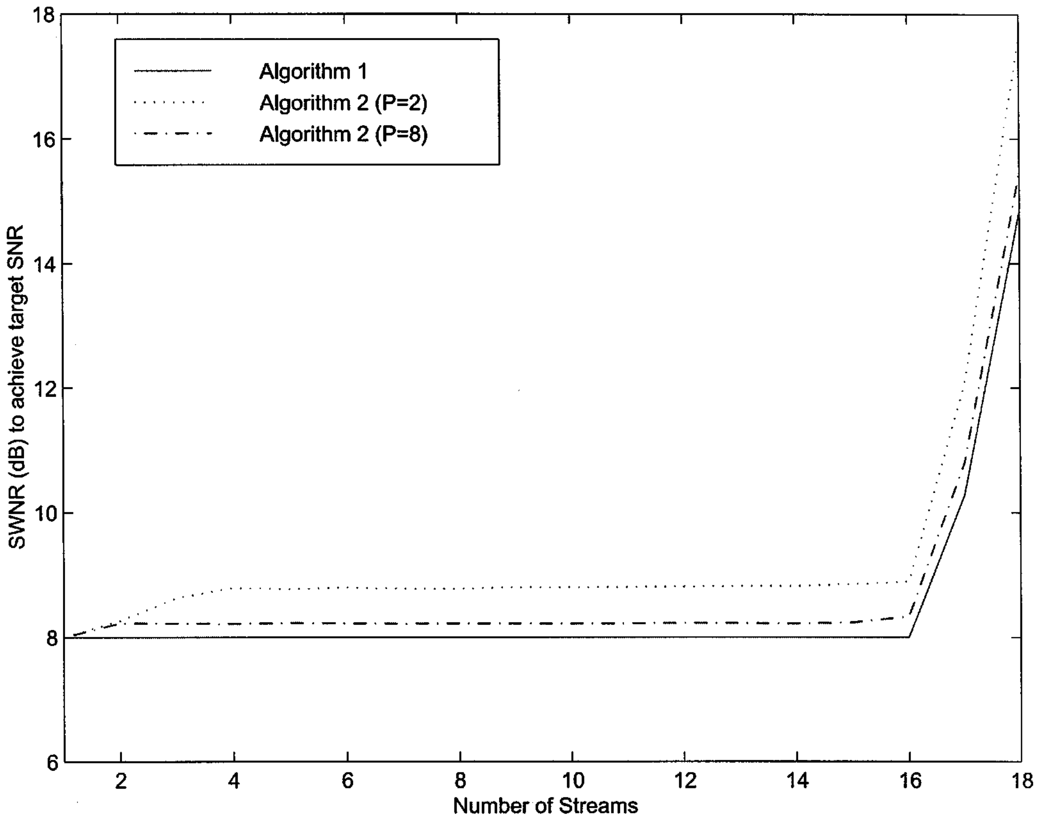

Fig. 5. Average power per stream needed for all data streams in the multiuser adaptation setting.

and the constant gain $g_m[j]$ is chosen so that

$$\mathbf{a}_m^H[j+1]\mathbf{R}_T^{-1}[j]\mathbf{a}_m[j+1] = \gamma(1-\mathbf{a}_m^H[j]\mathbf{R}_T^{-1}[j]\mathbf{a}_m[j]). \quad (53)$$

C. Nonuniform Channel Gains

In a typical wireless communication system, the transmission paths of different users' signals are usually different. Hence, the channel gains for these transmission paths may be different. A more accurate model of the received signal in (4) should contain the nonuniform channel gains for the transmission paths of the users. We note that it is trivial to generalize the results for single-stream and multistream adaptations to the nonuniform channel gain case. However, the generalization for multiuser adaptation presents a much harder problem. First, the characterization of the solution of the multiuser optimization problem for the general nonuniform gain case is still an open question. The main difficulty comes from the fact that the total correlation matrices observed by the receivers of different users are different. Second, iterative algorithms that converges to the optimal solution of the optimization problem are generally unknown. We remark that the iterative algorithm (or its modified form) can still be applied because of its distributed nature.

V. Simulation Examples

In the section, we study the performance of Algorithms 1 and 2 via computer simulations. Throughout the section, we assume that the spreading factor N is 16, and the target SNRs for all data streams are set at 8 dB.

First, we illustrate the convergence of the SNR obtained by the algorithms with two multiuser adaptation examples. We consider a system where there are 17 streams of data. In the first example, each user carries out Algorithm 1 on each stream. The typical SNR performance is shown in Fig. 2. Clearly, the SNR of each stream converges to the target SNR in a small number of iterations. In the second example, each user carries out Algorithm 2 with $P = 2$ on each stream, i.e., all sequences are restricted to binary. The typical performance is shown in Fig. 3. Again, the SNR of each stream converges to the target in a small number of iterations.

