

Iterative Construction of Optimum Signature Sequence Sets in Synchronous CDMA Systems

Sennur Ulukus, *Member, IEEE*, and Roy D. Yates, *Member, IEEE*

Abstract—Recently, optimum signature sequence sets that maximize the capacity of single-cell synchronous code division multiple access (CDMA) systems have been identified. Optimum signature sequences minimize the total squared correlation (TSC); they form a set of orthogonal sequences, if the number of users is less than or equal to the processing gain, and a set of Welch bound equality (WBE) sequences, otherwise. We present an algorithm where users update their transmitter signature sequences sequentially, in a distributed fashion, by using available receiver measurements. We show that each update decreases the TSC of the set, and produces better signature sequence sets progressively. We prove that the algorithm converges to a set of orthogonal signature sequences when the number of users is less than or equal to the processing gain. We observe and conjecture that the algorithm converges to a WBE set when the number of users is greater than the processing gain. At each step, the algorithm replaces one signature sequence from the set with the normalized minimum mean squared error (MMSE) receiver corresponding to that signature sequence. Since the MMSE filter can be obtained by a distributed algorithm for each user, the proposed algorithm is amenable to distributed implementation.

Index Terms—Code division multiple access (CDMA), distributed interference avoidance, minimum mean squared error (MMSE), optimum signature sequence sets, Welch bound equality (WBE) sequences.

I. INTRODUCTION

WE consider the uplink of a single-cell synchronous code division multiple access (CDMA) system with K users and processing gain N . In the presence of additive white Gaussian noise (AWGN) $n(t)$ with zero mean and power spectral density σ^2 , the received signal in one symbol interval is [3, Sec. 2.1]

$$r(t) = \sum_{i=1}^K \sqrt{p_i} b_i s_i(t) + n(t) \quad (1)$$

where, for user i , p_i is the received power, b_i is the information symbol with zero mean and variance $E[b_i^2] = 1$, and $s_i(t)$ is the signature waveform. The signature waveforms of the users, which are zero outside the symbol interval, have unit energies,

Manuscript received October 20, 1999; revised October 13, 2000. This work was supported by NSF Grant NCR 95-06505. This work contains material from S. Ulukus' Ph.D. dissertation and was presented in part at the 36th Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, September 1998.

S. Ulukus was with WINLAB, Rutgers–State University of New Jersey, Piscataway, NJ 08854-8060. She is now with AT&T Labs-Research, Florham Park, NJ 07932 USA (e-mail: ulukus@research.att.com).

R. D. Yates is with WINLAB, Rutgers–State University of New Jersey, Piscataway, NJ 08854-8060 USA (e-mail: ryates@winlab.rutgers.edu).

Communicated by M. L. Honig, Associate Editor for Communications.

Publisher Item Identifier S 0018-9448(01)04483-2.

and can be represented by N orthonormal basis waveforms [3, Sec. 2.3.6] $\{\psi_j(t)\}_{j=1}^N$ such that

$$s_i(t) = \sum_{j=1}^N s_{ij} \psi_j(t)$$

where $s_{ij} = \langle s_i(t), \psi_j(t) \rangle$. In standard CDMA notation, $\{\psi_j(t)\}_{j=1}^N$ could be chosen as the shifted versions of the basic chip waveform by multiples of the chip duration [3, Secs. 2.3.6 and 2.9.2]. Projecting the received signal onto these basis waveforms yields a set of sufficient statistics $\{r_j\}_{j=1}^N$ where $r_j = \langle r(t), \psi_j(t) \rangle$ [3, Sec. 2.9.2]. By defining the signature sequence of user i as $\mathbf{s}_i = [s_{i1}, \dots, s_{iN}]^T$ and the received signal vector $\mathbf{r} = [r_1, \dots, r_N]^T$, we can write (1) in the equivalent vector notation

$$\mathbf{r} = \sum_{i=1}^K \sqrt{p_i} b_i \mathbf{s}_i + \mathbf{n}. \quad (2)$$

Note that \mathbf{n} is a zero mean Gaussian random vector with $E[\mathbf{n}\mathbf{n}^T] = \sigma^2 \mathbf{I}_N$, where \mathbf{I}_N denotes the $N \times N$ identity matrix.

The information-theoretic capacity region of a single-cell synchronous CDMA system was derived in [4] (see also [5]). An important measure of overall information capacity of a multiaccess channel is the sum capacity [4]

$$C_{\text{sum}} = \frac{1}{2} \log[\det(\mathbf{I}_N + \sigma^{-2} \mathbf{S} \mathbf{P} \mathbf{S}^T)] \quad (3)$$

where $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_K]$ is an $N \times K$ matrix with the users' signature sequences as its columns and $\mathbf{P} = \text{diag}\{p_1, \dots, p_K\}$ is a $K \times K$ diagonal matrix of the users' received powers. Note that it will be convenient to use the matrix \mathbf{S} as a notation for the set of column vectors of \mathbf{S} . For example, we may write $\mathbf{s} \in \mathbf{S}$ when \mathbf{s} is a column of \mathbf{S} .

When the received powers of the users are the same, $p_i = p$ for all i , (3) reduces to

$$\begin{aligned} C_{\text{sum}} &= \frac{1}{2} \log \left[\det \left(\mathbf{I}_N + \frac{p}{\sigma^2} \mathbf{S} \mathbf{S}^T \right) \right] \\ &= \frac{1}{2} \log \left[\det \left(\mathbf{I}_K + \frac{p}{\sigma^2} \mathbf{S}^T \mathbf{S} \right) \right] \end{aligned} \quad (4)$$

where the equivalence of the two expressions in (4) follows from the equality

$$\det(\mathbf{I}_L + \mathbf{A} \mathbf{B}) = \det(\mathbf{I}_M + \mathbf{B} \mathbf{A})$$

for any two matrices $\mathbf{A}_{L \times M}$ and $\mathbf{B}_{M \times L}$.

For arbitrary (unequal) powers \mathbf{P} , the optimum signature sequence sets that maximize the sum capacity in (3) have recently been identified in [6]. For the case of equal powers, [7] shows

that the sum capacity is maximized if the signature sequences are chosen such that, if $K \leq N$

$$\mathbf{S}^T \mathbf{S} = \mathbf{I}_K \quad (5)$$

and if $K > N$

$$\mathbf{S}\mathbf{S}^T = \frac{K}{N} \mathbf{I}_N. \quad (6)$$

The signature sequence sets satisfying (5) contain K orthonormal signature sequences in N -dimensional vector space. The sequence sets satisfying (6) are named Welch bound equality (WBE) sequence sets in [7], because they satisfy the Welch bound on the sum of the squares of the cross correlations of unit energy sequences with equality. Note that there are infinitely many sets of sequences satisfying (5) or (6) in the continuous \mathcal{R}^N space.

In [8], the *user* capacity of a single-cell synchronous CDMA system is defined as the maximum number of admissible users, given the processing gain N , and a common signal to interference ratio (SIR) target β ; K users are said to be admissible if there exist positive powers p_i and signature sequences \mathbf{s}_i such that all users have SIRs that are at least as large as the target SIR β . The user capacity of a CDMA system was found for two kinds of linear receiver structures in [8]: matched filters and minimum mean squared error (MMSE) filters [9], [10]. In both cases, the user capacity of a single-cell synchronous CDMA system was found to be

$$K < N \left(1 + \frac{1}{\beta} \right). \quad (7)$$

It was shown in [8] that the user capacity is maximized if the signature sequence set is chosen to satisfy (5) if $K \leq N$ and (6) if $K > N$, and if the received powers of the users are chosen to be the same for all users, for both MMSE and matched-filter receivers cases. This is expected since when signature sequences satisfy (5), or when they satisfy (6) and $p_i = p$ for all i , the MMSE receivers reduce to matched filters [8]. Therefore, an important consequence of [8] is that once the optimum signature sequences are used, the optimum linear receiver filter selection is simplified: matched filters are optimum linear receiver filters.

The aim of this paper is to develop a simple iterative and distributed algorithm that uses receiver measurements to adapt transmitter signatures. We describe an algorithm that converges to an optimum signature sequence set (an orthonormal set for $K \leq N$ and a WBE set for $K > N$) for arbitrary K and N [1], [2]. When $K \leq N$, a simple Gram-Schmidt orthogonalization procedure can be used to obtain K orthonormal vectors starting with K linearly independent vectors, yielding an optimum signature sequence set. Unfortunately, this does not have a direct generalization to the case $K > N$.

