Copyright © 2001, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

ITERATIVE CONSTRUCTION OF OPTIMAL SIGNATURE SEQUENCES FOR CDMA

by

Pablo Anigstein and Venkat Anantharam

Memorandum No. UCB/ERL M01/24

1 February 2001

ITERATIVE CONSTRUCTION OF OPTIMAL SIGNATURE SEQUENCES FOR CDMA

by

Pablo Anigstein and Venkat Anantharam

Memorandum No. UCB/ERL M01/24

1 February 2001

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Iterative construction of optimal signature sequences for CDMA *

Pablo Anigstein and Venkat Anantharam
Dept. of Electrical Engineering and Computer Sciences
University of California at Berkeley
Berkeley, CA 94720
{pabloa,ananth}@eecs.berkeley.edu

February 2001

Abstract

Viswanath and Anantharam ([15]) characterize the sum capacity of multi-access vector channels. For given number of users, received powers, spreading gain and noise covariance matrix in a code-division multiple-access (CDMA) system, [15] presents a combinatorial algorithm to generate a set of signature sequences that achieves the maximum sum capacity. These sets also minimize a performance measure called total square correlation (TSC).

Ulukus and Yates ([10]) propose an iterative algorithm suitable for distributed implementation: at each step one signature sequence is replaced by its linear minimum mean square error (MMSE) filter. This algorithm results in a decrease of TSC at each step. The MMSE iteration has fixed points not only at the optimal configurations which attain the global minimum TSC but also at other configurations which are suboptimal. [10] claims that simulations show that when starting with random sequences, the algorithm converges to optimum sets of sequences, but gives no formal proof.

We show that the TSC function has no local minima, in the sense that given any suboptimal set of sequences, there exist arbitrarily close sets with lower TSC. Therefore, only the optimal sets are stable fixed points of the MMSE iteration. We define a noisy version of the MMSE iteration as follows: after replacing all the signature sequences one at a time by their linear MMSE filter, we add a bounded random noise to all the sequences. Using our observation about the TSC function, we can prove that if we choose the bound on the noise adequately, making it decrease to zero, the noisy MMSE iteration converges to the set of optimal configurations with probability one for any initial set of sequences.

*This research is supported by EPRI/DOD Complex Interactive Networks under contract number EPRI-W08333-04, NSF ANI 9872764, ECS 9873086 and IIS 9941569

1 Introduction and previous work

We consider the uplink of a symbol-synchronous CDMA system. An important performance measure of such system is the sum capacity, the maximum sum of rates of the users at which reliable communication can take place. If we fix the processing gain, number of users, and received user powers, we can regard the sum capacity as a function of the signature sequences assigned to the users. We will refer to such an assignment as a "configuration" of signature sequences. A signature sequence will be modeled as a unit-norm real vector of dimension equal to the spreading gain.

The capacity region of a symbol-synchronous CDMA channel was first obtained in [11]. Later [7] characterized the maximum sum capacity of a CDMA channel with white noise and equal user received powers. In [14] the case of different user received powers was solved using majorization theory. [15] also considers the case of asymmetric received powers with colored noise, and gives a recursive algorithm to construct an optimal configuration of signature sequences.

Another performance measure of the CDMA channel is the total square correlation (TSC). An iterative procedure called MMSE iteration, in which at each step one signature sequence is modified in a way such that TSC is non-increasing, was proposed in [9, 10]. Another iterative procedure with the same property is proposed in [6]. These algorithms are suitable for distributed implementation. The main idea is that the receiver for some user would periodically decide on an update for the signature sequence of that user and communicate it to the user through some feedback channel. The user transmitter would then switch to the new signature sequence. When these algorithms are applied, TSC is non-increasing, but there is no guarantee that the TSC will converge to its minimum possible value. Nevertheless, simulations suggest that when the initial signature sequences are chosen at random, the iteration converges to the minimum of TSC. A modification of the algorithm of [6] is proposed in [5] in order to guarantee convergence to the optimum TSC value. However, the modified algorithm has increased complexity and is not suitable for distributed implementation.

We will define a modified version of the MMSE iteration adding noise and prove almost sure convergence of the TSC to the global minimum. A short version of the results herein was presented in [1].

2 Outline

The rest of this report is organized as follows. In section 3 we present the CDMA channel model used. Section 4 obtains the linear MMSE filter. In section 5 we define the majorization partial order on $\mathbb{R}^n$ and state some results that will be used later. In section 6 the two performance measures used, sum capacity and TSC, are defined and basic properties are observed. Section 7 presents the MMSE iteration. The fixed configurations of this iteration are characterized, and we prove that the MMSE iteration asymptotically approaches the set of fixed configurations. In section 8 we state the recursive algorithm of [15] which obtains the maximum sum capacity and a configuration of signature sequences attaining it. We give a proof of the optimality of the algorithm which is different from the one in [15] and is useful later. In the process, we prove using results of section 6, that the optimal configurations that maximize sum capacity are the same that minimize TSC. In section 9 we observe and prove that TSC has no minima other than the global minima. Motivated by this result, in section 10 we define a modified version of the MMSE update adding noise. We prove that if the noise bound is chosen adequately, the noisy MMSE iteration converges to the optimum TSC almost surely regardless of the initial configuration.

3 Model

Consider a symbol-synchronous CDMA system with K users. Let T be the duration of the symbol interval and let $s_k : [0,T] \to \mathbb{R}$ represent the signature waveform assigned to user k, assumed unit-norm in $L^2[0,T]$, i.e.

$$\int_0^T (s_k(t))^2 dt = 1$$

The received signal at the base station in one symbol interval can then be expressed as

$$y(t) = \sum_{k=1}^{K} \sqrt{p_k} x_k s_k(t) + z(t) \quad , \quad t \in [0,T] \tag{1}$$

Here p_k is the power received from user k. The information transmitted by user k is modeled by the random variable x_k having zero mean and unit variance, and independent of the information transmitted by other users. The noise $z(t)$ is assumed a zero-mean Gaussian process independent of the user symbols $x_1, \ldots, x_K$.

Let the processing gain be N. The signature waveforms are then constrained to be of the form

$$s_k(t) = \sum_{n=1}^{N} s_{k,n} \psi_n(t)$$

where $\{\psi_1(t), \ldots, \psi_N(t)\}$ is an orthonormal set in $L^2[0,T]$, and as $s_k(t)$ is unit norm $\sum_{n=1}^{N} s_{k,n}^2 = 1$. If we write

$$\begin{aligned} y_n &= \int_0^T y(t)\psi_n(t)dt \\ z_n &= \int_0^T z(t)\psi_n(t)dt \end{aligned}$$

we have

$$y_n = \sum_{k=1}^{K} \sqrt{p_k} x_k s_{k,n} + z_n$$

Hence writing $s_k = \begin{bmatrix} s_{k,1} & \dots & s_{k,n} \end{bmatrix}^T$, $y = \begin{bmatrix} y_1 & \dots & y_N \end{bmatrix}^T$ and $z = \begin{bmatrix} z_1 & \dots & z_N \end{bmatrix}^T$ we obtain [1]

$$y = \sum_{k=1}^{K} \sqrt{p_k} x_k s_k + z \tag{2}$$

Here s_k is a unit-norm (i.e. $s_k^T s_k = 1$) N-dimensional column vector corresponding to the signature sequence of user k .

If we write $S = \begin{bmatrix} s_1 & \dots & s_K \end{bmatrix}$, $D = \text{diag}(p_1, \dots, p_K)$ and $x = \begin{bmatrix} x_1 & \dots & x_K \end{bmatrix}^T$ equation (2) can be rewritten as

$$y = SD^{\frac{1}{2}} x + z \tag{3}$$

Because of our assumption on the noise, z is a Gaussian distributed zero-mean N-dimensional column vector independent of x. We will denote the covariance of z as $E[zz^T] = W$, a $K \times K$ symmetric positive definite matrix. Usually the noise process $z(t)$ is assumed white. In that case, W is a multiple of the identity matrix and y is easily shown to be a sufficient statistic for estimating x. Note that if the noise is not white, then not only the different components of z, but also the vectors z corresponding to different symbol intervals will be correlated. Moreover, in this case y is not a sufficient statistic. Nevertheless, we will just consider the model (3) with an arbitrary symmetric positive definite noise covariance matrix W, and to compute the sum capacity, the noise vector z will be assumed uncorrelated across different symbol intervals. The solution of this case of colored noise may provide insight for the consideration of a system with multiple base stations, where users communicating with one base station could be modeled as noise at the other base station.

In the sequel we assume N, K, p_k ($k \in \{1, \dots, K\}$) and W are given and fixed. Thus a configuration is determined by the signatures matrix $S \in \mathcal{S}$ where

$$\mathcal{S} = \left\{ \begin{bmatrix} s_1 & \dots & s_K \end{bmatrix} : s_k \in \mathbb{S}^{N-1} \, \forall k \in \{1, \dots, K\} \right\} \tag{4}$$

with $\mathbb{S}^{N-1} = \{s \in \mathbb{R}^N : ||s|| = 1\}$ the unit-sphere in $\mathbb{R}^N$.

4 MMSE linear filter

A linear filtering of the received signal is represented by an N-dimensional column vector v, and has output $v^T y$. The MMSE linear filter for user k, which we will denote v_k, is defined as the linear filter that minimizes the mean squared difference between the information transmitted by user k (x_k) and the output of the filter. Let us write $\text{MSE}_k(v) = \text{E}[(v^T y - x_k)^2]$. Then

$$v_k = \arg \min_{v \in \mathbb{R}^N} \text{MSE}_k(v)$$

[1] Hopefully the abuse of notation of representing with the same symbol the functions of time and the corresponding column vectors will not lead to confusion.

Now

$$\text{MSE}_k(v) = v^T\text{E}[yy^T]v - 2v^T\text{E}[x_k y] + \text{E}[x_k^2]$$

Using (3), $\text{E}[xx^T] = I$ and the independence of x and z we get

$$\begin{aligned}\text{E}[yy^T] &= SD^{\frac{1}{2}}\text{E}[xx^T]D^{\frac{1}{2}}S^T + SD^{\frac{1}{2}}\text{E}[xz^T] + \text{E}[zx^T]D^{\frac{1}{2}}S\text{E}[xz^T] + \text{E}[zz^T] \\ &= SDS^T + W\end{aligned}$$

and

$$\begin{aligned}\text{E}[x_k y] &= SD^{\frac{1}{2}}\text{E}[x_k x] \\ &= SD^{\frac{1}{2}}e_k \\ &= \sqrt{p_k}s_k\end{aligned}$$

where e_k is the k-th canonical unit-norm vector in $\mathbb{R}^K$. Therefore

$$\text{MSE}_k(v) = v^T(SDS^T + W)v - 2\sqrt{p_k}v^T s_k + 1$$

We see that $\text{MSE}_k(v)$ is a convex function of v, and to compute v_k we can differentiate the above expression with respect to v and equate the gradient to zero, to obtain

$$2(SDS^T + W)v_k - 2\sqrt{p_k}s_k = 0$$

Solving for v_k:

$$v_k = \sqrt{p_k}(SDS^T + W)^{-1}s_k \tag{5}$$

Given S and $k \in \{1, \ldots, K\}$ we will write $D_k = \text{diag}(p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_K)$ and $S_k = \begin{bmatrix} s_1 & \ldots & s_{k-1} & s_{k+1} & \ldots & s_K \end{bmatrix}$.