Next, we consider the power performance of Algorithm 1 in the single-user adaptation setting. We consider the performance of the first user when Algorithm 1 is applied. We assume that there are ten streams of data being transmitted by other users who do not perform sequence adaptation. Fig. 4, which is obtained from the average of 500 realizations, shows the average power needed per stream when the first user transmits one to eight streams with Algorithm 1. The average power needed using the optimal sequences determined by Proposition 2 is also shown for comparison. We note that the powers shown in Fig. 4 are normalized by the power spectral density of the AWGN and are expressed in terms of the signal-to-white-noise ratio (SWNR). It can be seen that Algorithm 1 gives sequences that perform as well as the optimal ones characterized in Proposition 2 for one to six streams and slightly inferior for seven to eight streams. We also note that when the first user transmits one to six streams, both Algorithm 1 and Proposition 2 yield sequences which avoid all interference from other users, i.e., a SWNR of only 8 dB is needed to achieve a target SNR of 8 dB.

Finally, we consider the power performances of Algorithms 1 and 2 when they are applied by all users in the multiuser adaptation setting. Fig. 5, which is obtained from the average of 500 realizations, shows the average power per stream needed against the total number of streams in the system. The results of Algorithm 1, Algorithm 2 with $P = 2$ (binary sequence restriction), and Algorithm 2 with $P = 8$ (eight-phase sequence restriction) are shown together for comparison. We note that the power needed using Algorithm 1 is almost the same as the theoretical minimal power[3] as prescribed by Proposition 5. When the sequences are restricted to eight phases (Algorithm 2 with $P = 8$), only slightly higher power is needed. However, when the sequences are restricted to binary (Algorithm 2 with $P = 2$), 1–2 dB more power is required.

VI. Conclusion

In this paper, the optimization problem of selecting spreading sequences with minimum transmission powers to satisfy some predetermined SNR targets for the transmission streams of different users in a synchronous multicode DS-CDMA system is approached from the viewpoint of a single user as well as the viewpoint of the whole system. The solutions of the optimization problems from the two viewpoints are characterized and a distributed algorithm is developed to iteratively select the sequences. We show that the algorithm gives the optimal sequences in the cases of single-user single-stream and multiuser adaptations. For the case of single-user multistream adaptation, it gives a choice of sequences that is close to the optimal one. The effect of restricting the choice of sequences to polyphase sequences is shown to be minimal from simulations. Generalizations of the current results to the nonuniform channel gain case are trivial for single-user adaptation but difficult for multiuser adaptation. More research efforts are needed in this direction.

Appendix A
Proof of Proposition 1

We sketch the proof of Proposition 1. Denote the spectral factorization of $\mathbf{R}_1$ by $\mathbf{U}\mathbf{\Lambda}_1\mathbf{U}^H$, where the eigenvalues in $\mathbf{\Lambda}_1$ are arranged in a descending order. Consider the algorithm in (21). Using (14) and the spectral factorization above, it suffices to show that given any initial vector $\mathbf{b}_1[0]$, whose last component is nonzero, the iterative procedure

$$\mathbf{b}_1[j+1] = g\frac{1}{1+\mathbf{b}_1^H[j]\mathbf{\Lambda}_1^{-1}\mathbf{b}_1[j]}\,\mathbf{\Lambda}_1^{-1}\mathbf{b}_1[j] \qquad (54)$$

converges. To do so, let us denote the nth component of the vector $\mathbf{b}_1[j]$ as $b_{1,n}[j]$. Then, we see from (54) that

$$\frac{|b_{1,n}[j]|}{|b_{1,N}[j]|} = \left(\frac{\lambda_{1,N}}{\lambda_{1,n}}\right)^j \frac{|b_{1,n}[0]|}{|b_{1,N}[0]|}. \qquad (55)$$

Hence, as the iteration proceeds, the last component of $\mathbf{b}_1[j]$, namely $b_{1,N}[j]$, dominates all other components in the vector. Hence, it suffices to show the convergence of $b_{1,N}[j]$ (because the other components will converge to zero if $b_{1,N}[j]$ converges). It is easy to see from (54) that if $\mathbf{b}_1[j]$ is obtained at the jth iteration with the initialization $\mathbf{b}_1[0]$, then $e^{j\phi}\mathbf{b}_1[j]$ will be the vector obtained at the jth iteration with the initialization $e^{j\phi}\mathbf{b}_1[0]$ for any ϕ. As a result, it suffices to consider $b_{1,N}[0]$ to be real and positive. Moreover, as (54) proceeds, the iteration of $b_{1,N}[j+1]$ will get increasingly closer to the following:

$$b_{1,N}[j+1] = \frac{\dfrac{gb_{1,N}[j]}{\lambda_{1,N}}}{1+\dfrac{b_{1,N}^2[j]}{\lambda_{1,N}}} = \frac{gb_{1,N}[j]}{\lambda_{1,N}+b_{1,N}^2[j]}. \qquad (56)$$

Therefore, it is enough to establish the convergence of the iteration defined in (56), which is given in Lemma 1.