The idea that the transmitter signature sequences may be changed according to multiaccess interference (MAI) conditions has also been presented in [11]. Reference [11] studies the problem of finding jointly optimum transmitter signature sequence and the receiver filter of a *single* user in a system where the powers of all users and the signature sequences of all interfering users are assumed to be fixed. The mean squared error (MSE) of this user is taken to be the optimality criterion.

Reference [11] assumes that the receiver filter of the desired user at the base station is updated to be the MMSE filter corresponding to the current signature sequence of the user instantaneously, and presents an algorithm where the signature sequence of the desired user is updated to be the normalized MMSE receiver filter given the signature sequences and the powers of all of the interfering users. This signature sequence update combined with the simultaneous update of the receiver filter to the corresponding MMSE receiver is shown to decrease the MSE of the desired user. The significant difference between the approach of [11] and this paper is that we consider the joint optimization of all users' signature sequences and receiver filters with the intent of finding signature sequences that are an optimum *ensemble* for all users. We note, however, that although they address fundamentally different problems, a single signature sequence update of [11] and that of this paper are essentially the same.

A related topic to signature sequence design is transmitter precoding [12]. In transmitter precoding, the transmit signals intended for multiple users are passed through a linear transformation at the common transmitter (base station) to minimize the MAI at the remote receivers (users). Transmitter precoding assumes that the symbols of the users on the downlink are already modulated with some linearly independent signature sequences, and essentially undoes the cross correlation introduced by the nonorthogonality of these signature sequences. The optimum linear transformation was found to be the inverse of the cross correlation matrix in [12]. Intrinsically, [12] assumes that the number of users is less than the processing gain, so that this cross-correlation matrix is invertible. Receiver optimization in conjunction with transmitter precoding was studied in [13].

It is worth mentioning that transmitter beamforming in systems with multielement transmit antenna arrays [14]–[18] is also related to signature sequence design. In transmit beamforming, the transmitter signal out of the multielement antenna array is processed *spatially* to decrease the MAI at the receiver sites. A fundamental difference between transmit beamforming in multielement antenna systems and signature sequence design in CDMA systems is that the signature sequences of the users in CDMA systems are created and therefore can be fully controlled by the transmitter. The spatial signatures in multielement antenna systems, on the other hand, are created by the wireless communication channel and cannot be directly controlled by the transmitters. The transmitters can control the transmit beamforming weights, though, to improve the quality of communication metrics such as the MSE or the SIR at the receiver. References [14], [15] study the transmit beamforming problem, whereas [16]–[18] study transmit beamforming in combination with transmit power control.

In the following section, first we relate the Welch bound to the optimum signature sequence set design problem. In Section III, we will make some simple observations which will be the basis of the iterative algorithm that will be given in Section IV.

II. WELCH BOUND AND THE TOTAL SQUARED CORRELATION

Welch developed lower bounds for the $2m$ th power of the maximum correlation among a set of K unit energy vectors

[19]. These lower bounds were actually obtained from the lower bound

$$\sum_{i=1}^K \sum_{j=1}^K (\mathbf{s}_i^\top \mathbf{s}_j)^{2m} \geq \frac{K^2}{\binom{N+m-1}{m}} \quad (8)$$

on the sum of the $2m$ th powers of the correlations of a sequence set. When $m = 1$, the Welch bound (8) reduces to the following bound on the TSC:

$$\text{TSC} = \sum_{i=1}^K \sum_{j=1}^K (\mathbf{s}_i^\top \mathbf{s}_j)^2 \geq \frac{K^2}{N}. \quad (9)$$

For a simple derivation of the bound (9), see [20], [21]. We note that the Welch bound (9) is loose for $K < N$. When $K < N$, the minimum value $\text{TSC} = K$ is achieved by K orthonormal vectors and the Welch lower bound K^2/N is not achievable. On the other hand, when $K > N$, the Welch bound can be achieved and the sets of unit energy vectors satisfying (6) are precisely those achieving the bound [20], [21]. For general K and N , we can conclude that optimum signature sequence sets, the signature sequence sets satisfying (5) for $K \leq N$ and (6) for $K > N$, are those that achieve the minimum TSC.

The TSC can also be related to the total MSE in the system. Assume that the signature sequences of the users are \mathbf{s}_i , $i = 1, \dots, K$, the signals of the users at the base station are received by matched filters, and $p_i = p$. In this case, from the received signal \mathbf{r} given in (2), the MSE for the i th user is

$$\begin{aligned} \text{MSE}_i &= E [(\mathbf{r}^\top \mathbf{s}_i - b_i)^2] \quad (10) \\ &= \mathbf{s}_i^\top \left(p \sum_{j=1}^K \mathbf{s}_j \mathbf{s}_j^\top + \sigma^2 \mathbf{I}_N \right) \mathbf{s}_i - 2\sqrt{p} \mathbf{s}_i^\top \mathbf{s}_i + 1. \quad (11) \end{aligned}$$

The total MSE in the system is

$$\begin{aligned} \text{MSE} &= \sum_{i=1}^K \text{MSE}_i \quad (12) \\ &= p \sum_{i=1}^K \sum_{j=1}^K (\mathbf{s}_i^\top \mathbf{s}_j)^2 - (2\sqrt{p} - \sigma^2) \sum_{i=1}^K \mathbf{s}_i^\top \mathbf{s}_i + K. \quad (13) \end{aligned}$$

We observe that the first term in (13) is the TSC. Since the signature sequences are restricted to be of unit energy, we see that

$$\text{MSE} = p\text{TSC} + (1 + \sigma^2 - 2\sqrt{p})K. \quad (14)$$

Therefore, minimizing TSC subject to $\mathbf{s}_i^\top \mathbf{s}_i = 1$ is equivalent to minimizing MSE subject to the same constraint, and the sequence sets satisfying (5) for $K \leq N$, and (6) for $K > N$ minimize the MSE as well. In other words, orthogonal sequences for $K \leq N$ and WBE sequences for $K > N$ are the *global optimal* solutions of the following two equivalent problems:

$$\begin{aligned} \text{P1} \quad & \min \quad \text{TSC} \\ & \text{s. t.} \quad \mathbf{s}_i^\top \mathbf{s}_i = 1, \quad i = 1, \dots, K \\ \text{P2} \quad & \min \quad \text{MSE} \\ & \text{s. t.} \quad \mathbf{s}_i^\top \mathbf{s}_i = 1, \quad i = 1, \dots, K. \quad (15) \end{aligned}$$

When the TSC reaches its minimum value, we will have reached the optimum signature sequence set.

In this paper, we will construct an iterative algorithm to minimize TSC. The algorithm will update the signature sequences of the users sequentially, updating only one signature sequence at a time, in a way that is guaranteed to decrease (more precisely, not to increase) the TSC after every update. Although one can find many possible signature sequence update mechanisms for TSC reduction, we will concentrate on a particular method that we call the *MMSE update*. In the MMSE update, one of the K signature sequences in the set is replaced with the normalized MMSE receiver filter corresponding to that signature sequence. We will show that the set after the update will have a lower TSC compared to the set before the update. We focus on the MMSE update because the MMSE receiver filter is particularly well studied and understood in the context of multiuser signal detection [3]. Although the MMSE update algorithm that we propose and analyze is deterministic, it is motivated by the possibility of adaptive [10] and even blind [22] implementations of the MMSE update.

III. TSC REDUCTION

Given a set of signatures represented by the columns of the matrix $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_K]$, we first separate the TSC terms that depend on a particular signature sequence \mathbf{s}_k by writing

$$\text{TSC} = (\mathbf{s}_k^\top \mathbf{s}_k)^2 + 2\mathbf{s}_k^\top \left(\sum_{j \neq k} \mathbf{s}_j \mathbf{s}_j^\top \right) \mathbf{s}_k + \gamma_k \quad (16)$$

where

$$\gamma_k = \sum_{i \neq k} \sum_{j \neq k} (\mathbf{s}_i^\top \mathbf{s}_j)^2 \quad (17)$$

represents the squared correlation terms that do not depend on \mathbf{s}_k . Since we will always restrict ourselves to unit energy signature sequences, i.e., $\mathbf{s}_k^\top \mathbf{s}_k = 1$, we can add $2a^2 \mathbf{s}_k^\top \mathbf{s}_k$ and subtract $2a^2$ in (16) to obtain

$$\text{TSC} = 2\mathbf{s}_k^\top \mathbf{A}_k \mathbf{s}_k + \gamma_k + 1 - 2a^2 \quad (18)$$

where

$$\mathbf{A}_k = \sum_{j \neq k} \mathbf{s}_j \mathbf{s}_j^\top + a^2 \mathbf{I}_N. \quad (19)$$

Let us replace the signature sequence of user k with the unit energy vector

$$\mathbf{c}_k = \frac{\mathbf{A}_k^{-1} \mathbf{s}_k}{(\mathbf{s}_k^\top \mathbf{A}_k^{-2} \mathbf{s}_k)^{1/2}}. \quad (20)$$

This maps the set of signature sequences \mathbf{S} to a new set of signatures

$$\bar{\mathbf{S}} = [\mathbf{s}_1 \quad \dots \quad \mathbf{s}_{k-1} \quad \mathbf{c}_k \quad \mathbf{s}_{k+1} \quad \dots \quad \mathbf{s}_K]. \quad (21)$$

Note that \mathbf{c}_k is the normalized MMSE filter for user k [9], [10]. Also note that this MMSE filter is a generalized one; \mathbf{c}_k is the normalized MMSE filter for user k in a CDMA system where all other users have unit received power and the variance of the AWGN is a^2 . It will be apparent in what follows that any generalized normalized MMSE filter will be as good as any other in terms of constructing optimum sequences. The MMSE

filter coefficients in (20) can be obtained using an adaptive [10], [23]–[25] or a blind [22] algorithm.