Another expression for v_k can be obtained from (5) using the known formula [2]

$$(A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1}$$

which holds for matrices A, B, C of suitable dimensions whenever A and $A + BC$ are nonsingular. Taking $A = S_k D_k S_k^T + W$, $B = p_k s_k$ and $C = s_k^T$ we obtain

$$(SDS^T + W)^{-1} = (S_k D_k S_k^T + W)^{-1} - \frac{(S_k D_k S_k^T + W)^{-1} p_k s_k s_k^T (S_k D_k S_k^T + W)^{-1}}{1 + p_k s_k^T (S_k D_k S_k^T + W)^{-1} s_k}$$

Hence from (5) we get

$$v_k = \frac{\sqrt{p_k}}{1 + p_k s_k^T (S_k D_k S_k^T + W)^{-1} s_k}(S_k D_k S_k^T + W)^{-1} s_k \tag{6}$$

Another important property of the filter v_k is that it maximizes the output signal-to-interference ratio (SIR) of user k over all linear receivers ([12]). To see this, note

[2]This expression can be verified directly multiplying the right hand side by $A + BC$ and simplifying.

that the output of filter v is $\sum_{k=1}^{K} \sqrt{p_k} x_k v^T s_k + v^T z$. Hence

$$\begin{aligned}
\mathrm{SIR}_k(v) &= \frac{\mathrm{E}[(< v, \sqrt{p_k} x_k s_k >)^2]}{\mathrm{E}[(< v, y - \sqrt{p_k} x_k s_k >)^2]} \\
&= \frac{p_k (v^T s_k)^2}{\sum_{k' \neq k} p_{k'} (v^T s_{k'})^2 + v^T W v} \\
&= \frac{p_k v^T s_k s_k^T v}{v^T (S_k D_k S_k^T + W) v}
\end{aligned}$$

Let $\hat{v} = (S_k D_k S_k^T + W)^{\frac{1}{2}} v$. Then

$$\begin{aligned}
\mathrm{SIR}_k(v) &= \frac{p_k \hat{v}^T (S_k D_k S_k^T + W)^{-\frac{1}{2}} s_k s_k^T (S_k D_k S_k^T + W)^{-\frac{1}{2}} \hat{v}}{\hat{v}^T \hat{v}} \\
&= \frac{p_k \left(\hat{v}^T (S_k D_k S_k^T + W)^{-\frac{1}{2}} s_k \right)^2}{||\hat{v}||^2}
\end{aligned}$$

Using Cauchy-Schwartz inequality,

$$\mathrm{SIR}_k(v) \leq p_k s_k^T (S_k D_k S_k^T + W)^{-1} s_k$$

with equality if and only if $\hat{v} = \alpha (S_k D_k S_k^T + W)^{-\frac{1}{2}} s_k$, i.e. $v = \alpha (S_k D_k S_k^T + W)^{-1} s_k$ for some $\alpha \in \mathbb{R}$.

5 Majorization

In this section we define the majorization partial order on $\mathbb{R}^n$. This order makes precise the notion that the components of a vector are "less spread out" or "more nearly equal" than those of another.

Given $a \in \mathbb{R}^n$, the components of a in decreasing order, called the order statistics of a, will be denoted $a_{[1]}, \ldots, a_{[n]}$. In other words, $(a_{[1]}, \ldots, a_{[n]})$ is the permutation of $(a_1, \ldots, a_n)$ such that $a_{[1]} \geq \ldots \geq a_{[n]}$.

Given $a, b \in \mathbb{R}^n$, we say that a majorizes b iff

$$\begin{aligned}
\sum_{i=1}^{m} a_i &\geq \sum_{i=1}^{m} b_i \qquad \forall m \in \{1, \ldots, n-1\} \\
\sum_{i=1}^{n} a_i &= \sum_{i=1}^{n} b_i
\end{aligned}$$

As a trivial example, given any $a \in \mathbb{R}^n$,

$$(a_1, \ldots, a_n) \text{ majorizes } \left(\frac{1}{n} \sum_{i=1}^{n} a_i, \ldots, \frac{1}{n} \sum_{i=1}^{n} a_i \right)$$

The following theorem will be useful later.

Theorem 1. *Let $H \in \mathbb{R}^{n\times n}$ be symmetric with diagonal elements $h_1, \dots, h_n$ and eigenvalues $\lambda_1, \dots, \lambda_n$. Then λ majorizes h.*

Conversely, if $\lambda, h \in \mathbb{R}^n$ and λ majorizes h, then there exists a symmetric matrix $H \in \mathbb{R}^{n\times n}$ with diagonal elements $h_1, \dots, h_n$ and eigenvalues $\lambda_1, \dots, \lambda_n$.

Proof. See theorems 9.B.1 and 9.B.2 in [4]. □

In the sequel, given a symmetric matrix $H \in \mathbb{R}^{n\times n}$ we will denote $\lambda(H)$ the vector whose components are the eigenvalues of H in non-decreasing order. I.e. if $\lambda_1 \geq \dots \geq \lambda_n$ are the eigenvalues of H, we will write $\lambda(H) = (\lambda_1, \dots, \lambda_n)$.

The following lemma will be used later.

Lemma 1. *Let $H \in \mathbb{R}^{n\times n}$ be symmetric and nonnegative definite and let $v \in \mathbb{S}^{n-1}$ be a unit-norm eigenvector associated with the minimum eigenvalue of H. Then, for all $p > 0$ and all $s \in \mathbb{S}^{n-1}$,*

$$\lambda(H + pss^T) \text{ majorizes } \lambda(H + pvv^T)$$

Proof. See [16] or [13]. □

A function $f : A \to \mathbb{R}$ (with $A \subset \mathbb{R}^n$) is said to be Schur-convex iff for all $a, b \in A$ such that a majorizes b we have $f(a) \geq f(b)$. If $-f$ is Schur-convex, f is said to be Schur-concave.

Lemma 2. *Let $g : A \to \mathbb{R}$ (with $A \subset \mathbb{R}$ a convex set) be convex (concave). Then the symmetric function $f : A^n \to \mathbb{R}$ with $f(a) = \sum_{i=1}^n g(a_i)$ is Schur-convex (Schur-concave).*

Proof. See theorem 3.C.1 in [4]. □

Given a set $A \subset \mathbb{R}^n$ and an element $b \in A$ we say that b is a Schur-minimum of A if and only if for all $a \in A$, a majorizes b. Clearly, if $f : A \to \mathbb{R}$ is Schur-convex (Schur-concave) and $b \in A$ is a Schur-minimum of A, then f attains a global minimum (maximum) at b.

6 Sum capacity and TSC

In this section we define two important performance measures of a given configuration. Sum capacity (C_{sum}) is defined as the maximum sum of rates at which the users can transmit and be reliably decoded at the base station. All other parameters being thought fixed, we will regard C_{sum} as a function of the signature sequences, i.e. $C_{sum} : \mathcal{S} \to \mathbb{R}$. It can be shown that ([15])

$$C_{sum}(S) = \frac{1}{2}\log\det\left(I + W^{-1}SDS^T\right) = \frac{1}{2}\log\det\left(SDS^T + W\right) - \frac{1}{2}\log\det(W) \quad (7)$$

If we use the sum capacity as a measure of performance, an optimal configuration $S \in \mathcal{S}$ is one that maximizes C_{sum}.

Given $S \in \mathcal{S}$, let $\lambda = \lambda(SDS^T + W)$. Then

$$\det(SDS^T + W) = \prod_{n=1}^{N} \lambda_n$$

and so

$$C_{sum}(S) = \frac{1}{2}\sum_{n=1}^{N} \log \lambda_n - \frac{1}{2}\log\det(W)$$

As $\log(\cdot)$ is a concave function, lemma 2 implies that $C_{sum}(S)$ is a Schur-concave function of $\lambda(SDS^T + W)$.

We define a generalized total square correlation (TSC) as a function TSC : $\mathcal{S} \to \mathbb{R}$ with ([5])

$$\text{TSC}(S) = \text{tr}\left[\left(SDS^T + W\right)^2\right] \tag{8}$$

A motivation for the choice of TSC as a performance measure is the following. First write

$$\begin{aligned} \text{TSC}(S) &= \text{tr}\left[(SDS^T)^2\right] + \text{tr}(SDS^TW) + \text{tr}(WSDS^T) + \text{tr}(W^2) \\ &= \text{tr}\left[(SDS^T)^2\right] + 2\text{tr}(SDS^TW) + \text{tr}(W^2) \end{aligned}$$

Now $SDS^T = \sum_{k=1}^{K} p_k s_k s_k^T$, hence

$$\text{tr}\left[(SDS^T)^2\right] = \sum_{k=1}^{K}\sum_{m=1}^{K} p_k p_m (s_k^T s_m)^2 = \sum_{k=1}^{K} p_k^2 + 2\sum_{k=1}^{K-1}\sum_{m=k+1}^{K} p_k p_m (s_k^T s_m)^2$$

and

$$\text{tr}(SDS^TW) = \sum_{k=1}^{K} p_k s_k^T W s_k$$

So $\text{tr}\left[(SDS^T)^2\right]$ is a weighted sum of the interference power "seen" by all users (plus the constant term $\sum_{k=1}^{K} p_k^2$), and $\text{tr}(SDS^TW)$ is a weighted sum of the noise power "seen" by all users. Hence it seems reasonable to use TSC as a performance measure; the smaller the TSC the better.

Given $S \in \mathcal{S}$ and $\lambda = \lambda(SDS^T + W)$, the matrix $\left(SDS^T + W\right)^2$ has eigenvalues

$$\lambda\left((SDS^T + W)^2\right) = \left(\lambda_1^2, \ldots, \lambda_N^2\right)$$

and therefore, as the trace of a matrix is equal to the sum of its eigenvalues,

$$\text{TSC}(S) = \sum_{n=1}^{N} \lambda_n^2$$

As $(\cdot)^2$ is a convex function, lemma 2 implies that TSC(S) is a Schur-convex function of $\lambda(SDS^T + W)$.

From now on we will focus on TSC. We will prove later in section 8 that the set $\{\lambda(SDS^T + W) : S \in \mathcal{S}\}$ has a Schur-minimum element. Therefore, as C_{sum} is Schur-concave and TSC is Schur-convex, the configurations attaining this Schur-minimum element will achieve the maximum C_{sum} and the minimum TSC. Hence the optimal configurations are the same whether we use C_{sum} or TSC as performance measure.

7 MMSE iteration

We would like to obtain configurations that attain the minimum TSC. To this end we will define an iterative procedure that, starting with some initial configuration, modifies one of the signature sequences at each iteration in a way that reduces the TSC.

For a given configuration $S \in \mathcal{S}$ we will denote the normalized MMSE linear filter for user k as $c_k(S)$. Hence by (5),

$$c_k(S) = \frac{1}{\sqrt{s_k^T \left(S_k D_k S_k^T + W\right)^{-2} s_k}} \left(S_k D_k S_k^T + W\right)^{-1} s_k \tag{9}$$

or equivalently from (6),

$$c_k(S) = \frac{1}{\sqrt{s_k^T \left(S D S^T + W\right)^{-2} s_k}} \left(S D S^T + W\right)^{-1} s_k \tag{10}$$

We define the MMSE user k update function as

$$\Phi_k(S) = \left[\begin{array}{ccccccc} s_1 & \ldots & s_{k-1} & c_k(S) & s_{k+1} & \ldots & s_K \end{array}\right] \tag{11}$$

which replaces the signature sequence for user k by the corresponding normalized linear MMSE filter. The following lemma ([10][3]) states that this update strictly decreases the TSC except when the signature sequence for user k coincides with the MMSE filter.