Lemma 1: Consider the function

$$f(x) = \frac{gx}{x^2+\eta} \qquad (57)$$

where $g > \eta > 0$. Starting with $x > 0$, recursive applications of $f(x)$ converge to $x = \sqrt{g-\eta}$, which is the positive fixed point of the function.

Proof: By simple calculus, we can verify the following

- $f(x)$ has a fixed point at $x = \sqrt{g-\eta}$
- $f(x)$ has a maximum at $x = \sqrt{\eta}$ and $f(\sqrt{\eta}) = g/2\sqrt{\eta}$
- For $x > 0$, $f(x) > 0$, and its monotone increases to its maximum (at $x = \sqrt{\eta}$), and then monotone, decreases toward zero
- The curve of $f(x)$ lies above the line $y = x$ for $0 < x < \sqrt{g-\eta}$, and the curve of $f(x)$ lies below the line $y = x$ for $x > \sqrt{g-\eta}$.

To show the lemma, we have to consider two separate cases

Case 1) $\sqrt{g-\eta} \le \sqrt{\eta}$ (or $g \le 2\eta$)
Case 2) $\sqrt{g-\eta} > \sqrt{\eta}$ (or $g > 2\eta$).

For Case 1, let us divide the positive horizontal axis into three regions:

Region 1) $0 < x < \sqrt{g-\eta}$.
Region 2) $\sqrt{g-\eta} \le x \le \sqrt{\eta}$.
Region 3) $x > \sqrt{\eta}$.

In Region 1, $f(x)$ monotone increases from 0 to $\sqrt{g-\eta}$ and lies above the line $y = x$. Hence, starting in Region 1, repeat applications of $f(x)$ will increase to the fixed point at $x = \sqrt{g-\eta}$. In Region 2, $f(x)$ monotone increases from $\sqrt{g-\eta}$ to $\sqrt{\eta}$ and lies below the line $y = x$. Hence, starting in Region 2, repeat applications of $f(x)$ will decrease to the fixed point at $x = \sqrt{g-\eta}$. In Region 3, $f(x)$ monotone decreases from the maximum value to zero and lies below the line $y = x$. Hence, starting in Region 3, we will enter either Region 1 or 2 after a single application of $f(x)$.

For Case 2, we also divide the positive horizontal axis into three regions:

Region 1) $0 < x < \sqrt{\eta}$.
Region 2) $\sqrt{\eta} \le x \le g/2\sqrt{\eta}$.
Region 3) $x > g/2\sqrt{\eta}$.

In Region 1, $f(x)$ monotone increases from 0 to the maximum value $g/2\sqrt{\eta}$ and lies above the line $y = x$. Hence, starting

[3]The minimum SWNR needed as predicted by Proposition 5 is 8 dB for $M \le 16$, 10.18 dB for $M = 17$, or 14.75 dB for $M = 18$. For $M > 18$, the eight-dB target SNR cannot be achieved regardless of the transmission power.

in Region 1, repeat applications of $f(x)$ will eventually enter Region 2. In Region 2, $f(x)$ monotone decreases from $g/2\sqrt{\eta}$ to $f(g/2\sqrt{\eta})$. Using the condition that $g > 2\eta$, we can show that $f(g/2\sqrt{\eta}) > \sqrt{\eta}$. Hence, for any x in Region 2, $\sqrt{\eta} < f(x) \leq g/2\sqrt{\eta}$. In other words, starting in Region 2, repeat applications of $f(x)$ will never leave Region 2. In Region 3, $f(x)$ monotone decreases from $f(g/2\sqrt{\eta})$ to zero and lies below the line $y = x$. Hence, starting in Region 3, we will enter either Region 1 or 2 after a single application of $f(x)$. Therefore, it suffices to consider Region 2. Note that $f(x)$ lies above and below the line $y = x$ for $\sqrt{\eta} \leq x < \sqrt{g-\eta}$ and for $\sqrt{g-\eta} < x \leq g/2\sqrt{\eta}$, respectively. Hence, successive applications of $f(x)$ will enter each of these two subregions alternatively. For $\sqrt{\eta} \leq x < \sqrt{g-\eta}$, it is easy to verify that $f(f(x)) > x$. Hence, repeat applications of $f(x)$ in Region 2 converge to the fixed point. ■

Going back to the proof of Proposition 1, we conclude from Lemma 1 and the discussion above that the iterative algorithm in (21) converges to the eigenvector associated with $\lambda_{1,N}$ and the resulting power (the norm square of the vector) obtained is $g-\lambda_{1,N}$. To achieve the target SNR, we have to set this limiting power to $\lambda_{1,N}\gamma$.