The motivation for replacing \mathbf{s}_k with \mathbf{c}_k follows from the objective of the MMSE filter. Compared to the matched filter, the MMSE filter is designed to have lower cross correlation with the signature sequences of the other users, thereby suppressing the interference. Note that by the nature of the mapping in (20), $\mathbf{c}_k^\top \mathbf{c}_k = \mathbf{s}_k^\top \mathbf{s}_k = 1$. Thus, from (18) and (20), the modified signature set $\overline{\mathbf{S}}$ has total squared correlation

$$\overline{\text{TSC}} = 2 \frac{\mathbf{s}_k^\top \mathbf{A}_k^{-1} \mathbf{s}_k}{\mathbf{s}_k^\top \mathbf{A}_k^{-2} \mathbf{s}_k} + \gamma_k + 1 - 2a^2. \quad (22)$$

The following theorem verifies that replacing a particular sequence with its normalized MMSE receiver cannot increase the total squared correlation of the set. In other words, in terms of the TSC (or MSE) criteria, $\overline{\mathbf{S}}$ is a “better” set of signature sequences than \mathbf{S} .

Theorem 1: The MMSE signature update $\mathbf{s}_k \rightarrow \mathbf{c}_k$ reduces the total squared correlation: $\overline{\text{TSC}} \leq \text{TSC}$. The equality $\overline{\text{TSC}} = \text{TSC}$ occurs iff $\mathbf{c}_k = \mathbf{s}_k$.

Proof of Theorem 1, as well as proofs of all subsequent theorems and lemmas, can be found in the Appendix. It was shown in [11] that \mathbf{c}_k from (20) with $a^2 = \sigma^2$ decreased the MSE of user k at the output of the MMSE filter receiver corresponding to transmit signature \mathbf{c}_k . In fact, it is also true that \mathbf{c}_k decreases the MSE of user k with a matched-filter receiver in place, as defined by (11). It is interesting to note, however, that the MMSE update may increase or decrease the MSE of any other individual user, while being guaranteed to decrease the total MSE of the system by Theorem 1 and (14).

The following theorem states another significant property of the MMSE update.

Theorem 2: For any $x \geq 0$

$$\begin{aligned} \det(x\mathbf{I}_N + \overline{\mathbf{S}}\overline{\mathbf{S}}^\top) &\geq \det(x\mathbf{I}_N + \mathbf{S}\mathbf{S}^\top) \\ \det(x\mathbf{I}_K + \overline{\mathbf{S}}^\top \overline{\mathbf{S}}) &\geq \det(x\mathbf{I}_K + \mathbf{S}^\top \mathbf{S}). \end{aligned}$$

A simple consequence of Theorem 2 is the following.

Corollary 1: The MMSE signature update $\mathbf{s}_k \rightarrow \mathbf{c}_k$ increases the sum capacity.

Corollary 1 readily follows from Theorem 2 by choosing $x = \sigma^2/p$ and the monotonicity of the log function. Therefore, we observe from Theorem 1 and Corollary 1 that the MMSE update, when executed by the users sequentially, decreases the TSC and the MSE monotonically and increases C_{sum} monotonically. In other words, the signature sequence set after each update is a “better” set not only in terms of the TSC and the MSE, but in terms of the sum capacity as well.

IV. CONVERGENCE OF AN ITERATIVE ALGORITHM

We have observed in the previous section that given a set of unit energy vectors, if any one of these vectors is replaced with the corresponding normalized MMSE vector, then the TSC of the set decreases. We start with K unit length vectors

$\mathbf{S}(0) = [\mathbf{s}_1(0), \dots, \mathbf{s}_K(0)]$ at time 0. At iteration $n + 1$, the algorithm replaces the vectors $\mathbf{S}(n) = [\mathbf{s}_1(n), \dots, \mathbf{s}_K(n)]$ with their corresponding normalized MMSE filters one by one, and yields

$$\mathbf{S}(n + 1) = [\mathbf{s}_1(n + 1), \dots, \mathbf{s}_K(n + 1)].$$

A complete iteration includes K intermediate steps. At the k th intermediate step in iteration $n + 1$, the first $k - 1$ vectors have already been updated and the current vector set is

$$\mathbf{S}_{k-1}(n + 1) = [\mathbf{s}_1(n + 1) \quad \cdots \quad \mathbf{s}_{k-1}(n + 1) \\ \mathbf{s}_k(n) \quad \mathbf{s}_{k+1}(n) \quad \cdots \quad \mathbf{s}_K(n)]. \quad (23)$$

We update the signature \mathbf{s}_k employing the matrix

$$\mathbf{A}_k(n + 1) = \sum_{j < k} \mathbf{s}_j(n + 1)\mathbf{s}_j^\top(n + 1) + \sum_{j > k} \mathbf{s}_j(n)\mathbf{s}_j^\top(n) + a^2\mathbf{I}_N. \quad (24)$$

In particular, the k th signature is updated according to

$$\mathbf{s}_k(n + 1) = \frac{\mathbf{A}_k(n + 1)^{-1} \mathbf{s}_k(n)}{[\mathbf{s}_k^\top(n)\mathbf{A}_k(n + 1)^{-2} \mathbf{s}_k(n)]^{1/2}} \quad (25)$$

to yield the vector set

$$\mathbf{S}_k(n + 1) = [\mathbf{s}_1(n + 1) \quad \cdots \quad \mathbf{s}_{k-1}(n + 1) \\ \mathbf{s}_k(n + 1) \quad \mathbf{s}_{k+1}(n) \quad \cdots \quad \mathbf{s}_K(n)]. \quad (26)$$

To examine the convergence of this proposed algorithm, let $\text{TSC}_k(n)$ denote the TSC of the set $\mathbf{S}_k(n)$ after the k th intermediate step in iteration n . In addition, let $\text{TSC}(n)$ denote the TSC at the end of iteration n for set $\mathbf{S}(n)$. As a consequence of Theorem 1, we have for $k = 0, \dots, K - 1$

$$\text{TSC}_{k+1}(n) \leq \text{TSC}_k(n). \quad (27)$$

By recursive application of (27) we have

$$\text{TSC}(n + 1) \leq \text{TSC}(n). \quad (28)$$

For any K and N , TSC is lower-bounded by the Welch bound K^2/N . Therefore, the monotonically decreasing sequence $\text{TSC}(n)$ converges. This implies that at the fixed point where $\text{TSC}(n + 1) = \text{TSC}(n)$ we must have

$$\begin{aligned} \text{TSC}(n + 1) &= \text{TSC}_K(n + 1) = \text{TSC}_{K-1}(n + 1) = \cdots \\ &= \text{TSC}_1(n + 1) = \text{TSC}(n). \end{aligned} \quad (29)$$

From Theorem 1, this occurs iff $\mathbf{s}_k(n + 1) = \mathbf{s}_k(n)$ for all k . Thus, at the fixed point we must have

$$\mathbf{S}(n + 1) = \mathbf{S}_k(n + 1) = \mathbf{S}(n)$$

for all k .