Lemma 3.

$$\forall S \in \mathcal{S}: \ \mathrm{TSC}(\Phi_k(S)) \leq \mathrm{TSC}(S), \qquad \textit{with equality iff } s_k = c_k(S) \tag{12}$$

Proof. By direct calculation:

$$\begin{aligned} &\mathrm{TSC}(S) - \mathrm{TSC}(\Phi_k(S)) = \\ &2 p_k s_k^T \left[S_k D_k S_k^T + W - \frac{1}{s_k^T \left(S_k D_k S_k^T + W\right)^{-2} s_k} \left(S_k D_k S_k^T + W\right)^{-1} \right] s_k \end{aligned}$$

Given a symmetric positive definite matrix M and a unit-norm vector v, we claim that $v^T M v \geq \frac{v^T M^{-1} v}{v^T M^{-2} v}$ with equality iff v is an eigenvector of M. To see this apply Cauchy-Schwartz inequality twice:

$$\begin{aligned} 1 = \left(\left\langle M^{\frac{1}{2}} v, M^{-\frac{1}{2}} v \right\rangle\right)^2 &\leq \left\|M^{\frac{1}{2}} v\right\|^2 \left\|M^{-\frac{1}{2}} v\right\|^2 = \left(v^T M v\right)\left(v^T M^{-1} v\right) \\ \left(v^T M^{-1} v\right)^2 = \left(\left\langle v, M^{-1} v \right\rangle\right)^2 &\leq \|v\|^2 \left\|M^{-1} v\right\|^2 = v^T M^{-2} v \end{aligned}$$

with equality in both inequalities iff v is an eigenvector of M, and thus we get the desired claim.

Finally we apply our claim above to $M = S_k D_k S_k^T + W$ and $v = s_k$ to obtain (12). □

[3] [10] considers the case of white noise and equal received powers, but the proof holds for arbitrary noise covariance and received user powers.

Consider the MMSE update dynamics in $\mathcal{S}$:

$$S^{(t+1)} = \Phi_{t+1}(S^{(t)}) \tag{13}$$

where we define Φ_t for $t > K$ setting $\Phi_t = \Phi_{t-K}$. This corresponds to replacing each signature sequence using the MMSE update, one at a time. We remark that this iteration is amenable for a distributed[4] implementation. The linear MMSE filter for a user can be implemented blindly ([3]), i.e. without needing knowledge of received powers or signature sequences of other users.

Note that given any initial configuration $S^{(0)} \in \mathcal{S}$, the sequence $\mathrm{TSC}(S^{(t)})$ defined by equation (13) converges because it is non-increasing by lemma 3 and bounded below.

The MMSE update function is defined as

$$\Phi(S) = \Phi_K(\Phi_{K-1}(\dots\Phi_1(S))) \tag{14}$$

Let F_Φ be the set of fixed configurations of Φ:

$$F_\Phi = \{S \in \mathcal{S} : \Phi(S) = S\} \tag{15}$$

Lemma 4. *Let $S \in \mathcal{S}$. Then*

$$\mathrm{TSC}(\Phi(S)) \leq \mathrm{TSC}(S), \quad \textit{with equality iff } S \in F_\Phi \tag{16}$$

Moreover, $S \in F_\Phi$ if and only if $\Phi_k(S) = S$ for all $k \in \{1, \dots, K\}$.

Proof. Let $S \in \mathcal{S}$. Applying lemma 3 K times we get

$$\begin{aligned} \mathrm{TSC}(S) &\geq \mathrm{TSC}(\Phi_1(S)) \geq \mathrm{TSC}(\Phi_2(\Phi_1(S))) \geq \dots \\ &\geq \mathrm{TSC}(\Phi_{K-1}(\dots\Phi_1(S))) \geq \mathrm{TSC}(\Phi(S)) \end{aligned} \tag{17}$$

If $S \notin F_\Phi$, then there is some $k \in \{1, \dots, K\}$ such that $\Phi_k(\dots\Phi_1(S)) \neq \Phi_{k-1}(\dots\Phi_1(S))$, and so by lemma 3 $\mathrm{TSC}(\Phi_k(\dots\Phi_1(S))) < \mathrm{TSC}(\Phi_{k-1}(\dots\Phi_1(S)))$. Hence using (17) $\mathrm{TSC}(S) > \mathrm{TSC}(\Phi(S))$.

If $\mathrm{TSC}(S) = \mathrm{TSC}(\Phi(S))$ then equality must hold in all inequalities in (17). From lemma 3 we get $\Phi_1(S) = S$, $\Phi_2(\Phi_1(S)) = \Phi_1(S), \dots, \Phi(S) = \Phi_{K-1}(\dots\Phi_1(S))$. Hence we obtain $S \in F_\Phi$ and also $\Phi_k(S) = S$ for all $k \in \{1, \dots, K\}$.

Next we consider the last assertion. If $\Phi_k(S) = S$ for all $k \in \{1, \dots, K\}$, then clearly $S \in F_\Phi$. Now assume $S \in F_\Phi$. Then $\mathrm{TSC}(S) = \mathrm{TSC}(\Phi(S))$, and hence as proved above $\Phi_k(S) = S$ for all $k \in \{1, \dots, K\}$. □

The following lemma and theorem provide a characterization of the fixed configurations.

Lemma 5. *Let $S = \begin{bmatrix} s_1 & \dots & s_K \end{bmatrix} \in \mathcal{S}$. Then $S \in F_\Phi$ if and only if for all $k \in \{1, \dots, K\}$, s_k is an eigenvector of $SDS^T + W$.*

[4] Here distributed means that can be implemented in parallel modules with no interaction. The user receivers are in the base station, hence co-located.

Proof. Let $S \in F_\Phi$. We can apply lemma 4 to obtain $c_k(S) = s_k$ for all $k \in \{1, \dots, K\}$ and hence from (10),

$$\left(SDS^T + W\right) s_k = \lambda_k s_k \tag{18}$$

where $\lambda_k = \left[s_k^T \left(SDS^T + W\right)^{-2} s_k\right]^{-\frac{1}{2}}$.

Conversely suppose $\left(SDS^T + W\right) s_k = \lambda_k s_k$ for all $k \in \{1, \dots, K\}$. Then using (10) $c_k(S) = s_k$, hence $\Phi_k(S) = S$ for all $k \in \{1, \dots, K\}$ and so $S \in F_\Phi$. □

Theorem 2. *Let $S \in F_\Phi$. Then*

1. *There exists an orthonormal basis of (common) eigenvectors of SDS^T and W. Equivalently, matrices SDS^T and W commute.*

2. *Let $w_1, \dots, w_N$ be the eigenvalues of W, and let $\{q_1, \dots, q_N\}$ be an orthonormal basis of eigenvectors of SDS^T and W with $Wq_n = w_n q_n$ for all $n \in \{1, \dots, N\}$. There exist $L \in \{1, \dots, N\}$, a partition $\mathcal{J}_1, \dots, \mathcal{J}_L$ (with possibly some of the $\mathcal{J}_\ell$ empty) of the set $\{1, \dots, K\}$, a partition $\mathcal{I}_1, \dots, \mathcal{I}_L$ of the set $\{1, \dots, N\}$, and positive real numbers $\mu_1 \geq \dots \geq \mu_L$ such that for all $\ell \in \{1, \dots, L\}$:*

$$(SDS^T + W)s_k = \mu_\ell s_k \quad \forall k \in \mathcal{J}_\ell \tag{19}$$

$$(SDS^T + W)q_n = \mu_\ell q_n \quad \forall n \in \mathcal{I}_\ell \tag{20}$$

$$\lambda(SDS^T + W) = (\underbrace{\mu_1, \dots, \mu_1}_{|\mathcal{I}_1|}, \dots\dots, \underbrace{\mu_L, \dots, \mu_L}_{|\mathcal{I}_L|}) \tag{21}$$

$$\mu_\ell = \frac{1}{|\mathcal{I}_\ell|}\left(\sum_{k \in \mathcal{J}_\ell} p_k + \sum_{n \in \mathcal{I}_\ell} w_n\right) \tag{22}$$

$$s_{k_1}^T s_{k_2} = 0 \quad \forall k_1 \in \mathcal{J}_\ell,\ k_2 \in \{1, \dots, K\} \setminus \mathcal{J}_\ell \tag{23}$$

$$\{s_k : k \in \mathcal{J}_\ell\} \subset \operatorname{span}\{q_n : n \in \mathcal{I}_\ell\} \tag{24}$$

and

$$\operatorname{TSC}(S) = \sum_{\ell=1}^{L} \frac{1}{|\mathcal{I}_\ell|}\left(\sum_{k \in \mathcal{J}_\ell} p_k + \sum_{n \in \mathcal{I}_\ell} w_n\right)^2 \tag{25}$$

where $|\mathcal{I}_\ell|$ is the cardinality of $\mathcal{I}_\ell$.

Proof. Let L be the number of distinct eigenvalues of $SDS^T + W$, and $\mu_1 > \dots > \mu_L$ be such eigenvalues. From lemma 5 all s_k are eigenvectors of $SDS^T + W$, so we can partition the set $\{1, \dots, K\}$ grouping the signatures associated to the same eigenvalues. I.e. if we define for $\ell \in \{1, \dots, L\}$

$$\mathcal{J}_\ell = \{k \in \{1, \dots, K\} : (SDS^T + W)s_k = \mu_\ell s_k\} \tag{26}$$

the $\mathcal{J}_\ell$ are disjoint, $\bigcup_{\ell=1}^{L} \mathcal{J}_\ell = \{1, \dots, K\}$ and equation (19) is satisfied. As $SDS^T + W$ is a symmetric matrix, eigenvectors associated with distinct eigenvalues are orthogonal and (23) is proved. Consider any $\ell \in \{1, \dots, L\}$ with $\mathcal{J}_\ell \neq \emptyset$. If we write $S_{\mathcal{J}_\ell} = [s_k, k \in \mathcal{J}_\ell]$ and $D_{\mathcal{J}_\ell} = \operatorname{diag}(p_k, k \in \mathcal{J}_\ell)$ it follows

$$(S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T + W)s_k = \mu_\ell s_k \quad \forall k \in \mathcal{J}_\ell \tag{27}$$

Multiplying (27) on the right by $p_k s_k^T$ and summing over $k \in \mathcal{J}_\ell$ we obtain

$$(S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T + W) S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T = \mu_\ell S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$$

Hence $W S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T = \mu_\ell S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T - (S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T)^2$ is a symmetric matrix, which implies that W and $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ commute. As

$$SDS^T = \sum_{\substack{\ell = 1 \\ \mathcal{J}_\ell \neq \emptyset}}^{L} S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$$

we see that W and SDS^T commute. Therefore there exists an orthonormal basis $\{q_1, \ldots, q_N\}$ of eigenvectors of W and SDS^T (see e.g. corollary 3 of theorem 3' in chapter VIII of [2]). Hence $q_1, \ldots, q_N$ are eigenvectors of $SDS^T + W$. Now choose the partition of the set $\{1, \ldots, N\}$ as follows:

$$\mathcal{I}_\ell = \{n \in \{1, \ldots, N\} : \ (SDS^T + W)\, q_n = \mu_\ell q_n\}$$

Then (20) is satisfied and (21) follows. Fix $\ell \in \{1, \ldots, L\}$ and let $n \in \mathcal{I}_\ell$ and $k \in \{1, \ldots, K\} \setminus \mathcal{J}_\ell$. Then q_n and s_k are eigenvectors of $SDS^T + W$ associated with distinct eigenvalues and hence are orthogonal. Therefore

$$\mu_\ell q_n = (SDS^T + W) q_n = (S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T + W) q_n$$

and $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = (\mu_\ell - w_n) q_n$. By convention we will take $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ as the $N \times N$ zero matrix when $\mathcal{J}_\ell = \emptyset$. Note that with this convention, the previous equations hold even for such ℓ's. If we consider $n \in \{1, \ldots, N\} \setminus \mathcal{I}_\ell$, as for all $k \in \mathcal{J}_\ell$, q_n and s_k are orthogonal, $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = 0$. So we can write

$$S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = \begin{cases} (\mu_\ell - w_n) q_n & \text{if } n \in \mathcal{I}_\ell \\ 0 & \text{if } n \in \{1, \ldots, N\} \setminus \mathcal{I}_\ell \end{cases} \tag{28}$$

Multiplying on the right by q_n^T and summing over all $n \in \{1, \ldots, N\}$ we get

$$S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T = \sum_{n \in \mathcal{I}_\ell} (\mu_\ell - w_n) q_n q_n^T \tag{29}$$

where we have used the fact that $\sum_{n=1}^N q_n q_n^T$ is the $N \times N$ identity matrix because the q_n form an orthonormal basis. For $k \in \mathcal{J}_\ell$, using the same identity and that s_k is orthogonal to q_n for $n \notin \mathcal{I}_\ell$, we obtain

$$s_k = \sum_{n=1}^{N} q_n q_n^T s_k = \sum_{n \in \mathcal{I}_\ell} q_n q_n^T s_k$$

which proves (24).