APPENDIX B
PROOF OF PROPOSITION 2

First, let us rewrite the optimization problem in (29) into a more convenient form. Let $\mathbf{B} = \mathbf{R}_{1\cdots M_1}^{-1/2}\mathbf{A}_{1\cdots M_1}$, and rewrite the optimization problem (29) in terms of $\mathbf{B}$

$$\min \operatorname{tr}[\mathbf{B}^H\mathbf{R}_{1\cdots M_1}\mathbf{B}]$$

subject to

$$\operatorname{diag}[\mathbf{I} - (\mathbf{I}+\mathbf{B}^H\mathbf{B})^{-1}] = \zeta\mathbf{I} \tag{58}$$

where we have used the relation

$$\mathbf{B}^H(\mathbf{B}\mathbf{B}^H+\mathbf{I})^{-1}\mathbf{B} = \mathbf{I} - (\mathbf{I}+\mathbf{B}^H\mathbf{B})^{-1} \tag{59}$$

to obtain the constraint in (58).

Let $\mathbf{R}_{1\cdots M_1} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ be the spectral factorization as described in the statement of the proposition, i.e., the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ are arranged in a descending order, and $\beta_1, \beta_2, \ldots, \beta_{M_1}$ be the eigenvalues of $\mathbf{B}^H\mathbf{B}$ arranged in an ascending order. To proceed, we need to make use of the following result, whose proof can be found, for example, in [17, p. 249].

Lemma 2: Suppose $\mathbf{X}$ and $\mathbf{Y}$ are two Hermitian $N \times N$ matrices. Arrange the eigenvalues x_i of $\mathbf{X}$ in a descending order and the eigenvalues y_i of $\mathbf{Y}$ in an ascending order. Then

$$\operatorname{tr}[\mathbf{XY}] \geq \sum_{i=1}^{N} x_i y_i. \tag{60}$$

Apply this lemma to $\operatorname{tr}[\mathbf{B}^H\mathbf{R}_{1\cdots M_1}\mathbf{B}]$, we have

$$\operatorname{tr}[\mathbf{B}^H\mathbf{R}_{1\cdots M_1}\mathbf{B}] = \operatorname{tr}[\mathbf{R}_{1\cdots M_1}\mathbf{B}\mathbf{B}^H] \geq \sum_{i=1}^{M_1} \lambda_{N-M_1+i}\beta_i. \tag{61}$$

Now, consider the optimization problem (30) in the statement of Proposition 2. Based on (61), we see that the optimization problem in (30) concerns the minimization of the lower bound of $\operatorname{tr}[\mathbf{B}^H\mathbf{R}_{1\cdots M_1}\mathbf{B}]$ under a relaxed constraint

$$\operatorname{tr}[\mathbf{I} - (\mathbf{I}+\mathbf{B}^H\mathbf{B})^{-1}] = \sum_{i=1}^{M_1} \frac{\beta_i}{1+\beta_i} = M_1\zeta. \tag{62}$$

Assuming that $\beta_1, \beta_2, \cdots, \beta_{M_1}$ is the solution of the optimization problem in (30) and $\tilde{\mathbf{U}}$ and $\mathbf{V}$ are chosen as stated in Proposition 2, we proceed to show that the choice of

$$\mathbf{B} = \tilde{\mathbf{U}}\operatorname{diag}\left[\sqrt{\beta_1}, \sqrt{\beta_2}, \ldots, \sqrt{\beta_{M_1}}\right]\mathbf{V}^H \tag{63}$$

gives a solution to the optimization problem in (58).

First, notice that from direct substitution, $\operatorname{tr}[\mathbf{B}^H\mathbf{R}_{1\cdots M_1}\mathbf{B}] = \sum_{i=1}^{M_1}\lambda_{N-M_1+i}\beta_i$. Hence, this choice of $\mathbf{B}$ attains the lower bound in (61). Second, for this choice of $\mathbf{B}$, the lower bound of $\operatorname{tr}[\mathbf{B}^H\mathbf{R}_{1\cdots M_1}\mathbf{B}]$ is minimized since the eigenvalues of $\mathbf{B}^H\mathbf{B}$ are $\beta_1, \beta_2, \ldots, \beta_{M_1}$, which is the solution to the minimization of the lower bound. Combining these two observations, we see that the choice in (63) solves the optimization problem in (58) with the relaxed constraint. Hence, if we can show that this choice of $\mathbf{B}$ also satisfies the original constraint, it will also be a solution to the optimization problem in (58), but this is evident from the construction of $\mathbf{V}$ as

$$\begin{aligned}
&\operatorname{diag}[\mathbf{I} - (\mathbf{I}+\mathbf{B}^H\mathbf{B})^{-1}] \\
&\quad = \mathbf{B}^H(\mathbf{B}\mathbf{B}^H+\mathbf{I})^{-1}\mathbf{B} \\
&\quad = \operatorname{diag}\left[\mathbf{V}\operatorname{diag}\left[\frac{\beta_1}{1+\beta_1}, \frac{\beta_2}{1+\beta_2}, \ldots, \frac{\beta_{M_1}}{1+\beta_{M_1}}\right]\mathbf{V}^H\right] \\
&\quad = \zeta\mathbf{I}.
\end{aligned} \tag{64}$$

Obviously, with $\mathbf{A}_{1\cdots M_1} = \mathbf{R}_{1\cdots M_1}^{1/2}\mathbf{B}$, the original optimization problem in (23) is also solved.