Let us denote the fixed-point set of vectors by the matrix $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_K]$. For evaluation of the fixed point, it will prove convenient to write the MMSE update of (20) in the form

$$\mathbf{c}_k = \frac{\mathbf{B}^{-1} \mathbf{s}_k}{(\mathbf{s}_k^\top \mathbf{B}^{-2} \mathbf{s}_k)^{1/2}} \quad (30)$$

where

$$\mathbf{B} = \sum_{j=1}^K \mathbf{s}_j \mathbf{s}_j^\top + a^2 \mathbf{I}_N = \mathbf{S}\mathbf{S}^\top + a^2 \mathbf{I}_N. \quad (31)$$

Therefore, the fixed point \mathbf{S} satisfies $\mathbf{B}^{-1}\mathbf{s}_i = \kappa_i\mathbf{s}_i$ or, equivalently, $\mathbf{B}\mathbf{s}_i = \mathbf{s}_i/\kappa_i$. It follows from (31) that

$$\mathbf{S}\mathbf{S}^\top\mathbf{s}_i = \phi_i\mathbf{s}_i, \quad i = 1, \dots, K. \quad (32)$$

That is, at the fixed point, each \mathbf{s}_i is an eigenvector of $\mathbf{S}\mathbf{S}^\top$. In the following, we will investigate the properties of the fixed point for the cases $K \leq N$ and $K > N$ separately. We will need the following claim.

Corollary 2: If the signature \mathbf{s}_k is replaced by \mathbf{c}_k , at step k of iteration n , then

- a) $\det(\mathbf{S}_{k+1}^\top(n)\mathbf{S}_{k+1}(n)) \geq \det(\mathbf{S}_k^\top(n)\mathbf{S}_k(n))$
- b) $\det(\mathbf{S}_{k+1}(n)\mathbf{S}_{k+1}^\top(n)) \geq \det(\mathbf{S}_k(n)\mathbf{S}_k^\top(n))$.

Corollary 2 readily follows from Theorem 2 by choosing $x = 0$. Corollary 2 a) implies that if $\det(\mathbf{S}^\top(0)\mathbf{S}(0)) > 0$, then at the fixed point, $\det(\mathbf{S}^\top\mathbf{S}) > 0$. Similarly, Corollary 2 b) says that if $\det(\mathbf{S}(0)\mathbf{S}^\top(0)) > 0$, then at the fixed point, $\det(\mathbf{S}\mathbf{S}^\top) > 0$. Although these results are valid whether $K \leq N$ or $K > N$, their utility depends on K and N . The following corollaries which are immediate consequences of Corollary 2 will be useful in the discussion of the convergence of the MMSE algorithm.

Corollary 3: If $K \leq N$ and $\mathbf{S}^\top(0)\mathbf{S}(0)$ has rank K , then for all n , $\mathbf{S}^\top(n)\mathbf{S}(n)$ has rank K , and at the fixed point, $\mathbf{S}^\top\mathbf{S}$ has rank K .

Corollary 4: If $K > N$ and $\mathbf{S}(0)\mathbf{S}^\top(0)$ has rank N , then for all n , $\mathbf{S}(n)\mathbf{S}^\top(n)$ has rank N , and at the fixed point, $\mathbf{S}\mathbf{S}^\top$ has rank N .

Corollaries 3 and 4 ensure that if $\mathbf{S}(0)$ is full rank, then the fixed point \mathbf{S} is full rank. For $K \leq N$, this leads to the following result.

Theorem 3: If $K \leq N$ and $\mathbf{S}(0)$ has rank K , then at the fixed point

$$\mathbf{S}^\top\mathbf{S} = \mathbf{I}_K. \quad (33)$$

Theorem 3 simply says that when $K \leq N$, the signature sequences converge to an orthonormal set. The only condition for this to happen is that the algorithm must be started with a full rank signature sequence set.

We also see in Theorem 3 that the MMSE update algorithm looks to move the users' signatures toward orthogonal subspaces. We will say that \mathbf{S} is *partitioned into orthogonal subsets* $\mathbf{S}_1, \dots, \mathbf{S}_L$, or simply *partitioned*, if we can order the signatures of \mathbf{S} so as to write $\mathbf{S} = [\mathbf{S}_1|\mathbf{S}_2|\dots|\mathbf{S}_L]$ where for $i \neq j$, $\mathbf{S}_i^\top\mathbf{S}_j = \mathbf{0}$. In the next claim, we observe that the MMSE algorithm preserves a partition of the signature set.

Lemma 1: If $\mathbf{S}(0)$ is partitioned into $\mathbf{S}_1(0), \dots, \mathbf{S}_L(0)$, then, for all n , $\mathbf{S}(n)$ is partitioned as $[\mathbf{S}_1(n)|\dots|\mathbf{S}_L(n)]$ such that for each l , $\text{span}\mathbf{S}_l(n) = \text{span}\mathbf{S}_l(0)$.

For $K \leq N$, the property that the MMSE update preserves partitions plays no role since the eventual fixed point has every signature in an orthogonal subspace. When $K > N$, the preservation of partitions complicates the description of the fixed points. For $K > N$, we now characterize the fixed point

\mathbf{S} of the MMSE algorithm. We will assume in the following discussion that $\mathbf{S}(0)$ has rank N , so that by Corollary 4, the fixed point \mathbf{S} has rank N .

From (32), ϕ_i is the eigenvalue of the eigenvector \mathbf{s}_i of $\mathbf{S}\mathbf{S}^\top$. Since $\mathbf{S}\mathbf{S}^\top$ is real and symmetric, $\phi_i \neq \phi_j$ implies \mathbf{s}_i and \mathbf{s}_j are orthogonal. That is, the distinct eigenvalues of $\mathbf{S}\mathbf{S}^\top$ define an orthogonal partition of \mathbf{S} . Since $\mathbf{S}\mathbf{S}^\top$ is $N \times N$, we let $\{\omega_l\}_{l=1}^L$ denote its $L \leq N$ distinct eigenvalues. We use \mathbf{S}_l to denote the $N \times K_l$ matrix whose K_l columns are the signature sequences with common eigenvalue ω_l . Since each user must belong to one subset, we have $\sum_{l=1}^L K_l = K$. Let N_l denote the rank of \mathbf{S}_l . Since we assume that $\mathbf{S}(0)$ is full rank, \mathbf{S} must be full rank as well, and we must have $\sum_{l=1}^L N_l = N$.

Theorem 4: For each l , the matrix $\mathbf{S}_l\mathbf{S}_l^\top$ has eigenvalues of K_l/N_l and 0 with multiplicities of N_l and $N - N_l$, respectively. The signature sequences in \mathbf{S}_l constitute a WBE set, in the sense that, for K_l signatures occupying N_l signal space dimensions, they achieve the Welch bound K_l^2/N_l on TSC with equality.

Theorem 4 says that, when $K > N$, the properties of the convergence point signature sequence set depend on whether they are partitioned into orthogonal subsets, and the condition that the initial signature sequence set must have full rank is necessary but not sufficient to ensure convergence to a WBE set. A simple consequence of Theorem 4 is that if the convergence point signature sequence set does not include any orthogonal subsets, then the algorithm has converged to a WBE set, as stated in the following corollary.

Corollary 5: If $K > N$, and the fixed point \mathbf{S} is not partitioned, then $\mathbf{S}\mathbf{S}^\top = (K/N)\mathbf{I}_N$.

Corollary 5 follows from the fact that an $N \times N$ symmetric matrix with all of its N eigenvalues equal to K/N must be $(K/N)\mathbf{I}_N$.

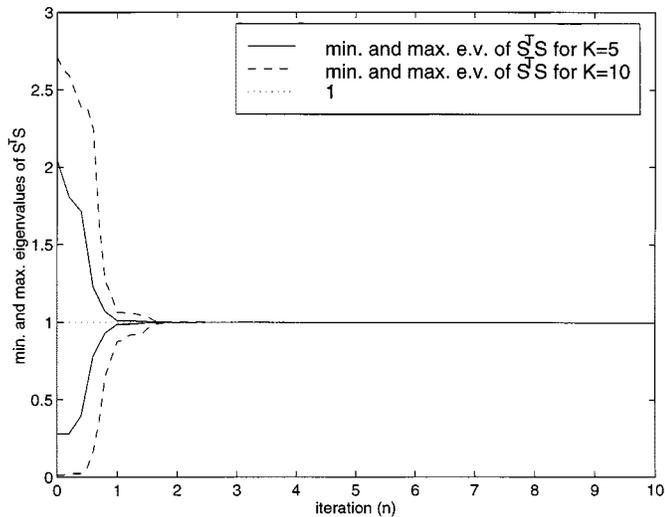
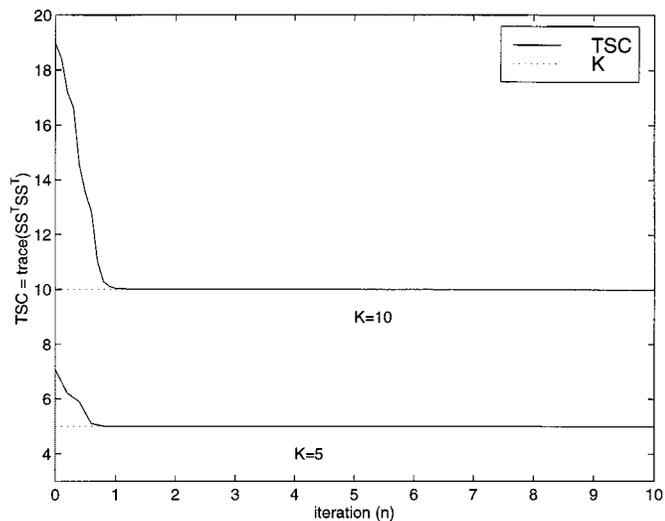
If the fixed-point signature sequence set contains orthogonal partitions, then the algorithm might have converged to a suboptimum point. That is, an orthogonally partitioned set, in general, has a larger TSC than a WBE set. As stated in the following lemma, when $K > N$, an orthogonally partitioned signature sequence set is strictly suboptimal, unless the particular orthogonal partition happens to satisfy the "equal loading" condition, i.e., the number of users in each subset is proportional to the dimensionality of the subset.