Now, as for any matrices A and B of appropriate dimensions $\text{tr}(AB) = \text{tr}(BA)$,

$$\text{tr}(S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T) = \text{tr}(D_{\mathcal{J}_\ell}^{\frac{1}{2}} S_{\mathcal{J}_\ell}^T S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell}^{\frac{1}{2}}) = \sum_{k \in \mathcal{J}_\ell} p_k$$

where we have used the fact that the diagonal elements of $S_{\mathcal{J}_\ell}^T S_{\mathcal{J}_\ell}$ are all 1 because the signatures are unit-norm. Also $\text{tr}(q_n q_n^T) = q_n^T q_n = 1$, so from equation (29) we obtain (22). Equation (25) is obtained noting from (21) that $\text{TSC}(S) = \sum_{\ell=1}^{L} |\mathcal{I}_\ell| \mu_\ell^2$. □

We remark that the characterization obtained in the proof of theorem 2 may in general not be the only one satisfying (19)-(25). As an example, let $K = 2$, $N = 2$, $p_1 = p_2 = 4$, $W = \text{diag}(1, 9)$ and

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Then $SDS^T + W = 9I$ and hence, by lemma 5, S is a fixed configuration. The characterization obtained in the proof of theorem 2 is $L = 1$, $\mu_1 = 9$, $\mathcal{J}_1 = \{1, 2\}$, $\mathcal{I}_1 = \{1, 2\}$. Another characterization which verifies (19)-(25) is $L = 2$, $\mu_1 = \mu_2 = 9$, $\mathcal{J}_1 = \{1, 2\}$, $\mathcal{J}_2 = \emptyset$, $\mathcal{I}_1 = \{1\}$, $\mathcal{I}_2 = \{2\}$.

The characterization obtained in the proof of theorem 2 is clearly the most economical one in the sense that L is as small as possible (because all μ's are distinct). However we will find it convenient to use the characterization of the fixed configurations as in the following lemma.

Lemma 6. *Let $S \in F_\Phi$. Then there exists a characterization as in theorem 2 satisfying equations (19)-(25) that also verifies the following for all $\ell \in \{1, \ldots, L\}$:*

1. *If $\mathcal{J}_\ell \neq \emptyset$ then $|\mathcal{J}_\ell| \geq |\mathcal{I}_\ell|$ and for all $n \in \mathcal{I}_\ell$, $\mu_\ell > w_n$.*
2. *If $\mathcal{J}_\ell = \emptyset$ then $|\mathcal{I}_\ell| = 1$.*
3. *If $\ell < L$ and $\mathcal{J}_\ell \neq \emptyset$ then $\mu_\ell > \mu_{\ell+1}$.*

Proof. Take the partitions in the proof of theorem 2. Consider any $\ell \in \{1, \ldots, L\}$ with $\mathcal{J}_\ell \neq \emptyset$, and any $n \in \mathcal{I}_\ell$. From equation (28),

$$S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = (\mu_\ell - w_n) q_n$$

As $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ is nonnegative definite, $\mu_\ell \geq w_n$.

Assume $\mu_\ell = w_n$. Then $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = 0$. This implies $q_n^T S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = \left\| D_{\mathcal{J}_\ell}^{\frac{1}{2}} S_{\mathcal{J}_\ell}^T q_n \right\|^2 = 0$ and hence, as $D_{\mathcal{J}_\ell}$ is invertible, $S_{\mathcal{J}_\ell}^T q_n = 0$. Therefore, q_n is orthogonal to the signature sequences of all users in $\mathcal{J}_\ell$. Define:

$$\begin{aligned} \mathcal{J}'_\ell &= \emptyset \\ \mathcal{J}''_\ell &= \mathcal{J}_\ell \\ \mathcal{I}'_\ell &= \{n \in \mathcal{I}_\ell : \mu_\ell = w_n\} \\ \mathcal{I}''_\ell &= \{n \in \mathcal{I}_\ell : \mu_\ell > w_n\} \end{aligned}$$

Note that $|\mathcal{J}''_\ell| \geq |\mathcal{I}''_\ell|$ because $\{q_n : n \in \mathcal{I}''_\ell\}$ are orthonormal eigenvectors of $S_{\mathcal{J}''_\ell} D_{\mathcal{J}''_\ell} S^T_{\mathcal{J}''_\ell}$ associated with nonzero eigenvalues, and hence $S_{\mathcal{J}''_\ell}$ has rank $|\mathcal{I}''_\ell|$ and $|\mathcal{J}''_\ell|$ columns.

A new characterization satisfying (19)-(25) (with L increased by $|\mathcal{I}'_\ell|$) is obtained by dividing $(\mathcal{J}_\ell, \mathcal{I}_\ell)$ in $|\mathcal{I}'_\ell|+1$ parts: $(\mathcal{J}''_\ell, \mathcal{I}''_\ell)$ and for each $n \in \mathcal{I}'_\ell$, $(\mathcal{J}''_\ell, \{n\})$.

If we do the same for all ℓ for which there is at least one $n \in \mathcal{I}_\ell$ with $w_n = \mu_\ell$, we obtain the desired result. Let $\hat{L}$, $(\hat{\mathcal{J}}_1, \ldots, \hat{\mathcal{J}}_{\hat{L}})$, $(\hat{\mathcal{I}}_1, \ldots, \hat{\mathcal{I}}_{\hat{L}})$, $\hat{\mu}_1 \geq \ldots \geq \hat{\mu}_{\hat{L}}$ be the new characterization. Note that in our construction given any λ there can be at most one ℓ with $\hat{\mathcal{J}}_\ell \neq \emptyset$ and $\hat{\mu}_\ell = \lambda$. Hence condition *3* is satisfied ordering the partitions so that if $\hat{\mu}_\ell = \hat{\mu}_{\ell+1}$ then $\hat{\mathcal{J}}_\ell = \emptyset$. □

Given $S^{(0)} \in \mathcal{S}$ we can define the ω-limit set ([8]) with respect to the dynamics (13) as:

$$\omega_\Phi(S^{(0)}) = \{S \in \mathcal{S} : \exists t_1 < t_2 < \ldots \text{ s.t. } \lim_{m\to\infty} S^{(t_m)} = S\} \tag{30}$$

In words, $\omega_\Phi(S^{(0)})$ is the set of all limit points of the trajectory $S^{(t)}$.

The following lemma shows that for any initial set of signature sequences, the MMSE iteration (13) converges to the set of fixed configurations.

Lemma 7. *Given any* $S^{(0)} \in \mathcal{S}$,

$$\omega_\Phi(S^{(0)}) \subset F_\Phi \tag{31}$$

Proof. If $S \in \omega_\Phi(S^{(0)})$ then $\exists t_1 < t_2 < \ldots$ s.t. $\lim_{m\to\infty} S^{(t_m)} = S$. For some $k \in \{1, \ldots, K\}$, t_m is a multiple of k for infinitely many m, let t'_m be the corresponding subsequence. Then $S^{(t'_m)} \to S$ as $m \to \infty$. By continuity of Φ_{k+1}, $\Phi_{k+1}(S^{(t'_m)}) \to \Phi_{k+1}(S)$ as $m \to \infty$.

Now assume $\Phi_{k+1}(S) \neq S$. Then by lemma 3, $\text{TSC}(\Phi_{k+1}(S)) < \text{TSC}(S)$. Let $\Delta = \text{TSC}(S) - \text{TSC}(\Phi_{k+1}(S))$. Then, as TSC is continuous, there exists p such that $\forall m > p$ it is $\text{TSC}(S^{(t'_m+1)}) < \text{TSC}(S^{(t'_m)}) - \frac{\Delta}{2}$. Thus $\text{TSC}(S^{(t'_{m+1})}) < \text{TSC}(S^{(t'_m)}) - \frac{\Delta}{2}$ for $p > m$ and therefore $\text{TSC}(S^{(t'_m)}) \to -\infty$ as $m \to \infty$. This is a contradiction because TSC is positive, and thus $\Phi_{k+1}(S) = S$.

But then $\Phi_{k+1}(S^{(t'_m)}) = S^{(t'_m+1)} \to \Phi_{k+1}(S) = S$ as $m \to \infty$. Recurring to the same argument as before we now get $\Phi_{k+2}(S) = S$. Repeating this argument $(K-2)$ more times we get $\Phi(S) = S$ as we wanted to prove. □

We conclude that for any initial condition the MMSE iteration approaches the set of fixed configurations as $t \to \infty$. As TSC is a continuous function, this implies that $\lim_{t\to\infty} \text{TSC}(S^{(t)}) \in T_F$ where

$$T_F = \{\text{TSC}(S) : S \in F_\Phi\} \tag{32}$$

Note that from theorem 2, T_F has a finite number of elements because there is a finite number of ways of partitioning the sets $\{1, \ldots, K\}$ and $\{1, \ldots, N\}$. A loose upper bound on $|T_F|$ can be found by noting that for a given L, there are less than L^N ways of partitioning the set $\{1, \ldots, N\}$ in L subsets: for each element in $\{1, \ldots, N\}$, we can choose one of the L subsets in the partition to put that element. Analogously, there are at most L^K ways of partitioning the set $\{1,\dots,K\}$ in L subsets. Hence, as $L \in \{1,\dots,N\}$,

$$|T_F| \le \sum_{L=1}^{N} L^{K+N} \tag{33}$$

Let τ be the minimum of the TSC:

$$\tau = \min_{S\in\mathcal{S}} \mathrm{TSC}(S) \tag{34}$$

As $\mathcal{S}$ is a compact set and TSC is continuous, the minimum is attained and we can define the set of optimal configurations:

$$\Omega = \{S \in \mathcal{S} : \mathrm{TSC}(S) = \tau\} \tag{35}$$

Clearly we have $\Omega \subset F_\Phi$: if $S \in \Omega$ then $\tau = \mathrm{TSC}(S) \ge \mathrm{TSC}(\Phi(S))$ and by lemma 4 $\mathrm{TSC}(S) \le \mathrm{TSC}(\Phi(S))$, and therefore $\mathrm{TSC}(\Phi(S)) = \mathrm{TSC}(S)$ which again by lemma 4 implies $S \in F_\Phi$. But it is easy to see that F_Φ contains non-optimal configurations, i.e. $F_\Phi \ne \Omega$ except for the trivial case $N = 1$. As an example take $N \ge 2$ and let $w_1 \le \dots \le w_N$ be the ordered eigenvalues of W, and $q_1,\dots,q_N$ be an orthogonal basis of associated eigenvectors. Then, if we take $s_k = q_N$ for all $k \in \{1,\dots,K\}$ we obtain a fixed configuration $S \in F_\Phi$. It is easy to see that if $s'_1 = q_1$ and $s'_k = q_N$ for $k \in \{2,\dots,K\}$, the new configuration S' attains a lower TSC value: $\mathrm{TSC}(S') < \mathrm{TSC}(S)$. Hence $S \notin \Omega$. Actually, S attains the global maximum of the TSC over $\mathcal{S}$.

Therefore, for $N \ge 2$, the set T_F has more than one element and we cannot conclude that $\lim_{t\to\infty} \mathrm{TSC}(S^{(t)}) = \tau$ as we would like. Simulations suggest that if the initial condition $S^{(0)}$ is chosen randomly, then $\mathrm{TSC}(S^{(t)})$ converges to τ with probability one ([10]), but no formal proof has been given.