To complete the proof, the following lemma shows that the matrix $\mathbf{V}$ can actually be constructed.

Lemma 3: Suppose that $x_1, x_2, \ldots, x_{M_1}$ are positive numbers such that

$$\sum_{i=1}^{M_1} x_i = M_1\zeta. \tag{65}$$

Then, there exists an M_1-by-M_1 unitary matrix $\mathbf{V}$ such that

$$\operatorname{diag}[\mathbf{V}\operatorname{diag}[x_1, x_2, \ldots, x_{M_1}]\mathbf{V}^H] = \zeta\mathbf{I}. \tag{66}$$

Proof: We prove the lemma by providing an iterative algorithm (with at most M_1 iterations) to construct the matrix $\mathbf{V}$. The construction relies upon the fact that we can redistribute the norm squares of any pair of rows of a matrix by a unitary transform. For example, consider the unitary transform

$$\mathbf{\Phi} = \begin{bmatrix} \cos\phi & \sin\phi & \mathbf{0} \\ -\sin\phi & \cos\phi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{M_1-2\times M_1-2} \end{bmatrix}. \tag{67}$$

Let $\mathbf{X}$ be an M_1-by-M_1 matrix with its ith row denoted by $\mathbf{x}_i^H$. Apply the transform above to obtain $\tilde{\mathbf{X}} = \mathbf{\Phi}\mathbf{X}$ and denote its ith row by $\tilde{\mathbf{x}}_i^H$. Then, $\tilde{\mathbf{x}}_i^H = \mathbf{x}_i^H$ for $i = 3, 4, \ldots, M_1$, and

$$\tilde{\mathbf{x}}_1^H = \mathbf{x}_1^H \cos\phi + \mathbf{x}_2^H \sin\phi \tag{68}$$
$$\tilde{\mathbf{x}}_2^H = \mathbf{x}_2^H \cos\phi - \mathbf{x}_1^H \sin\phi. \tag{69}$$

Note that $\|\tilde{\mathbf{x}}_1\|^2$ can vary between $\|\mathbf{x}_1\|^2$ and $\|\mathbf{x}_2\|^2$ by suitably choosing ϕ. The norm square of $\|\tilde{\mathbf{x}}_1\|^2$ is then determined by $\|\tilde{\mathbf{x}}_2\|^2 = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 - \|\tilde{\mathbf{x}}_1\|^2$.

Based on the weight-redistributing property of the unitary transforms described above, we apply the following algorithm.

- *Initialization:*
  1) Set $\mathbf{X} = \mathrm{diag}[\sqrt{x_1}, \sqrt{x_2}, \ldots, \sqrt{x_{M_1}}]$.
  2) Set the first pointer $p_1 = 1$.
- *Loop:*
  1) If $\|\mathbf{x}_{p_1}\|^2 = \zeta$, go to Step 6.
  2) If $\|\mathbf{x}_{p_1}\|^2 > \zeta$, choose $p_2 > p_1$ such that $\|\mathbf{x}_{p_2}\|^2 < \zeta$, and go to Step 4.
  3) If $\|\mathbf{x}_{p_1}\|^2 < \zeta$, choose $p_2 > p_1$ such that $\|\mathbf{x}_{p_2}\|^2 > \zeta$.
  4) Obtain $\tilde{\mathbf{X}} = \mathbf{\Phi}\mathbf{X}$ by a suitably chosen unitary transform $\mathbf{\Phi}$ so that the p_1th row of $\tilde{\mathbf{X}}$ has norm square ζ.
  5) Set $\mathbf{X} = \tilde{\mathbf{X}}$.
  6) Set $p_1 = p_1 + 1$.
  7) If $p_1 \geq M_1$, end. Otherwise, go back to *Loop*.

Note that we can always find p_2 as described in the algorithm because $\sum_{i=1}^{M_1} x_i = M_1\zeta$. Since the transform applied at each step is unitary, it is easy to see that after $M_1 - 1$ iterations, the desired matrix $\mathbf{V}$ can be obtained as $\mathbf{X}\mathrm{diag}[1/\sqrt{x_1}, 1/\sqrt{x_2}, \ldots, 1/\sqrt{x_{M_1}}]$. ■

Appendix C
Proof of Proposition 3

First, we claim that

$$\frac{\sqrt{\lambda_{N-M_1+i}}}{\sum_{m=M^*+1}^{M_1} \sqrt{\lambda_{N-M_1+m}}} < \frac{1}{M_1 - M^* - M_1\zeta} \tag{70}$$

for $i = M^* + 1, \ldots, M_1$. Since λ_i, for $i = 1, \cdots, M$, are arranged in a descending order, we only need to show that (70) is true for $i = M^* + 1$. Indeed, if this were not true, we would have

$$0 < \frac{\sum_{m=M^*+2}^{M_1} \sqrt{\lambda_{N-M_1+m}}}{\sqrt{\lambda_{N-M_1+M^*+1}}} \leq M_1 - M^* - 1 - M_1\zeta \leq 0 \tag{71}$$

by the definition of M^*. Because of (70), the first statement in the proposition has to be true. Indeed, M_* is the smallest integer ($\leq M^*$) that (34) is true.