Lemma 2: If the convergence point signature sequence set \mathbf{S} is partitioned into L orthogonal subsets with K_l users and rank N_l for subset l , then the TSC of this set is larger than or equal to K^2/N , the TSC achievable with a WBE set. The TSC equals K^2/N iff $K_l/N_l = K/N$ for all l .

V. CONJECTURE, OBSERVATIONS, AND CONCLUSION

Because the MMSE algorithm is completely deterministic, whether the convergence point signature sequence set contains orthogonal partitions or not depends only on the initial signature sequence set, $\mathbf{S}(0)$.

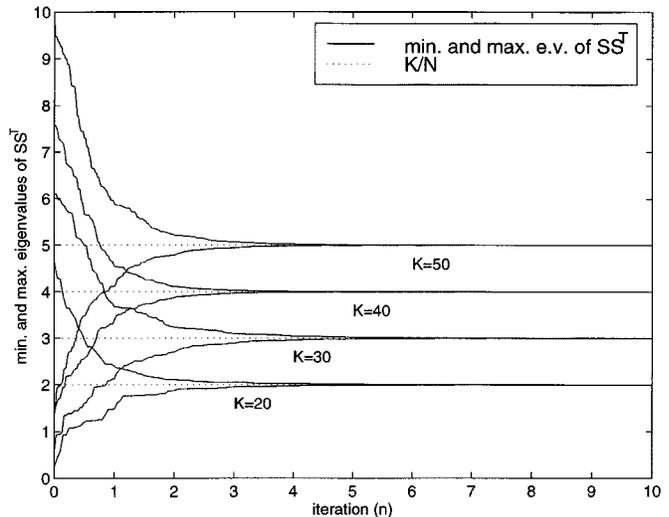
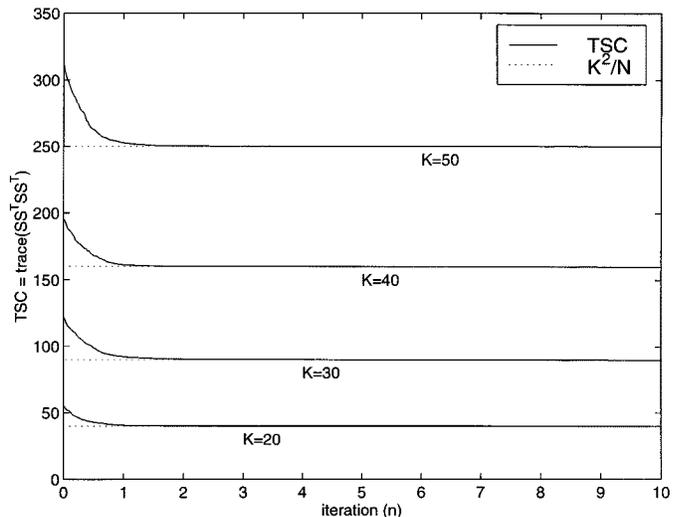
In this section, we describe our experiments with MMSE signature update. We will see that our experimental evidence encourages us to believe that if the initial signature sequence

Fig. 1. Minimum and maximum eigenvalues of matrix $\mathbf{S}^T(n)\mathbf{S}(n)$.Fig. 2. Total squared correlation, $\text{TSC}(n)$.

set $\mathbf{S}(0)$ is not partitioned then the signature sequence set will never, including the fixed point, be partitioned. Although we have observed this through a substantial number of experiments, we have not been able to prove it formally. Therefore, we make the conjecture below.

Conjecture 1: For $K > N$, $\mathbf{S}(n)$ for all n , and the fixed point $\mathbf{S} = \lim_{n \rightarrow \infty} \mathbf{S}(n)$, is partitioned only if $\mathbf{S}(0)$ is partitioned.

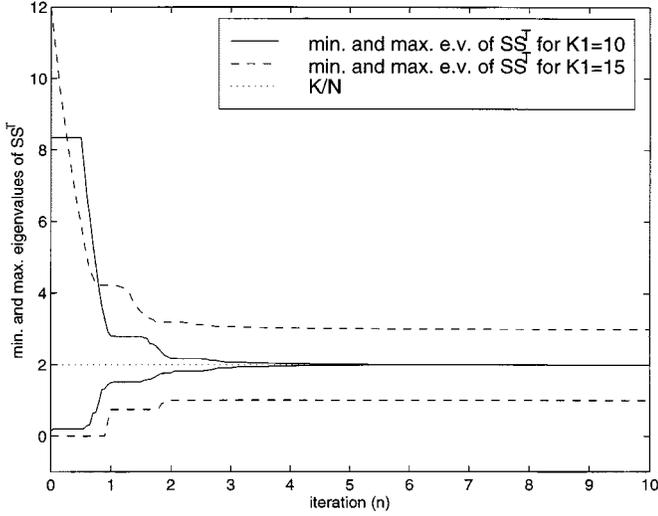
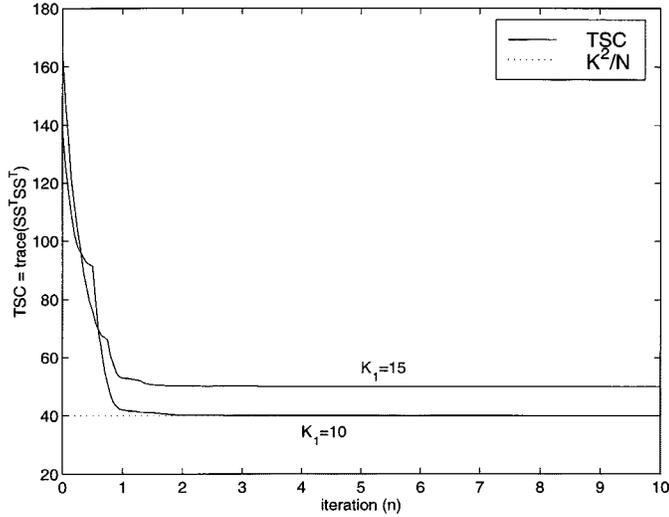
If the above conjecture is true, the fact that we have never observed orthogonal partitioning of the signature set at the fixed point in our experiments when we started our algorithm with randomly generated signature sequences in \mathcal{R}^N is understandable. Because, when we view the initial set of signature sequences as a set of continuous random vectors, the probability that it has orthogonal subsets is zero. Also worth mentioning is the fact that, as in the case of orthogonal partitioning problem, when the initial set is created randomly, the probability that the matrices $\mathbf{S}^T(0)\mathbf{S}(0)$ for $K \leq N$ and $\mathbf{S}(0)\mathbf{S}^T(0)$ for $K > N$ will have nonfull rank is zero. Thus, the condition that $\mathbf{S}(0)$ must have $\min\{K, N\}$ linearly independent signature

Fig. 3. Minimum and maximum eigenvalues of matrix $\mathbf{S}(n)\mathbf{S}^T(n)$.Fig. 4. Total squared correlation, $\text{TSC}(n)$.

sequences is also satisfied with probability one when $\mathbf{S}(0)$ is created randomly.