8 Global optimal configurations

We have seen in the previous section that the global minimum of the TSC over all configurations $S \in \mathcal{S}$ is attained for some fixed configuration of the MMSE update $S \in F_\Phi$, i.e.

$$\min_{S\in\mathcal{S}} \mathrm{TSC}(S) = \min_{S\in F_\Phi} \mathrm{TSC}(S)$$

Any fixed configuration is associated with a partition of the set of users and a partition of the set of signal dimensions as shown in theorem 2. Conversely, given such a pair of partitions, we could try to find a corresponding configuration $S \in F_\Phi$. This is not always feasible, as the following simple example shows.

Let $K = 2$, $N = 2$, $p_1 = p_2 = 1$, $w_1 = 3$ and $w_2 = 0.2$. Consider $L = 1$, $\mathcal{J}_1 = \{1,2\}$ and $\mathcal{I}_1 = \{1,2\}$. For this partition pair we should have according to theorem 2 that $SDS^T + W$ has eigenvalue $\mu_1 = 2.6$ with multiplicity 2 (i.e. $SDS^T + W$ is 2.6 times the 2×2 identity matrix). But, being SDS^T and W symmetric and nonnegative definite, the maximum eigenvalue of $SDS^T + W$ has to be at least as large as the maximum eigenvalue of W, w_1. As $2.6 < 3$ we see that it is not possible to find s_1 and s_2 such that $SDS^T + W = 2.6I$ and hence the proposed partition pair is not feasible.

The following lemma characterizes the feasible partition pairs.

Lemma 8. *Let $\{q_1, \ldots, q_N\}$ be an orthonormal basis of eigenvectors of W respectively associated with eigenvalues $w_1, \ldots, w_N$. Suppose we are given $L \in \{1, \ldots, N\}$, real numbers $\mu_1 \geq \ldots \geq \mu_L$, a partition $\mathcal{J}_1, \ldots \mathcal{J}_L$ of $\{1, \ldots, K\}$ (with possibly some $\mathcal{J}_\ell$ empty) and a partition $\mathcal{I}_1, \ldots, \mathcal{I}_L$ of $\{1, \ldots, N\}$ with*

$$\mu_\ell = \frac{1}{|\mathcal{I}_\ell|}\left(\sum_{k \in \mathcal{J}_\ell} p_k + \sum_{n \in \mathcal{I}_\ell} w_n\right)$$

Then following are equivalent:

1. *There exists a configuration $S \in \mathcal{S}$ satisfying equations (19)-(25).*
2. *For each $\ell \in \{1, \ldots, L\}$,*

$$\mu_\ell \geq \max\left(\{w_n : n \in \mathcal{I}_\ell\} \cup \left\{\frac{1}{M}\sum_{m=1}^{M}\left(p_m^\ell + w_m^\ell\right) : M \in \{1, \ldots, \min(|\mathcal{I}_\ell|, |\mathcal{J}_\ell|)\}\right\}\right) \tag{36}$$

where p_m^ℓ is the m-th largest component of $(p_k : k \in \mathcal{J}_\ell)$ and w_m^ℓ is the m-th smallest component of $(w_n : n \in \mathcal{I}_\ell)$.

Proof.

(*1* $\Rightarrow$ *2*) Consider any $\ell \in \{1, \ldots, L\}$. Using (20) and (24), as $\{q_1, \ldots, q_N\}$ is an orthonormal set, we can obtain as in the proof of theorem 2 (see (28)),

$$S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = \begin{cases} (\mu_\ell - w_n) q_n & \text{if } n \in \mathcal{I}_\ell \\ 0 & \text{if } n \in \{1, \ldots, N\} \setminus \mathcal{I}_\ell \end{cases}$$

This implies that $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ has eigenvalues (in non-increasing order)

$$\lambda(S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T) = (\mu_\ell - w_1^\ell, \ldots, \mu_\ell - w_{|\mathcal{I}_\ell|}^\ell, \underbrace{0, \ldots, 0}_{N - |\mathcal{I}_\ell|})$$

As $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ is nonnegative definite, all eigenvalues must be nonnegative and hence $\mu_\ell \geq \max\{w_n : n \in \mathcal{I}_\ell\}$.

Consider the $|\mathcal{J}_\ell| \times |\mathcal{J}_\ell|$ matrix $D_{\mathcal{J}_\ell}^{\frac{1}{2}} S_{\mathcal{J}_\ell}^T S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell}^{\frac{1}{2}}$. It has the same nonzero eigenvalues as $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell}^{\frac{1}{2}} D_{\mathcal{J}_\ell}^{\frac{1}{2}} S_{\mathcal{J}_\ell}^T = S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ and the diagonal elements are $(p_k : k \in \mathcal{J}_\ell)$. From theorem 1 we obtain that

$$(\mu_\ell - w_1^\ell, \ldots, \mu_\ell - w_{\min(|\mathcal{I}_\ell|, |\mathcal{J}_\ell|)}^\ell, \underbrace{0, \ldots, 0}_{|\mathcal{J}_\ell| - \min(|\mathcal{I}_\ell|, |\mathcal{J}_\ell|)}) \text{ majorizes } (p_1^\ell, \ldots, p_{|\mathcal{J}_\ell|}^\ell)$$

where for convenience we have written the eigenvalues and diagonal elements in non-increasing order. The above majorization relation implies that for all $M \in \{1, \dots, \min(|\mathcal{I}_\ell|, |\mathcal{J}_\ell|)\}$ we have

$$\sum_{m=1}^{M}(\mu_\ell - w_m^\ell) \geq \sum_{m=1}^{M} p_m^\ell$$

or equivalently

$$\mu_\ell \geq \frac{1}{M}\sum_{m=1}^{M}(p_m^\ell + w_m^\ell)$$

($2 \Rightarrow 1$) Fix any $\ell \in \{1, \dots, L\}$. Define

$$\mathcal{I}'_\ell = \{n \in \mathcal{I}_\ell : \mu_\ell > w_n\}$$

From (36) $\mu_\ell \geq w_n$ for all $n \in \mathcal{I}_\ell$, thus for $n \in \mathcal{I}_\ell \setminus \mathcal{I}'_\ell$ we have $\mu_\ell = w_n$. Also $|\mathcal{I}'_\ell| \leq |\mathcal{J}_\ell|$. Otherwise we would have

$$\mu_\ell \geq \frac{1}{|\mathcal{J}_\ell|}\sum_{m=1}^{|\mathcal{J}_\ell|}(p_m^\ell + w_m^\ell) = \frac{1}{|\mathcal{J}_\ell|}\left(\sum_{k\in\mathcal{J}_\ell} p_k + \sum_{m=1}^{|\mathcal{J}_\ell|} w_m^\ell\right)$$

and $\mu_\ell > w^\ell_{|\mathcal{J}_\ell|+1}$ hence we would get

$$\mu_\ell > \frac{1}{|\mathcal{I}_\ell|}\left(\sum_{k\in\mathcal{J}_\ell} p_k + \sum_{n\in\mathcal{I}_\ell} w_n\right)$$

which is a contradiction.

Equation (36) also implies that

$$(\mu_\ell - w_1^\ell, \dots, \mu_\ell - w^\ell_{|\mathcal{I}'_\ell|}, \underbrace{0, \dots, 0}_{|\mathcal{J}_\ell| - |\mathcal{I}'_\ell|}) \text{ majorizes } (p_1^\ell, \dots, p^\ell_{|\mathcal{J}_\ell|})$$

Hence from theorem 1 we can find a symmetric matrix $H_\ell \in \mathbb{R}^{|\mathcal{J}_\ell| \times |\mathcal{J}_\ell|}$ with eigenvalues

$$\lambda(H_\ell) = (\mu_\ell - w_1^\ell, \dots, \mu_\ell - w^\ell_{|\mathcal{I}'_\ell|}, \underbrace{0, \dots, 0}_{|\mathcal{J}_\ell| - |\mathcal{I}'_\ell|})$$

and diagonal elements $(p_k : k \in \mathcal{J}_\ell)$. As H_ℓ is symmetric and has $|\mathcal{I}'_\ell|$ nonzero eigenvalues, it can be written as $H_\ell = V_\ell \Lambda_\ell V_\ell^T$ where $\Lambda_\ell = \mathrm{diag}(\mu_\ell - w_n : n \in \mathcal{I}'_\ell)$ and $V_\ell \in \mathbb{R}^{|\mathcal{J}_\ell| \times |\mathcal{I}'_\ell|}$ satisfies $V_\ell^T V_\ell = I$. Let $S_{\mathcal{J}_\ell} = Q_{\mathcal{I}'_\ell} \Lambda_\ell^{\frac{1}{2}} V_\ell^T D_{\mathcal{J}_\ell}^{-\frac{1}{2}}$, where $Q_{\mathcal{I}'_\ell} = [q_n, n \in \mathcal{I}'_\ell]$. Now using $Q^T_{\mathcal{I}'_\ell} Q_{\mathcal{I}'_\ell} = I$ we get $S^T_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell} = D_{\mathcal{J}_\ell}^{-\frac{1}{2}} H_\ell D_{\mathcal{J}_\ell}^{-\frac{1}{2}}$ has unit diagonal entries, so $S_{\mathcal{J}_\ell}$ has unit norm columns. For $k \in \mathcal{J}_\ell$ take s_k as the corresponding column of $S_{\mathcal{J}_\ell}$. The columns of $S_{\mathcal{J}_\ell}$ are linear combinations of the columns of $Q_{\mathcal{I}'_\ell}$ and $\mathcal{I}'_\ell \subset \mathcal{I}_\ell$, hence (24) is verified. For $n \in \mathcal{I}_\ell$,

$$\begin{aligned}(SDS^T + W)q_n &= \sum_{\ell'=1}^{L} Q_{\mathcal{I}'_{\ell'}} \Lambda_{\ell'} Q^T_{\mathcal{I}'_{\ell'}} q_n + w_n q_n \\ &= Q_{\mathcal{I}'_\ell} \Lambda_\ell Q^T_{\mathcal{I}'_\ell} q_n + w_n q_n \\ &= (\mu_\ell - w_n) q_n + w_n q_n = \mu_\ell q_n\end{aligned}$$

and so (20) and (21) are satisfied. Equations (19), (23) and (25) are readily verified.

□

Hence the problem of minimizing TSC(S) over $S \in \mathcal{S}$ is equivalent to minimizing the expression (25) over all partition pairs that satisfy (36). Next we present an algorithm proposed in [14] that solves this optimization problem.

Without loss of generality, from now on we will assume p_k and w_n are ordered so that $p_1 \geq p_2 \geq \ldots \geq p_K$ and $w_1 \leq w_2 \leq \ldots w_N$.

Algorithm 1 ($\mathcal{A}$).

Input $K, N, (p_1, \ldots, p_K), (w_1, \ldots, w_N)$.

Output L, $(\mathcal{J}_1, \ldots, \mathcal{J}_L)$, $(\mathcal{I}_1, \ldots, \mathcal{I}_L)$, $(\mu_1, \ldots, \mu_L)$

Call syntax

$$[L, (\mathcal{J}_1, \ldots, \mathcal{J}_L), (\mathcal{I}_1, \ldots, \mathcal{I}_L), (\mu_1, \ldots, \mu_L)] = \mathcal{A}\left(K, N, (p_1, \ldots, p_K), (w_1, \ldots, w_N)\right)$$

Update

1. If $N = 0$ then let $L = 0$ and exit.

2. Let

$$\mu_1 = \max\left(\left\{w_N, \frac{1}{N}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right)\right\} \cup \left\{\frac{1}{M}\sum_{m=1}^{M}(p_m + w_m) : M \in \{1, \ldots, \min(N-1, K)\}\right\}\right) \tag{37}$$

3. (a) If $\mu_1 = w_N$ then:

- *Let $\mathcal{J}_1 = \emptyset, \mathcal{I}_1 = \{N\}$.*
- *Call*

$$[L', (\mathcal{J}'_1, \ldots, \mathcal{J}'_{L'}), (\mathcal{I}'_1, \ldots, \mathcal{I}'_{L'}), (\mu'_1, \ldots, \mu'_{L'})] = \mathcal{A}\left(K, N-1, (p_1, \ldots, p_K), (w_1, \ldots, w_{N-1})\right)$$

- *Let* $\hat{M} = 0$.