Now, let $\alpha_i = \beta_i/(1 + \beta_i)$ for $i = 1, \cdots, M_1$, and rewrite (30) as

$$\min \sum_{i=1}^{M_1} \frac{\lambda_{N-M_1+i}\alpha_i}{1 - \alpha_i}$$

subject to

$$\sum_{i=1}^{M_1} \alpha_i = M_1\zeta \qquad 0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{M_1} < 1. \tag{72}$$

Writing $f(\alpha) = \sum_{i=1}^{M_1} (\lambda_{N-M_1+i}\alpha_i)/(1 - \alpha_i)$, it is easy to see that $f(\alpha)$ is convex and continuously differentiable in the region in $\mathcal{R}^{M_1}$ defined by the constraints in (72). Based on this observation, the minimization problem in (72) can be solved by applying [18, p. 87, th. 4.4],[4] which essentially states that α' is a solution to (72) if and only if there is a μ such that

$$\frac{\partial f(\alpha')}{\partial \alpha_i} = \mu \qquad \text{for all } i \text{ that } \alpha_i' > 0 \tag{73}$$
$$\frac{\partial f(\alpha')}{\partial \alpha_i} \geq \mu \qquad \text{for all } i \text{ that } \alpha_i' = 0. \tag{74}$$

We claim that the following choices

$$\alpha_i' = \begin{cases} 0 & \text{if } 1 \leq i \leq M_* \\ 1 - (M_1 - M_* - M_1\zeta) \cdot \dfrac{\sqrt{\lambda_{N-M_1+i}}}{\sum_{m=M_*+1}^{M_1} \sqrt{\lambda_{N-M_1+m}}} & \text{if } M_* + 1 \leq i \leq M_1 \end{cases} \tag{75}$$

$$\mu = \left(\frac{1}{M_1 - M_* - M_1\zeta} \sum_{m=M_*+1}^{M_1} \sqrt{\lambda_{N-M_1+m}} \right)^2 \tag{76}$$

satisfy (73) and (74). Moreover, it is easy to see that the choice in (75) satisfies the constraints in (72). Hence, (75) is a solution to the minimization problem in (72). Transforming back to β_i, the second statement in the proposition that (35) is a solution to the original minimization problem in (30) is proved.

It remains to be shown that the two conditions (73) and (74) are indeed satisfied. To see this, notice that

$$\frac{\partial f(\alpha')}{\partial \alpha_i} = \frac{\lambda_{N-M_1+i}}{(1 - \alpha_i')^2} = \begin{cases} \lambda_{N-M_1+i}, & \text{if } 1 \leq i \leq M_* \\ \mu, & \text{if } M_* + 1 \leq i \leq M_1. \end{cases} \tag{77}$$

Hence, the proof will be completed if we can show that $\lambda_{N-M_1+i} \geq \mu$ for $i = 1, \ldots, M_*$. Again, it suffices to show that $\lambda_{N-M_1+M_*} \geq \mu$ or equivalently

$$\frac{\sum_{m=M_*+1}^{M_1} \sqrt{\lambda_{N-M_1+m}}}{\sqrt{\lambda_{N-M_1+M_*}}} \leq M_1 - M_* - M_1\zeta.$$

[4]This theorem is a special case of the well-known Karush–Kuhn–Tucker condition for constrained optimization problems [19].

This can be easily seen to be true by the very definition of M_* that

$$\frac{\sqrt{\lambda_{N-M_1+M_*}}}{\sum_{m=M_*}^{M_1}\sqrt{\lambda_{N-M_1+m}}} \geq \frac{1}{M_1 - M_* + 1 - M_1\zeta}.$$

APPENDIX D
PROOF OF PROPOSITION 4

Proposition 4 can be obtained as a corollary of the following result which is just a restatement of [20, th. 5.2.2] for our case.

Lemma 4: Suppose that $\mathbf{R}$ is an $N \times N$ Hermitian matrix with eigenvector $\mathbf{q}_1, \cdots, \mathbf{q}_N$ and associated eigenvalues $\lambda_1, \cdots, \lambda_N$ satisfying

$$\lambda_1 \geq \cdots \geq \lambda_{N-M_1} > \lambda_{N-M_1+1} \geq \lambda_{N-M_1+2} \geq \cdots \geq \lambda_N. \quad (78)$$

Let $\mathcal{V}$ be the subspace spanned by the first $N - M_1$ eigenvectors $\mathbf{q}_1, \ldots, \mathbf{q}_{N-M_1}$ and $\mathcal{U}$ be the subspace spanned by the last M_1 eigenvectors $\mathbf{q}_{N-M_1+1}, \ldots, \mathbf{q}_N$. If $\mathcal{A}$ is any M_1-dimensional subspace that intersects trivially with $\mathcal{V}$, then $\mathbf{R}^{-j}\mathcal{A}$ converges to $\mathcal{U}$.