We now present some simple numerical results to verify our analysis and support our conjecture. We take the processing gain to be $N = 10$. The initial signature sequences are created randomly. In all the figures, K updates take place between iterations n and $(n + 1)$. In each update, the i th user's signature sequence is replaced with the corresponding normalized MMSE filter, for $i = 1, \dots, K$. Figs. 1 and 2 show the minimum and maximum eigenvalues of the matrix $\mathbf{S}^T(n)\mathbf{S}(n)$, and $\text{TSC}(n)$, respectively, as a function of the iteration index for number of users $K = 5, 10$. As expected, the minimum and maximum eigenvalues of $\mathbf{S}^T(n)\mathbf{S}(n)$ converge to 1 implying that the matrix converges to $\mathbf{S}^T \mathbf{S} = \mathbf{I}_K$, and the TSC converges to K .

Figs. 3 and 4 show the minimum and maximum eigenvalues of the matrix $\mathbf{S}(n)\mathbf{S}^T(n)$ and $\text{TSC}(n)$, respectively, for a number of users $K = 20, 30, 40, 50$. As expected, the minimum and maximum eigenvalues of the matrix $\mathbf{S}(n)\mathbf{S}^T(n)$ converge to K/N implying that the matrix converges to $\mathbf{S}\mathbf{S}^T = (K/N)\mathbf{I}_N$, and the TSC converges to K^2/N .


 Fig. 5. Minimum and maximum eigenvalues of matrix $\mathbf{S}(n)\mathbf{S}^T(n)$.

 Fig. 6. Total squared correlation, $\text{TSC}(n)$.

In the following experiment, we observe the effects of orthogonal partitioning mentioned previously. The number of users is $K = 20$ and the processing gain is $N = 10$. Figs. 5 and 6 show the minimum and maximum eigenvalues of the matrix $\mathbf{S}(n)\mathbf{S}^T(n)$ and $\text{TSC}(n)$, respectively. Two kinds of orthogonal partitioning are examined. In the first, the users are partitioned into two orthogonal subspaces with $K_1 = K_2 = 10$ and $N_1 = N_2 = 5$. Since this is an “equal loading” case, the whole signature sequence set converges to a WBE sequence set: the matrix converges to $\mathbf{S}\mathbf{S}^T = (K/N)\mathbf{I}_N = 2\mathbf{I}_N$ and the TSC converges to $K^2/N = 40$; see the solid curves in Fig. 5 and the curve corresponding to $K_1 = 10$ in Fig. 6. In the second orthogonal partitioning we examined, the users are partitioned into two orthogonal sets according to $K_1 = 15$ and $K_2 = 5$ and, again, $N_1 = N_2 = 5$. Note that this partitioning does not yield equal loads to both subsets and the matrix $\mathbf{S}(n)\mathbf{S}^T(n)$ does not converge to a multiple of the identity matrix (see dashed lines in Fig. 5), and the $\text{TSC}(n)$ does not decrease down to $K^2/N = 40$, but converges to 50. Note that the signature sequences of $K_1 = 15$ users in the first subset converge to a WBE

sequence set achieving a TSC of $K_1^2/N_1 = 15^2/5 = 45$, and the signature sequences of $K_2 = 5$ users converge to orthogonal set yielding a TSC of $K_2 = 5$. Thus, the combined TSC is $45 + 5 = 50$ which can be seen in Fig. 6.

A. Further Remarks

From the numerical experiments, it appears that the deterministic MMSE update results in fast convergence to an optimal set of signatures. However, we must keep in mind that a practical implementation of the MMSE algorithm will rely on stochastic receiver measurements and adaptive implementations of the MMSE receiver filter. Stochastic convergence of an MMSE update algorithm needs to be examined. Practice also dictates that the feedback transmission of filter coefficients to the transmitter must not require excessive bandwidth. These concerns will need to be addressed for signature optimization to be useful in practical systems. An intriguing aspect of the MMSE update is that the basic approach extends readily to systems with multiple receivers. However, in this case, appropriate objectives for signature optimization for an ensemble of users must be formulated.

APPENDIX ADDITIONAL PROOFS

Proof of Theorem 1: From (18) and (22), we observe that $\overline{\text{TSC}} \leq \text{TSC}$ iff

$$\mathbf{s}_k^T \mathbf{A}_k^{-1} \mathbf{s}_k \leq \left(\mathbf{s}_k^T \mathbf{A}_k^{-2} \mathbf{s}_k \right) \left(\mathbf{s}_k^T \mathbf{A}_k \mathbf{s}_k \right). \quad (34)$$

Since \mathbf{A}_k is symmetric and positive definite we can write $\mathbf{A}_k = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ with $\mathbf{\Lambda}$ being the diagonal matrix of eigenvalues of \mathbf{A}_k and with the columns of \mathbf{V} being the orthonormal eigenvectors of the same matrix. Since $\mathbf{A}_k^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T$ and $\mathbf{A}_k^{-2} = \mathbf{V}\mathbf{\Lambda}^{-2}\mathbf{V}^T$, by defining $\mathbf{x} = \mathbf{V}^T \mathbf{s}_k$, we see that $\overline{\text{TSC}} \leq \text{TSC}$ iff

$$\mathbf{x}^T \mathbf{\Lambda}^{-1} \mathbf{x} \leq \left(\mathbf{x}^T \mathbf{\Lambda}^{-2} \mathbf{x} \right) \left(\mathbf{x}^T \mathbf{\Lambda} \mathbf{x} \right). \quad (35)$$

Since $\mathbf{x}^T \mathbf{x} = 1$, we can employ the Cauchy–Schwarz inequality to write

$$\begin{aligned} 1 &= \left(\mathbf{x}^T \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{-1/2} \mathbf{x} \right)^2 \leq \|\mathbf{\Lambda}^{-1/2} \mathbf{x}\|^2 \|\mathbf{\Lambda}^{1/2} \mathbf{x}\|^2 \\ &= \left(\mathbf{x}^T \mathbf{\Lambda}^{-1} \mathbf{x} \right) \left(\mathbf{x}^T \mathbf{\Lambda} \mathbf{x} \right). \end{aligned} \quad (36)$$

Once again, we use the Cauchy–Schwarz inequality to find that

$$\left(\mathbf{x}^T \mathbf{\Lambda}^{-1} \mathbf{x} \right)^2 = \left((\mathbf{\Lambda}^{-1} \mathbf{x})^T \mathbf{x} \right)^2 \leq \|\mathbf{\Lambda}^{-1} \mathbf{x}\|^2 \|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{\Lambda}^{-2} \mathbf{x}. \quad (37)$$

Applying (36) and (37) in succession yields

$$\mathbf{x}^T \mathbf{\Lambda}^{-1} \mathbf{x} \leq \left(\mathbf{x}^T \mathbf{\Lambda}^{-1} \mathbf{x} \right)^2 \left(\mathbf{x}^T \mathbf{\Lambda} \mathbf{x} \right) \leq \left(\mathbf{x}^T \mathbf{\Lambda}^{-2} \mathbf{x} \right) \left(\mathbf{x}^T \mathbf{\Lambda} \mathbf{x} \right). \quad (38)$$

Thus, we have shown $\overline{\text{TSC}} \leq \text{TSC}$. Note that inequalities (36) and (37) are equalities iff $\mathbf{x} = \beta \mathbf{\Lambda}^{-1} \mathbf{x}$ or, equivalently, $\mathbf{s}_k =$

$\beta \mathbf{A}_k^{-1} \mathbf{s}_k$. That is, $\overline{\text{TSC}} \leq \text{TSC}$ is satisfied with equality iff the normalized MMSE operation on \mathbf{s}_k yields \mathbf{s}_k . \square

Proof of Theorem 2: We will only show that

$$\det(x\mathbf{I}_N + \overline{\mathbf{S}}\overline{\mathbf{S}}^T) \geq \det(x\mathbf{I}_N + \mathbf{S}\mathbf{S}^T). \quad (39)$$

The second inequality in Theorem 2 follows from the determinant equality $\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$. Assume for the time being that $x > 0$. We will treat the case of $x = 0$ separately at the end.