(b) *Else if* $\mu_1 = \frac{1}{N}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right)$ *then:*
- *Let* $\mathcal{J}_1 = \{1, \ldots, K\}, \mathcal{I}_1 = \{1, \ldots, N\}$.
- *Let* $L' = 0$.
- *Let* $\hat{M} = 0$.

(c) *Else if* $\mu_1 = \frac{1}{M}\sum_{m=1}^{M}(p_m + w_m)$ *for some* $M \in \{1, \ldots, \min(N-1, K)\}$ *then:*
- *Let* $\hat{M}$ *be the maximum such* M.
- *Let* $\mathcal{J}_1 = \left\{1, \ldots, \hat{M}\right\}, \mathcal{I}_1 = \left\{1, \ldots, \hat{M}\right\}$.
- *Call*

$$[L', (\mathcal{J}'_1, \ldots, \mathcal{J}'_{L'}), (\mathcal{I}'_1, \ldots, \mathcal{I}'_{L'}), (\mu'_1, \ldots, \mu'_{L'})] = \mathcal{A}(K-\hat{M}, N-\hat{M}, (p_{\hat{M}+1}, \ldots, p_K), (w_{\hat{M}+1}, \ldots, w_N))$$

4. *Let* $L = L' + 1$.
5. *For all* $\ell \in \{2, \ldots, L\}$, *let*

$$\begin{aligned} \mu_\ell &= \mu'_{\ell-1} \\ \mathcal{J}_\ell &= \mathcal{J}'_{\ell-1} + \hat{M} \\ \mathcal{I}_\ell &= \mathcal{I}'_{\ell-1} + \hat{M} \end{aligned}$$

where $\mathcal{J}'_{\ell-1} + \hat{M} = \{k + \hat{M} : k \in \mathcal{J}'_{\ell-1}\}$ *and analogously for* $\mathcal{I}'_{\ell-1} + \hat{M}$.

6. *Exit.*

We first state a simple fact about the output of algorithm 1.

Lemma 9. *Let*

$$[L, (\mathcal{J}_1, \ldots, \mathcal{J}_L), (\mathcal{I}_1, \ldots, \mathcal{I}_L), (\mu_1, \ldots, \mu_L)] = \mathcal{A}(K, N, (p_1, \ldots, p_K), (w_1, \ldots, w_N))$$

Then $\mu_1 \geq \ldots \geq \mu_L$.

Proof. By the recursive nature of algorithm 1, we only need to prove that if $L \geq 2$ then $\mu_1 \geq \mu_2$. Note that if $\mu_1 = \frac{1}{N}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right)$ and $\mu_1 > w_N$ then $L = 1$. Therefore there are only six possibilities:

- $\mu_1 = w_N$ and:
  - $\mu_2 = w_{N-1}$. Then clearly $\mu_1 \geq \mu_2$.
  - $\mu_2 = \frac{1}{N-1}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N-1} w_n\right)$. From (37)

$$\mu_1 \geq \frac{1}{N}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right) = \frac{1}{N}\mu_1 + \frac{N-1}{N}\mu_2$$

    This implies $\mu_1 \geq \mu_2$.

– $\mu_2 = \frac{1}{M}\sum_{m=1}^{M}(p_m + w_m)$ for some $M \in \{1,\ldots,\min(N-2,K)\}$. From (37) $\mu_1 \geq \frac{1}{M}\sum_{m=1}^{M}(p_m + w_m)$, hence $\mu_1 \geq \mu_2$.

- $\mu_1 = \frac{1}{M_1}\sum_{m=1}^{M_1}(p_m + w_m)$ for some $M_1 \in \{1,\ldots,\min(N-1,K)\}$ and:

  – $\mu_2 = w_N$. From (37) $\mu_1 \geq w_N$, i.e. $\mu_1 \geq \mu_2$.

  – $\mu_2 = \frac{1}{N-M_1}\left(\sum_{k=M_1+1}^{K} p_k + \sum_{n=M_1+1}^{N} w_n\right)$. From (37)

$$\mu_1 \geq \frac{1}{N}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right) = \frac{M_1}{N}\mu_1 + \frac{N-M_1}{N}\mu_2$$

  This implies $\mu_1 \geq \mu_2$.

  – $\mu_2 = \frac{1}{M_2}\sum_{m=M_1+1}^{M_1+M_2}(p_m + w_m)$ for some $M_2 \in \{1,\ldots,\min(N-M_1-1,K-M_1)\}$. From (37)

$$\mu_1 \geq \frac{1}{M_1+M_2}\sum_{m=1}^{M_1+M_2}(p_m + w_m) = \frac{M_1}{M_1+M_2}\mu_1 + \frac{M_2}{M_1+M_2}\mu_2$$

  Thus $\mu_1 \geq \mu_2$.

□

As proved in the following lemma, the partitions output by algorithm 1 satisfy conditions (36) and therefore we can construct a configuration S corresponding to this pair of partitions.

Lemma 10. *Let*

$$[L, (\mathcal{J}_1,\ldots,\mathcal{J}_L), (\mathcal{I}_1,\ldots,\mathcal{I}_L), (\mu_1,\ldots,\mu_L)] = \mathcal{A}(K, N, (p_1,\ldots,p_K), (w_1,\ldots,w_N))$$

There exists $S \in F_\Phi$ such that equations (19)-(25) are satisfied. In particular,

$$\lambda(SDS^T + W) = (\underbrace{\mu_1,\ldots,\mu_1}_{|\mathcal{I}_1|},\ldots\ldots,\underbrace{\mu_L,\ldots,\mu_L}_{|\mathcal{I}_L|})$$

Proof. We use lemma 8. By the recursive nature of algorithm 1 we only need to prove (36) holds for $\mu_1, \mathcal{J}_1, \mathcal{I}_1$. It is straightforward to see that

$$\mu_1 = \frac{1}{|\mathcal{I}_1|}\left(\sum_{k\in\mathcal{J}_1} p_k + \sum_{n\in\mathcal{I}_1} w_n\right)$$

and by (37), equation (36) is satisfied for $\ell = 1$. □

Our next goal is to prove that such an S corresponds to a vector of eigenvalues of $SDS^T + W$ which is a Schur-minimum of the set of vectors of eigenvalues of $S'DS'^T + W$ over $S' \in \mathcal{S}$.

Definition 1. *We will say a characterization as in lemma 6 is* efficient *if for all* $\ell_1 < \ell_2 \in \{1, \ldots, L\}$ *the following conditions are satisfied:*

1. $|\mathcal{J}_{\ell_1}| \leq |\mathcal{I}_{\ell_1}|$.
2. $p_{k_1} \geq p_{k_2}$ *for all* $k_1 \in \mathcal{J}_{\ell_1}$ *and* $k_2 \in \mathcal{J}_{\ell_2}$.
3. *If* $\mathcal{J}_{\ell_1} \neq \emptyset$ *then* $w_{n_1} \leq w_{n_2}$ *for all* $n_1 \in \mathcal{I}_{\ell_1}$ *and* $n_2 \in \mathcal{I}_{\ell_2}$.

Lemma 11. *The characterization output by the algorithm 1 is* efficient.

Proof. Follows directly from algorithm 1. □

Lemma 12. *For all* efficient *characterizations, given any* $\ell' \in \{1, \ldots, L-1\}$ *there exist* $M \in \{1, \ldots, \min(N-1, K)\}$ *and* $R \in \{0, \ldots, N-M-1\}$ *such that*

$$\sum_{\ell=1}^{\ell'} \mu_\ell |\mathcal{I}_\ell| = \sum_{m=1}^{M} (p_m + w_m) + \sum_{r=0}^{R-1} w_{N-r} \tag{38}$$

and

$$M + R = \sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell| \tag{39}$$

Proof. Consider any $\ell' \in \{1, \ldots, L-1\}$. From (22),

$$\sum_{\ell=1}^{\ell'} \mu_\ell |\mathcal{I}_\ell| = \sum_{\ell=1}^{\ell'} \left(\sum_{k \in \mathcal{J}_\ell} p_k + \sum_{n \in \mathcal{I}_\ell} w_n \right)$$

Define

$$\begin{aligned} \mathcal{J} &= \bigcup_{\ell=1}^{\ell'} \mathcal{J}_\ell \\ \mathcal{I} &= \bigcup_{\ell=1}^{\ell'} \mathcal{I}_\ell \end{aligned}$$

Then

$$\sum_{\ell=1}^{\ell'} \mu_\ell |\mathcal{I}_\ell| = \sum_{k \in \mathcal{J}} p_k + \sum_{n \in \mathcal{I}} w_n$$

Define

$$\begin{aligned} \mathcal{L} &= \{\ell \in \{1, \ldots, \ell'\} : \mathcal{J}_\ell \neq \emptyset\} \\ \mathcal{I}' &= \bigcup_{\ell \in \mathcal{L}} \mathcal{I}_\ell \\ \mathcal{I}'' &= \mathcal{I} \setminus \mathcal{I}' \end{aligned}$$

Hence

$$\sum_{\ell=1}^{\ell'} \mu_\ell |\mathcal{I}_\ell| = \sum_{k\in\mathcal{J}} p_k + \sum_{n\in\mathcal{I}'} w_n + \sum_{n\in\mathcal{I}''} w_n \tag{40}$$

Consider $\ell \in \mathcal{L}$. As $\mathcal{J}_\ell \neq \emptyset$, $|\mathcal{J}_\ell| \geq |\mathcal{I}_\ell|$ (see condition *1* in lemma 6). As $\ell \leq \ell' < L$, by condition *1* in definition 1, $|\mathcal{J}_\ell| \leq |\mathcal{I}_\ell|$. Therefore $|\mathcal{J}_\ell| = |\mathcal{I}_\ell|$. This implies $|\mathcal{J}| = |\mathcal{I}'|$. Let $M = |\mathcal{I}'|$ and $R = |\mathcal{I}''|$. Clearly $M + R = |\mathcal{I}| = \sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell|$, so (39) is verified.