Now, let us look back at the iteration in (36). From (27), we have

$$\mathbf{A}_{1\cdots M_1}[j+1] = \mathbf{R}_{1\cdots M_1}^{-1}\tilde{\mathbf{A}}_{1\cdots M_1}[j] \quad (79)$$

where

$$\tilde{\mathbf{A}}_{1\cdots M_1}[j] = \mathbf{A}_{1\cdots M_1}[j](\mathbf{A}_{1\cdots M_1}^H[j]\mathbf{R}_{1\cdots M_1}^{-1}\mathbf{A}_{1\cdots M_1}[j] + \mathbf{I})^{-1}\mathbf{G}[j]. \quad (80)$$

As long as $\mathbf{G}[j]$ is nonsingular, the column spaces of $\tilde{\mathbf{A}}_{1\cdots M_1}[j]$ and $\mathbf{A}_{1\cdots M_1}[j]$ are identical. Moreover, if $\mathbf{G}[j]$ is chosen to guarantee that the norm of $\tilde{\mathbf{A}}_{1\cdots M_1}[j]$ is bounded away from 0, then the column space of $\tilde{\mathbf{A}}_{1\cdots M_1}[j]$ will never shrink to the trivial null subspace. Therefore, we get the desired convergence result in Proposition 4 by simply applying the lemma stated.

APPENDIX E
PROOF OF PROPOSITION 5

Apply singular value decomposition (SVD) to $\mathbf{A}_T[j]$

$$\mathbf{A}_T[j] = \mathbf{U}\mathbf{S}[j]\mathbf{V}^H \quad (81)$$

where $\mathbf{U}$ is an $N \times N$ unitary matrix, $\mathbf{S}[j]$ is an $N \times M$ diagonal matrix with main diagonal containing the singular values, and $\mathbf{V}$ is an $M \times M$ unitary matrix. Then, the update in (43) is given by

$$\begin{aligned}\mathbf{A}_T[j+1] &= g(\mathbf{A}_T[j]\mathbf{A}_T^H[j] + \eta\mathbf{I})^{-1}\mathbf{A}_T[j]\\ &= (\mathbf{U}\mathbf{S}[j]\mathbf{V}^H\mathbf{V}\mathbf{S}^H[j]\mathbf{U}^H + \eta\mathbf{I})^{-1}\mathbf{U}\mathbf{S}[j]\mathbf{V}^H\\ &= \mathbf{U}\{g(\mathbf{S}[j]\mathbf{S}^H[j] + \eta\mathbf{I})^{-1}\mathbf{S}[j]\}\mathbf{V}^H. \end{aligned} \quad (82)$$

Notice that the result in (82) is the SVD of the updated matrix $\mathbf{A}_T[j+1]$. The kth singular value of $\mathbf{A}_T[j+1]$ is related to the kth singular value of $\mathbf{A}_T[j]$ by

$$s_k[j+1] = \frac{gs_k[j]}{s_k^2[j] + \eta}. \quad (83)$$

By Lemma 1, we see that all the singular values converge to the fixed point $\sqrt{g-\eta}$. Hence, $\mathbf{A}_T[j]$ converges to the limit $\mathbf{A}_\infty = \mathbf{U}\mathbf{S}_\infty\mathbf{V}^H$, where the diagonal elements of $\mathbf{S}_\infty$ are $\sqrt{g-\eta}$. Moreover, if we denote the mth diagonal element of $\mathbf{A}_T[j]\mathbf{R}_T^{-1}[j]\mathbf{A}_T[j]$ by $\zeta_m[j]$ and the SNR achieved at the jth iteration for the mth stream by $\gamma_m[j]$, then, similar to (15), we have

$$\zeta_m[j] = \frac{\gamma_m[j]}{1+\gamma_m[j]}. \quad (84)$$

Also, notice that $\mathbf{A}_T[j]\mathbf{R}_T^{-1}[j]\mathbf{A}_T[j]$ converges to $\mathbf{A}_\infty^H(\mathbf{A}_\infty\mathbf{A}_\infty^H + \eta\mathbf{I})^{-1}\mathbf{A}_\infty$.

First, consider the case of $M \leq N$. In this case

$$\mathbf{A}_\infty^H(\mathbf{A}_\infty\mathbf{A}_\infty^H + \eta\mathbf{I})^{-1}\mathbf{A}_\infty = \frac{g-\eta}{g}\mathbf{I}. \quad (85)$$

Hence, from (84), the achieved SNR for each data stream is the solution of

$$\frac{g-\eta}{g} = \frac{\gamma}{1+\gamma}. \quad (86)$$

Therefore, if g is chosen as in (44), the desired SNR can be achieved. Moreover, it is easy to verify that

$$\mathbf{A}_\infty^H\mathbf{A}_\infty = (g-\eta)\mathbf{I} \quad (87)$$

i.e., the vectors $\mathbf{a}_m$, for $m = 1, \ldots, M$, converges to the state where they are orthogonal. Comparing this with the form of the optimal solution given in (41), we see that $\mathbf{A}_\infty$ is, in fact, a solution to the optimization in (37).