By defining $\mathbf{Z}_k = \sum_{j \neq k} \mathbf{s}_j \mathbf{s}_j^T$, the right-hand side of (39) can be written as

$$\begin{aligned} \det(x\mathbf{I}_N + \mathbf{S}\mathbf{S}^T) &= \det(x\mathbf{I}_N + \mathbf{Z}_k + \mathbf{s}_k \mathbf{s}_k^T) \\ &= \det(\mathbf{Z}_k + x\mathbf{I}_N) \det(\mathbf{I}_N + (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} \mathbf{s}_k \mathbf{s}_k^T) \\ &= \det(\mathbf{Z}_k + x\mathbf{I}_N) (1 + \mathbf{s}_k^T (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} \mathbf{s}_k). \end{aligned} \quad (40)$$

Similarly, the left-hand side of (39) can be written as

$$\det(x\mathbf{I}_N + \overline{\mathbf{S}}\overline{\mathbf{S}}^T) = \det(\mathbf{Z}_k + x\mathbf{I}_N) (1 + \mathbf{c}_k^T (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} \mathbf{c}_k) \quad (41)$$

where \mathbf{c}_k is given by (20). Thus, (39) holds iff

$$\mathbf{c}_k^T (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} \mathbf{c}_k \geq \mathbf{s}_k^T (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} \mathbf{s}_k. \quad (42)$$

Applying \mathbf{c}_k from (20) with $\mathbf{A}_k = \mathbf{Z}_k + a^2 \mathbf{I}_N$, we must show that

$$\frac{\mathbf{s}_k^T (\mathbf{Z}_k + a^2 \mathbf{I}_N)^{-1} (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} (\mathbf{Z}_k + a^2 \mathbf{I}_N)^{-1} \mathbf{s}_k}{\mathbf{s}_k^T (\mathbf{Z}_k + a^2 \mathbf{I}_N)^{-2} \mathbf{s}_k} \geq \mathbf{s}_k^T (\mathbf{Z}_k + x\mathbf{I}_N)^{-1} \mathbf{s}_k. \quad (43)$$

Let us denote $\mathbf{Z}_k = \mathbf{V}\mathbf{D}\mathbf{V}^T$, where diagonal matrix $\mathbf{D} = \text{diag}\{d_1, \dots, d_N\}$ contains the eigenvalues of \mathbf{Z}_k and \mathbf{V} contains the eigenvectors of \mathbf{Z}_k as its columns. Defining $\mathbf{w} = \mathbf{V}^T \mathbf{s}_k$, we can write (43) as

$$\sum_{i=1}^N \frac{w_i^2}{(d_i + x)(d_i + a^2)^2} \geq \left(\sum_{i=1}^N \frac{w_i^2}{d_i + x} \right) \left(\sum_{i=1}^N \frac{w_i^2}{(d_i + a^2)^2} \right). \quad (44)$$

Thus, in order to prove (39) and therefore the claims of Theorem 2 for $x > 0$, we have to prove (44). To do so, we use the following Lemma which is proved immediately following this proof.

Lemma 3: Let Y be a discrete random variable with finite range $\{y_1, \dots, y_n\}$. If $g(\cdot)$ and $h(\cdot)$ are both nonincreasing functions or $g(\cdot)$ and $h(\cdot)$ are both nondecreasing functions, then

$$E[g(Y)h(Y)] \geq E[g(Y)]E[h(Y)]. \quad (45)$$

Since $\sum_i w_i^2 = 1$, we can define a discrete random variable Y such that

$$P_Y(y) = \begin{cases} w_i^2, & y = d_i, 1 \leq i \leq N \\ 0, & \text{otherwise.} \end{cases} \quad (46)$$

Thus, (44) is true iff

$$E \left[\frac{1}{(Y+x)(Y+a^2)^2} \right] \geq E \left[\frac{1}{Y+x} \right] E \left[\frac{1}{(Y+a^2)^2} \right] \quad (47)$$

which follows from Lemma 3.

Thus far, we proved (39) for $x > 0$. Let us denote the eigenvalues of $\mathbf{S}\mathbf{S}^T$ by $\lambda_1, \dots, \lambda_N$ and the eigenvalues of $\overline{\mathbf{S}}\overline{\mathbf{S}}^T$ by μ_1, \dots, μ_N . Thus, both sides of (39) can be represented as

$$\begin{aligned} \tilde{f}(x) &= \det(x\mathbf{I}_N + \mathbf{S}\mathbf{S}^T) = \prod_{i=1}^N (x + \lambda_i) \\ \bar{f}(x) &= \det(x\mathbf{I}_N + \overline{\mathbf{S}}\overline{\mathbf{S}}^T) = \prod_{i=1}^N (x + \mu_i). \end{aligned} \quad (48)$$

Note that both $\tilde{f}(x)$ and $\bar{f}(x)$ are continuous functions of x . In particular, they are N th-order polynomials in x . Similarly, $f(x)$ defined as $f(x) = \bar{f}(x) - \tilde{f}(x)$ is also an N th-order polynomial function of x and is continuous. We have already proved that $f(x) \geq 0$ for $x > 0$. From the continuity of $f(x)$ we must have $f(0) \geq 0$, proving the desired result. \square

Proof of Lemma 3: We will prove the lemma by induction on n . When $n = 1$, $Y = y_1$, and the claim holds trivially. Assume the claim holds for any discrete random variable with range $\{y_1, \dots, y_{n-1}\}$ and let Y be a random variable with range $\{y_1, \dots, y_n\}$. Without loss of generality, we can assume $y_1 \leq y_2 \leq \dots \leq y_n$. Defining $p_n = P_Y(y_n)$, $h_n = h(y_n)$, and $g_n = g(y_n)$ we can write

$$E[g(Y)h(Y)] = (1 - p_n) \sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} g_i h_i + p_n g_n h_n. \quad (49)$$

By our induction hypothesis

$$\begin{aligned} E[g(Y)h(Y)] &\geq (1 - p_n) \left(\sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} g_i \right) \left(\sum_{j=1}^{n-1} \frac{p_j}{1 - p_n} h_j \right) + p_n g_n h_n \quad (50) \\ &= \frac{(E[g(Y)] - p_n g_n)(E[h(Y)] - p_n h_n)}{1 - p_n} + p_n g_n h_n \quad (51) \\ &= E[g(Y)]E[h(Y)] + \frac{p_n}{1 - p_n} (E[g(Y)] - g_n)(E[h(Y)] - h_n). \end{aligned} \quad (52)$$

Note that $y_i \leq y_n$ for all i . If $g(\cdot)$ and $h(\cdot)$ are both nonincreasing, $g_i \geq g_n$ and $h_i \geq h_n$ for all i . In this case, $E[g(Y)] \geq g_n$ and $E[h(Y)] \geq h_n$. This implies

$$(E[g(Y)] - g_n)(E[h(Y)] - h_n) \geq 0 \quad (53)$$

and thus the claim follows. If $g(\cdot)$ and $h(\cdot)$ are both nondecreasing, $g_i \leq g_n$ and $h_i \leq h_n$ for all i . In this case, $E[g(Y)] \leq g_n$ and $E[h(Y)] \leq h_n$. Thus, (53) remains true and the claim follows. \square

Proof of Theorem 3: Rewriting (32) in a matrix form, we obtain

$$(\mathbf{S}\mathbf{S}^T) \mathbf{S} = \mathbf{S}\Phi \quad (54)$$

where $\Phi = \text{diag}\{\phi_1, \dots, \phi_K\}$. Multiplying (54) from left by \mathbf{S}^\top yields

$$(\mathbf{S}^\top \mathbf{S})^2 = \mathbf{S}^\top \mathbf{S} \Phi. \quad (55)$$

If the algorithm is started with K linearly independent vectors $\mathbf{S}(0)$, then by Corollary 3, $\mathbf{S}^\top \mathbf{S}$ will be invertible. Multiplying both sides of (55) with $(\mathbf{S}^\top \mathbf{S})^{-1}$ yields

$$\mathbf{S}^\top \mathbf{S} = \Phi. \quad (56)$$

Thus, $\mathbf{S}^\top \mathbf{S}$ is a diagonal matrix. By the nature of the algorithm, $\mathbf{s}_i^\top \mathbf{s}_i = 1$ for all i , and the diagonal elements of $\mathbf{S}^\top \mathbf{S}$ are all unity. Therefore, we have $\mathbf{S}^\top \mathbf{S} = \mathbf{I}_K$. \square

Proof of Lemma 1: Let us denote $\mathbf{S}(0)$ as \mathbf{S} which is partitioned into orthogonal subsets \mathbf{S}_l , $l = 1, \dots, L$, as $\mathbf{S} = [\mathbf{S}_1] \cdots [\mathbf{S}_L]$ so that $\mathbf{S}_l^\top \mathbf{S}_m = \mathbf{0}$ for $l \neq m$. Suppose we update \mathbf{s} in \mathbf{S}_m to \mathbf{c} in the current iteration. From (30), the new signature sequence is

$$\mathbf{c} = \alpha^{-1} \mathbf{B}^{-1} \mathbf{s} \quad (57)$$

where α is chosen to normalize \mathbf{c} to unit length, and

$$\mathbf{B} = \mathbf{S} \mathbf{S}^\top + a^2 \mathbf{I}_N = \sum_{l=1}^L \mathbf{S}_l \mathbf{S}_l^\top + a^2 \mathbf{I}_N. \quad (58)$$