As $|\mathcal{J}| = M$ (recall $p_1 \geq \ldots \geq p_K$),

$$\sum_{k\in\mathcal{J}} p_k \leq \sum_{m=1}^{M} p_m$$

Assume the above inequality is strict. This implies that there exist $k \in \mathcal{J}$ and $m \in \{1, \ldots, M\} \setminus \mathcal{J}$ with $p_k < p_m$. But then $k \in \mathcal{J}_{\ell_1}$ for some $\ell_1 \leq \ell'$ and $m \in \mathcal{J}_{\ell_2}$ for some $\ell_2 > \ell'$, which contradicts condition *2* of definition 1. Therefore

$$\sum_{k\in\mathcal{J}} p_k = \sum_{m=1}^{M} p_m \tag{41}$$

As $|\mathcal{I}''| = R$ (recall $w_1 \leq \ldots \leq w_N$),

$$\sum_{n\in\mathcal{I}''} w_n \leq \sum_{r=0}^{R-1} w_{N-r}$$

Assume the above inequality is strict. This implies that there exist $n \in \mathcal{I}''$ and $m \in \{N - R + 1, \ldots, N\} \setminus \mathcal{I}''$ with $w_n < w_m$. Let $\ell_1, \ell_2 \in \{1, \ldots, L\}$ with $m \in \mathcal{I}_{\ell_1}$ and $n \in \mathcal{I}_{\ell_2}$. As $n \in \mathcal{I}''$, we have $\ell_2 \leq \ell'$, $\mathcal{J}_{\ell_2} = \emptyset$ and $\mathcal{I}_{\ell_2} = \{n\}$ (see lemma 6). Then $\mu_{\ell_2} = w_n$. As $m \in \mathcal{I}_{\ell_1}$, $\mu_{\ell_1} \geq w_m$. Therefore $\mu_{\ell_1} \geq w_m > w_n = \mu_{\ell_2}$. Hence (recall $\mu_1 \geq \ldots \geq \mu_L$) $\ell_1 < \ell_2$. So $\ell_1 < \ell'$ and as $m \notin \mathcal{I}''$ we must have $\ell_1 \in \mathcal{L}$, i.e. $\mathcal{J}_{\ell_1} \neq \emptyset$. But then by condition *3* of definition 1 we should have $w_m \leq w_n$, a contradiction. Therefore

$$\sum_{n\in\mathcal{I}''} w_n = \sum_{r=0}^{R-1} w_{N-r} \tag{42}$$

As $|\mathcal{I}'| = M$ (recall $w_1 \leq \ldots \leq w_N$),

$$\sum_{n\in\mathcal{I}'} w_n \geq \sum_{m=1}^{M} w_m$$

Assume the above inequality is strict. This implies that there exist $n \in \mathcal{I}'$ and $m \in \{1, \ldots, M\} \setminus \mathcal{I}'$ with $w_n > w_m$. Let $\ell_1, \ell_2 \in \{1, \ldots, L\}$ with $n \in \mathcal{I}_{\ell_1}$ and $m \in \mathcal{I}_{\ell_2}$. We claim that $\ell_1 < \ell_2$. First assume $\ell_2 \leq \ell'$. Then, as $m \notin \mathcal{I}'$, we have $\mathcal{J}_{\ell_2} = \emptyset$ and so $\mu_{\ell_2} = w_m < w_n \leq \mu_{\ell_1}$. Hence $\ell_1 < \ell_2$. Now assume $\ell_2 > \ell'$. Then also

$\ell_1 < \ell_2$ because as $n \in \mathcal{I}'$, $\ell_1 \leq \ell'$. As $n \in \mathcal{I}'$ we have $\ell_1 \in \mathcal{L}$, so $\mathcal{J}_{\ell_1} \neq \emptyset$. But then by condition *3* of definition 1 we should have $w_n \leq w_m$, a contradiction. Therefore

$$\sum_{n \in \mathcal{I}'} w_n = \sum_{m=1}^{M} w_m \tag{43}$$

Now (38) follows from (40), (41), (42) and (43). □

Theorem 3. *Let an* efficient *characterization (of some $S^* \in F_\Phi$) be given by L^*, $(\mathcal{J}_1^*, \ldots, \mathcal{J}_{L^*}^*)$, $(\mathcal{I}_1^*, \ldots, \mathcal{I}_{L^*}^*)$, $\mu_1^* \geq \cdots \geq \mu_{L^*}^*$.*

Then for all $S \in F_\Phi$,

$$\lambda(SDS^T + W) \text{ majorizes } (\underbrace{\mu_1^*, \ldots, \mu_1^*}_{|\mathcal{I}_1^*|}, \ldots\ldots, \underbrace{\mu_{L^*}^*, \ldots, \mu_{L^*}^*}_{|\mathcal{I}_{L^*}^*|})$$

Proof. Let

$$\lambda^* = (\underbrace{\mu_1^*, \ldots, \mu_1^*}_{|\mathcal{I}_1^*|}, \ldots\ldots, \underbrace{\mu_{L^*}^*, \ldots, \mu_{L^*}^*}_{|\mathcal{I}_{L^*}^*|})$$

Consider any $S \in F_\Phi$ along with its characterization of theorem 2. For $n \in \mathcal{I}_\ell$, take $\lambda_n = \mu_\ell$. I.e. λ_n is the eigenvalue of $SDS^T + W$ associated with q_n [5]. Then

$$\lambda(SDS^T + W) = (\lambda_{[1]}, \ldots, \lambda_{[N]})$$

We want to prove that λ majorizes λ^*. Assume not. Then there exists $V \in \{1, \ldots, N-1\}$ such that

$$\sum_{m=1}^{V} \lambda_{[m]} < \sum_{m=1}^{V} \lambda_m^*$$

Take the smallest such V. Hence $\lambda_{[V]} < \lambda_V^*$. Take $\ell' \in \{1, \ldots, L\}$ such that $\sum_{\ell=1}^{\ell'-1} |\mathcal{I}_\ell^*| < V$ and $\sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell^*| \geq V$. Define $\hat{V} = \sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell^*|$. For all $m \in \left\{V+1, \ldots, \hat{V}\right\}$ we have $\lambda_m^* = \mu_{\ell'}^* = \lambda_V^* > \lambda_{[V]} \geq \lambda_{[m]}$. Therefore

$$\sum_{m=1}^{\hat{V}} \lambda_{[m]} < \sum_{m=1}^{\hat{V}} \lambda_m^* \tag{44}$$

Clearly $\hat{V} < N$ because

$$\sum_{m=1}^{N} \lambda_{[m]} = \sum_{m=1}^{N} \lambda_m^* = \sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n$$

[5]Note that the components of λ^* are ordered non-increasing, but the components of λ are ordered according to the noise eigenvalues.

Therefore $\ell' < L$. Hence we can apply lemma 12 to obtain

$$\begin{aligned}\sum_{m=1}^{\hat{V}} \lambda_m^* &= \sum_{\ell=1}^{\ell'} \mu_\ell^* |\mathcal{I}_\ell^*| \\ &= \sum_{m=1}^{M} (p_m + w_m) + \sum_{r=0}^{R-1} w_{N-r} \end{aligned} \tag{45}$$

for some $M \in \{1, \ldots, \min(N-1, K)\}$ and $R \in \{0, \ldots, N-M-1\}$ with $M + R = \hat{V}$. Hence by (44),

$$\sum_{m=1}^{\hat{V}} \lambda_{[m]} < \sum_{m=1}^{M} (p_m + w_m) + \sum_{r=0}^{R-1} w_{N-r} \tag{46}$$

Now for $n \in \{1, \ldots, N\}$ let γ_n be the eigenvalue of SDS^T associated with q_n, i.e. $\gamma_n = \lambda_n - w_n$. As $D^{\frac{1}{2}} S^T S D^{\frac{1}{2}}$ has diagonal elements $(p_1, \ldots, p_K)$ and the same nonzero eigenvalues as SDS^T, from theorem 1

$$(\gamma_{[1]}, \ldots, \gamma_{[\min(K,N)]}, \underbrace{0, \ldots, 0}_{K-\min(K,N)}) \text{ majorizes } (p_1, \ldots, p_K) \tag{47}$$

Let $A_M \subset \{1, \ldots, N\}$ such that $|A_M| = M$ and $\sum_{n \in A_M} \gamma_n = \sum_{m=1}^{M} \gamma_{[m]}$. Define $B_M = \{N - R + 1, \ldots, N\} \setminus A_M$. Clearly $|B_M| \leq R$. Take any subset $C_M \subset \{1, \ldots, N\} \setminus (A_M \cup B_M)$ with $|C_M| = R - |B_M|$. This is always possible because

$$|\{1, \ldots, N\} \setminus (A_M \cup B_M)| = N - M - |B_M| = N - \hat{V} + R - |B_M| > R - |B_M|$$

as $\hat{V} < N$.

Now from the definition of A_M and using (47) we get

$$\sum_{m \in A_M} \lambda_m = \sum_{m \in A_M} (\gamma_m + w_m) = \sum_{m=1}^{M} \gamma_{[m]} + \sum_{m \in A_M} w_m \geq \sum_{m=1}^{M} p_m + \sum_{m \in A_M} w_m \tag{48}$$

As SDS^T is non-negative definite, γ_n is non-negative and therefore $\lambda_n = \gamma_n + w_n \geq w_n$ for all $n \in \{1, \ldots, N\}$. Hence

$$\sum_{m \in B_M \cup C_M} \lambda_m \geq \sum_{m \in B_M \cup C_M} w_m$$

and from (48),

$$\sum_{m \in A_M \cup B_M \cup C_M} \lambda_m \geq \sum_{m=1}^{M} p_m + \sum_{m \in A_M \cup B_M \cup C_M} w_m \tag{49}$$

Note that $\{N - R + 1, \ldots, N\} \subset (A_M \cup B_M \cup C_M)$. Define

$$E_M = (A_M \cup B_M \cup C_M) \setminus \{N - R + 1, \ldots, N\}$$

Then $|E_M| = |A_M \cup B_M \cup C_M| - R = M$. Therefore

$$\sum_{m\in A_M\cup B_M\cup C_M} w_m = \sum_{r=0}^{R-1} w_{N-r} + \sum_{m\in E_M} w_m \geq \sum_{r=0}^{R-1} w_{N-r} + \sum_{m=1}^{M} w_m$$

Introducing this inequality in (49) we obtain

$$\sum_{m\in A_M\cup B_M\cup C_M} \lambda_m \geq \sum_{m=1}^{M}(p_m + w_m) + \sum_{r=0}^{R-1} w_{N-r} = \sum_{m=1}^{\hat{V}} \lambda_m^*$$

But $|A_M \cup B_M \cup C_M| = M + R = \hat{V}$, hence

$$\sum_{m\in A_M\cup B_M\cup C_M} \lambda_m \leq \sum_{m=1}^{\hat{V}} \lambda_{[m]}$$

So we get

$$\sum_{m=1}^{\hat{V}} \lambda_{[m]} \geq \sum_{m=1}^{\hat{V}} \lambda_m^*$$

This contradicts (44). Therefore λ majorizes λ^* as we wanted to prove. □

Theorem 4. *Given any $S \in \mathcal{S}$ there exists $S' \in F_\Phi$ such that $\lambda(SDS^T + W)$ majorizes $\lambda(S'DS'^T + W)$.*

Proof. Consider any $S \in \mathcal{S}$. We will recursively generate a sequence of configurations. Take $S^{(0)} = S$. Given $S^{(t)}$ we will compute $S^{(t+1)}$ as follows. For each $k \in \{1, \dots, K\}$, let $v_k \in \mathbb{S}^{N-1}$ be a unit-norm eigenvector of $S_k^{(t)} D_k (S_k^{(t)})^T + W$ associated with the minimum eigenvalue. Let

$$\hat{S}^{(t+1,k)} = \begin{bmatrix} s_1^{(t)} & \dots & s_{k-1}^{(t)} & v_k & s_{k+1}^{(t)} & \dots & s_K^{(t)} \end{bmatrix}$$

Take any $k^* \in \{1, \dots, K\}$ such that

$$\mathrm{TSC}(\hat{S}^{(t+1,k^*)}) = \min\{\mathrm{TSC}(\hat{S}^{(t+1,k)}) : k \in \{1, \dots, K\}\}$$

and define $S^{(t+1)} = \hat{S}^{(t+1,k^*)}$.