Now, consider the case of $M > N$. In this case

$$\begin{aligned}&\mathbf{A}_\infty^H(\mathbf{A}_\infty\mathbf{A}_\infty^H + \eta\mathbf{I})^{-1}\mathbf{A}_\infty\\ &= \mathbf{V}\begin{bmatrix}\frac{g-\eta}{g}\mathbf{I}_{N\times N} & \\ & \mathbf{0}_{M-N\times M-N}\end{bmatrix}\mathbf{V}^H\\ &= \frac{N}{M}\frac{g-\eta}{g}\mathbf{I}_{M\times M}\end{aligned} \quad (88)$$

if all the conditions stated in the proposition are satisfied. Again, based on (84), the achieved SNR for each data stream is the solution of

$$\frac{N}{M}\frac{g-\eta}{g} = \frac{\gamma}{1+\gamma}. \quad (89)$$

Therefore, if g is chosen as in (46), the desired SNR can be achieved. To see that $\mathbf{A}_\infty$ is a solution to the optimization in (37), note that

$$\mathbf{A}_\infty\mathbf{A}_\infty^H = (g-\eta)\mathbf{I} \quad (90)$$

i.e., the rows of $\mathbf{A}_\infty$ are orthogonal. This is exactly the form specified in (42).

REFERENCES

[1] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[2] A. Duel-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser detection for CDMA systems," *IEEE Pers. Commun.*, vol. 2, pp. 46–58, Apr. 1995.

[3] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[4] B. Suard, A. Naguib, G. Xu, and A. Paulraj, "Performance analysis of CDMA mobile communication systems using antenna arrays," in *Proc. ICASSP*, vol. VI, Apr. 1993, pp. 153–156.

[5] T. F. Wong, T. M. Lok, J. S. Lehnert, and M. D. Zoltowski, "A linear receiver for direct-sequence spread-spectrum multiple-access systems with antenna arrays and blind adaptation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 659–676, Mar. 1998.

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

[7] 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.

[8] 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. 12, pp. 1332–1340, Sept. 1995.

[9] S. Ulukus and R. Yates, "Adaptive power control and mmse interference suppression," *Baltzer/ACM Wireless Networks*, vol. 4, no. 6, pp. 489–496, 1998.

[10] A. F. Almutairi, H. A. Latchman, T. F. Wong, and S. L. Miller, "MMSE based fully distributed power control algorithm," in *Proc. MILCOM*, Atlantic City, NJ, Nov. 1999.

[11] W. M. Jang, B. R. Vojcic, and R. L. Pickholtz, "Joint transmitter-receiver optimization in synchronous multiuser communications over multipath channels," *IEEE Trans. Commun.*, vol. 46, pp. 269–278, Feb. 1998.

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

[13] P. Viswanath, V. Anantharam, and D. Tse, "Optimal sequences, power control and capacity of spread-spectrum systems with multiuser linear receivers," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1968–1983, Sept. 1999.

[14] S. Ulukus and R. Yates, "Iterative signature adaptation for capacity maximization of CDMA systems," in *Proc. 36th Allerton Conf. Commun., Contr., and Comput.*, Monticello, IL, Sept. 1998.

[15] 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.

[16] G. Strang, *Linear Algebra and Its Applications*, 3rd ed. Orlando, FL: Harcourt Brace Jovanovich, 1988.

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

[18] R. G. Gallagher, *Information Theory and Reliable Commmunications*. New York: Wiley, 1968.

[19] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*. New York: Wiley, 1996.

[20] D. S. Watkins, *Fundamentals of Matrix Computations*. New York: Wiley, 1991.

Tan F. Wong received the B.Sc. degree (with first class honors) in electronic engineering from the Chinese University of Hong Kong in 1991 and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1992 and 1997, respectively.

He was a Research Engineer working on the high speed wireless networks project with the Department of Electronics at Macquarie University, Sydney, Australia. He also served as a Post-Doctoral Research Associate in the School of Electrical and Computer Engineering at Purdue. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Florida, Gainesville. His research interests include spread-spectrum communication systems, multiuser communications, and wireless cellular networks.

Tat M. Lok received the B.Sc. degree in electronic engineering from the Chinese University of Hong Kong in 1991. He continued his education at Purdue University, West Lafayette, IN, receiving the M.S.E.E. degree in electrical engineering and the Ph.D. degree in electrical and computer engineering in 1992 and 1995, respectively.

He had been a Research Assistant and a Post-Doctoral Research Associate at Purdue. Since 1996, he has been an Assistant Professor with the Department of Information Engineering, the Chinese University of Hong Kong. His research interests include code-division multiple-access systems, multiuser detection, adaptive antenna arrays, and communication theory.