We define $\mathbf{B}_m = \mathbf{S}_m \mathbf{S}_m^\top + a^2 \mathbf{I}_N$. Since (57) says that $\alpha \mathbf{B} \mathbf{c} = \mathbf{s}$, we have that $\alpha \mathbf{B}_m \mathbf{B} \mathbf{c} = \mathbf{B}_m \mathbf{s}$. We observe that \mathbf{B} and \mathbf{B}_m commute, and that $\mathbf{B}_m \mathbf{s} = \mathbf{B} \mathbf{s}$. Thus, $\alpha \mathbf{B} \mathbf{B}_m \mathbf{c} = \mathbf{B} \mathbf{s}$. Since \mathbf{B} is invertible, it follows that $\mathbf{c} = \alpha^{-1} \mathbf{B}_m^{-1} \mathbf{s}$. This implies that the span of \mathbf{S}_m is unchanged. Now note that for $l \neq m$, $\mathbf{B}_m \mathbf{S}_l = a^2 \mathbf{S}_l$; hence, $\mathbf{B}_m^{-1} \mathbf{S}_l = a^{-2} \mathbf{S}_l$. Therefore, for $l \neq m$

$$\mathbf{c}^\top \mathbf{S}_l = \alpha^{-1} \mathbf{s}^\top \mathbf{B}_m^{-1} \mathbf{S}_l = \alpha^{-1} a^{-2} \mathbf{s}^\top \mathbf{S}_l = 0 \quad (59)$$

which proves that the new signature \mathbf{c} is also orthogonal to all \mathbf{S}_l for $l \neq m$. Repeated application of this result yields the desired result. \square

Proof of Theorem 4: We note that we can write $\mathbf{S} \mathbf{S}^\top = \mathbf{S}_1 \mathbf{S}_1^\top + \cdots + \mathbf{S}_L \mathbf{S}_L^\top$. Since a signature $\mathbf{s}_j \in \mathbf{S}_l$ is an eigenvector of $\mathbf{S} \mathbf{S}^\top$ with eigenvalue ω_l and is orthogonal to all signatures not in \mathbf{S}_l , we have

$$\mathbf{S} \mathbf{S}^\top \mathbf{s}_j = \mathbf{S}_l \mathbf{S}_l^\top \mathbf{s}_j = \omega_l \mathbf{s}_j. \quad (60)$$

Since every eigenvector of $\mathbf{S}_l \mathbf{S}_l^\top$ with nonzero eigenvalue is a linear combination of columns of \mathbf{S}_l , every nonzero eigenvalue of $\mathbf{S}_l \mathbf{S}_l^\top$ is ω_l . Since \mathbf{S}_l has rank N_l , the eigenvalue ω_l has multiplicity N_l and the eigenvalue 0 has multiplicity of $N - N_l$. Using

$$\omega_l N_l = \text{trace}(\mathbf{S}_l \mathbf{S}_l^\top) = \text{trace}(\mathbf{S}_l^\top \mathbf{S}_l) = K_l$$

one finds $\omega_l = K_l/N_l$. The TSC of the subset \mathbf{S}_l is

$$\text{TSC}_l = \text{trace}[(\mathbf{S}_l \mathbf{S}_l^\top)^2] = K_l^2/N_l$$

which is lower than the TSC of any other matrix $\overline{\mathbf{S}}_l$ which contains K_l signature sequences and has rank N_l . This can be seen by first noting that for any other $\overline{\mathbf{S}}_l$ having K_l unit-energy signature sequences and a rank of N_l , the set of eigenvalues of $\overline{\mathbf{S}}_l \overline{\mathbf{S}}_l^\top$ has the form $\{\lambda_1, \dots, \lambda_{N_l}, 0, \dots, 0\}$ where

there are N_l nonzero eigenvalues and $N - N_l$ zero eigenvalues with $\sum_{j=1}^{N_l} \lambda_j = K_l$. Now, the TSC of set $\overline{\mathbf{S}}_l$ is

$$\overline{\text{TSC}}_l = \text{trace}[(\overline{\mathbf{S}}_l \overline{\mathbf{S}}_l^\top)^2] = \sum_{j=1}^{N_l} \lambda_j^2.$$

We define a random variable X which takes values of λ_j for $j = 1, \dots, N_l$ with probability $1/N_l$. Then, the simple inequality $E[X^2] \geq (E[X])^2$ is equivalent to $\overline{\text{TSC}}_l \geq \text{TSC}_l$. This inequality is satisfied with equality iff X is deterministic, i.e., all λ_j are equal. When they are equal, they must all be equal to K_l/N_l since $\sum_{j=1}^{N_l} \lambda_j = K_l$. \square

Proof of Lemma 2: The TSC of the set is equal to the sum of the TSCs of the orthogonal subsets since $\mathbf{S}_l^\top \mathbf{S}_m = \mathbf{0}$ for $l \neq m$

$$\text{TSC} = \text{trace}[(\mathbf{S} \mathbf{S}^\top)^2] = \sum_{l=1}^L \text{trace}[(\mathbf{S}_l \mathbf{S}_l^\top)^2] = \sum_{l=1}^L \frac{K_l^2}{N_l} \quad (61)$$

since the eigenvalues of $\mathbf{S}_l \mathbf{S}_l^\top$ are K_l/N_l and 0 with multiplicities of N_l and $N - N_l$, as stated in Theorem 4. Similar to the proof of Theorem 4, define a random variable X which takes the values of K_l/N_l for $l = 1, \dots, L$ with probability N_l/N for $l = 1, \dots, L$. Note that $\sum_{l=1}^L N_l = N$, by Theorem 4. Then, $\text{TSC} \geq K^2/N$ follows from $E[X^2] \geq (E[X])^2$; and $\text{TSC} = K^2/N$ iff X is deterministic, i.e., iff all K_l/N_l coincide. \square

REFERENCES

- [1] S. Ulukus, "Power control, multiuser detection and interference avoidance in CDMA systems," Ph.D. dissertation, Dept. Elec. Comput. Eng., Rutgers Univ., Piscataway, NJ, July 1998. Available: [Online] <http://www.research.att.com/~ulukus>.
- [2] S. Ulukus and R. D. Yates, "Iterative signature adaptation for capacity maximization of CDMA systems," in *36th Annu. Allerton Conf. Communications, Control and Computing*, Sept. 1998.
- [3] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [4] —, "Capacity region of Gaussian CDMA channels: The symbol-synchronous case," in *24th Allerton Conf. Communication, Control and Computing*, Oct. 1986, pp. 1025–1034.
- [5] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Tran. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
- [6] 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.
- [7] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1261–1266, July 1994.
- [8] P. Viswanath, V. Anantharam, and D. 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.
- [9] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 683–690, May 1990.
- [10] 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.
- [11] 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.
- [12] B. R. Vojcic and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Commun.*, vol. 46, pp. 1346–1355, Oct. 1998.

- [13] 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.
- [14] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna arrays with feedback," *IEEE Signal Processing Lett.*, vol. 1, pp. 150–152, Oct. 1994.
- [15] —, "Adaptive transmitting antenna methods for multipath environments," in *Proc. IEEE Global Telecommun. Conf.*, 1994, pp. 425–429.
- [16] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming for cellular communication systems," in *Proc. Conf. Information Sciences and Systems*, 1997, pp. 92–97.
- [17] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter and receiver beamforming for maximum capacity in spatial division multi-access," in *Proc. 35th Annu. Allerton Conf. Communications, Control and Computing*, 1997, pp. 93–101.
- [18] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Vehicular Technology Conf.*, May 1999, pp. 851–856.
- [19] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.
- [20] J. L. Massey, "Welch's bound for the correlation of a sequence set," in *Proc. IEEE Int. Symp. Information Theory*, 1991, p. 385.
- [21] J. L. Massey and T. Mittelholzer, "Welch's bound and sequence sets for code-division multiple-access systems," in *Sequences II Methods in Communication, Security and Computer Science*, R. Capocelli, A. De Santis, and U. Vaccaro., Eds. New York: Springer-Verlag, 1991.
- [22] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.
- [23] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685–697, May 1994.
- [24] M. Abdulrahman, A. U. H. Sheikh, and D. D. Falconer, "Decision feedback equalization for CDMA in indoor wireless communications," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 698–706, May 1994.
- [25] 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.