Applying lemma 1 with $H = S_{k^*}^{(t)} D_{k^*} (S_{k^*}^{(t)})^T + W$, $v = v_{k^*}$ and $s = s_{k^*}$ we obtain

$$\lambda(S^{(t)} D (S^{(t)})^T + W) \text{ majorizes } \lambda(S^{(t+1)} D (S^{(t+1)})^T + W) \tag{50}$$

Also for any $k \in \{1, \dots, K\}$, we can apply lemma 1 with $H = S_k^{(t)} D_k (S_k^{(t)})^T + W$, $v = v_k$ and $s = c_k(S)$ (i.e. s is the normalized MMSE linear filter for user k) to obtain

$$\lambda(\Phi_k(S^{(t)}) D (\Phi_k(S^{(t)}))^T + W) \text{ majorizes } \lambda(\hat{S}^{(t+1,k)} D (\hat{S}^{(t+1,k)})^T + W)$$

and therefore as TSC$(\cdot)$ is Schur-convex, TSC$(\hat{S}^{(t+1,k)}) \leq$ TSC$(\Phi_k(S^{(t)}))$. Hence for all $k \in \{1, \ldots, K\}$,

$$\mathrm{TSC}(S^{(t+1)}) \leq \mathrm{TSC}(\Phi_k(S^{(t)})) \tag{51}$$

As $\mathcal{S}$ is a compact set, there exist $S' \in \mathcal{S}$ and a subsequence $\{S^{(t_m)}\}_{m=1}^{\infty}$ such that $\lim_{m\to\infty} S^{(t_m)} = S'$. By continuity and transitivity of the majorization relation, equation (50) implies

$$\lambda(SDS^T + W) \text{ majorizes } \lambda(S'DS'^T + W)$$

Take any $k \in \{1, \ldots, K\}$. Then from (51) for all $k \in \{1, \ldots, K\}$,

$$\mathrm{TSC}(S^{(t_{m+1})}) \leq \mathrm{TSC}(S^{(t_m+1)}) \leq \mathrm{TSC}(\Phi_k(S^{(t_m)})) \leq \mathrm{TSC}(S^{(t_m)}) \tag{52}$$

where the first inequality follows from (50) because TSC$(\cdot)$ is Schur-convex and the last one from lemma 3. Letting $m \to \infty$ in (52), by continuity of TSC$(\cdot)$ and $\Phi_k(\cdot)$ we obtain

$$\mathrm{TSC}(S') = \mathrm{TSC}(\Phi_k(S'))$$

and hence by lemma 3, $S' = \Phi_k(S')$. As this holds for all $k \in \{1, \ldots, K\}$, we have $S' = \Phi(S')$, i.e. $S' \in F_\Phi$ as we wanted to prove. □

Theorem 5. *Let*

$$[L^*, (\mathcal{J}_1^*, \ldots, \mathcal{J}_{L^*}^*), (\mathcal{I}_1^*, \ldots, \mathcal{I}_{L^*}^*), (\mu_1^*, \ldots, \mu_{L^*}^*)] = \mathcal{A}(K, N, (p_1, \ldots, p_K), (w_1, \ldots, w_N))$$

Then for all $S \in \mathcal{S}$,

$$\lambda(SDS^T + W) \text{ majorizes } (\underbrace{\mu_1^*, \ldots, \mu_1^*}_{|\mathcal{I}_1^*|}, \ldots\ldots, \underbrace{\mu_{L^*}^*, \ldots, \mu_{L^*}^*}_{|\mathcal{I}_{L^*}^*|})$$

Proof. Take any $S \in \mathcal{S}$. By theorem 4 there exists $S' \in F_\Phi$ such that

$$\lambda(SDS^T + W) \text{ majorizes } \lambda(S'DS'^T + W)$$

By lemma 11 and theorem 3 we obtain

$$\lambda(S'DS'^T + W) \text{ majorizes } (\underbrace{\mu_1^*, \ldots, \mu_1^*}_{|\mathcal{I}_1^*|}, \ldots\ldots, \underbrace{\mu_{L^*}^*, \ldots, \mu_{L^*}^*}_{|\mathcal{I}_{L^*}^*|})$$

Hence by transitivity of the majorization relation,

$$\lambda(SDS^T + W) \text{ majorizes } (\underbrace{\mu_1^*, \ldots, \mu_1^*}_{|\mathcal{I}_1^*|}, \ldots\ldots, \underbrace{\mu_{L^*}^*, \ldots, \mu_{L^*}^*}_{|\mathcal{I}_{L^*}^*|})$$

□

Corollary 1. *Let*

$$[L^*, (\mathcal{J}_1^*, \ldots, \mathcal{J}_{L^*}^*), (\mathcal{I}_1^*, \ldots, \mathcal{I}_{L^*}^*), (\mu_1^*, \ldots, \mu_{L^*}^*)] = \mathcal{A}(K, N, (p_1, \ldots, p_K), (w_1, \ldots, w_N))$$

Then

$$(\underbrace{\mu_1^*, \ldots, \mu_1^*}_{|\mathcal{I}_1^*|}, \ldots\ldots, \underbrace{\mu_{L^*}^*, \ldots, \mu_{L^*}^*}_{|\mathcal{I}_{L^*}^*|})$$

is a Schur-minimal element of the set $\{\lambda(SDS^T + W) : S \in \mathcal{S}\}$ *and*

$$\begin{aligned}
\min_{S\in\mathcal{S}} \mathrm{TSC}(S) &= \sum_{\ell=1}^{L^*} |\mathcal{I}_\ell^*| \, (\mu_\ell^*)^2 \\
\max_{S\in\mathcal{S}} C_{sum}(S) &= \frac{1}{2}\sum_{\ell=1}^{L^*} |\mathcal{I}_\ell^*| \log(\mu_\ell^*) - \frac{1}{2}\log\det(W)
\end{aligned}$$

Proof. Follows from theorem 5 and lemma 10 because TSC is Schur-convex and C_{sum} is Schur-concave. □

9 Local minima of TSC

In this section we will prove an important property of the TSC function: that it has no local minima other than the global minima. To state this formally, let us first define a metric on $\mathcal{S}$. Given $S, S' \in \mathcal{S}$, we define the distance between S and S' as the maximum over the users of the angle between the two signatures assigned to the user:

$$d(S, S') = \max_{k=1\ldots K} \arccos(s_k^T s_k') \tag{53}$$

Note that the triangle inequality holds: given $S, S', S'' \in \mathcal{S}$,

$$\begin{aligned}
d(S, S'') = & \max_{k=1\ldots K} \arccos(s_k^T s_k'') \le \max_{k=1\ldots K} \left[\arccos(s_k^T s_k') + \arccos(s_k'^T s_k'')\right] \\
& \le \max_{k=1\ldots K} \arccos(s_k^T s_k') + \max_{k=1\ldots K} \arccos(s_k'^T s_k'') = d(S, S') + d(S', S'')
\end{aligned}$$

and hence $d(\cdot,\cdot)$ is a metric. Given $S \in \mathcal{S}$ and $\theta \in (0, \pi]$ let $B[S, \theta]$ be the closed ball of radius θ centered at S:

$$B[S, \theta] = \{S' \in \mathcal{S} : d(S', S) \le \theta\} \tag{54}$$

In order to state the main result of this section, we will proceed with some lemmas.

Lemma 13. *If* TSC *has a local minimum at* $S \in \mathcal{S}$, *then for all* $k \in \{1, \ldots, K\}$, s_k *is an eigenvector of* $S_k D_k S_k^T + W$ *associated with the minimum eigenvalue.*

Proof. Assume there exists $k \in \{1, \ldots, K\}$ such that s_k is not an eigenvector associated with the minimum eigenvalue of $(S_k D_k S_k^T + W)$. Let λ be the minimum eigenvalue of $(S_k D_k S_k^T + W)$ and let v be a unit-norm eigenvector associated with λ.

Consider any $\epsilon \in (0, \pi)$ and take S' with $s'_m = s_m$ for $m \neq k$ and $s'_k = \alpha s_k + \beta v$, where $\alpha = \cos\epsilon$ and $\beta = -\alpha s_k^T v + \text{sign}(s_k^T v)\sqrt{\alpha^2 (s_k^T v)^2 + 1 - \alpha^2}$. This is valid because $s_k'^T s_k' = \alpha^2 + \beta^2 + 2\alpha\beta s_k^T v = 1$. We see that $\beta s_k^T v \geq 0$ and thus $s_k^T s_k' = \alpha + \beta s_k^T v \geq \alpha$ and $d(S, S') = \arccos\left(s_k^T s_k'\right) = \epsilon$. Direct computation shows:

$$\text{TSC}(S) - \text{TSC}(S') = 2p_k(1-\alpha^2)\left[s_k^T\left(S_k D_k S_k^T + W\right)s_k - \lambda\right] \tag{55}$$

As $\left(S_k D_k S_k^T + W\right)$ is a symmetric matrix with minimum eigenvalue λ and s_k is not an eigenvector associated with λ, $s_k^T\left(S_k D_k S_k^T + W\right)s_k > \lambda$. Also $1 - \alpha^2 = \sin^2\epsilon > 0$ and therefore $\text{TSC}(S) > \text{TSC}(S')$. Hence there are configurations arbitrarily close to S which attain a smaller TSC. This implies that TSC does not have a local minimum at S. As this followed from assuming that s_k is *not* an eigenvector associated with the minimum eigenvalue of $(S_k D_k S_k^T + W)$ for some k, the lemma is proved. □

Corollary 2. *If* TSC *has a local minimum at* $S \in \mathcal{S}$*, then* $S \in F_\Phi$*.*

Proof. Apply lemmas 13 and 5. □

By corollary 2 all local minima of TSC are fixed configurations of the MMSE update. Hence in what follows, we can associate with each local minimum of TSC the characterization of lemma 6. The next three lemmas, which use the same ideas as in [5], present necessary conditions on this characterization for a configuration to be a local minimum of TSC.

Lemma 14. *Let* TSC *have a local minimum at* $S \in \mathcal{S}$ *and consider the characterization of lemma 6. Then given* $\ell_1, \ell_2 \in \{1, \ldots, L\}$ *with* $\mu_{\ell_1} > \mu_{\ell_2}$*,* $k_1 \in \mathcal{J}_{\ell_1}$ *and* $k_2 \in \mathcal{J}_{\ell_2}$ *we must have* $p_{k_1} \geq p_{k_2}$*.*

Proof. Assume not, i.e. $p_{k_1} < p_{k_2}$. Consider any $\epsilon > 0$ and let $\alpha = \sin\epsilon$ and $\beta = -\frac{p_{k_1}}{p_{k_2}}\alpha$. Take S' with $s'_k = s_k$ for $k \notin \{k_1, k_2\}$, $s'_{k_1} = \sqrt{1-\alpha^2}s_{k_1} + \alpha s_{k_2}$ and $s'_{k_2} = \sqrt{1-\beta^2}s_{k_2} + \beta s_{k_1}$. This can be done because s_{k_1} is orthogonal to s_{k_2} and therefore $\|s'_{k_1}\| = \|s'_{k_2}\| = 1$. Then

$$S'DS'^T = SDS^T + \Delta$$

where

$$\Delta = \left(\beta^2 p_{k_2} - \alpha^2 p_{k_1}\right)\left(s_{k_1}s_{k_1}^T - s_{k_2}s_{k_2}^T\right) + \left(p_{k_1}\alpha\sqrt{1-\alpha^2} + p_{k_2}\beta\sqrt{1-\beta^2}\right)\left(s_{k_1}s_{k_2}^T + s_{k_2}s_{k_1}^T\right)$$

Hence $\text{TSC}(S') = \text{TSC}(S) + 2\text{tr}((SDS^T + W)\Delta) + \text{tr}(\Delta^2)$. Using (19) and (23) we obtain:

$$\text{TSC}(S) - \text{TSC}(S') =$$
$$2\left[\left(\mu_{\ell_1} - \mu_{\ell_2}\right)\left(\alpha^2 p_{k_1} - \beta^2 p_{k_2}\right) - \left(\alpha^2 p_{k_1} - \beta^2 p_{k_2}\right)^2 - \left(p_{k_1}\alpha\sqrt{1-\alpha^2} + p_{k_2}\beta\sqrt{1-\beta^2}\right)^2\right]$$

Now replace for $\beta = -\frac{p_{k_1}}{p_{k_2}}\alpha$ and $\alpha = \sin\epsilon$, and observe that

$$\mathrm{TSC}(S) - \mathrm{TSC}(S') = 2(\mu_{\ell_1} - \mu_{\ell_2})\epsilon^2 p_{k_1}\left(1 - \frac{p_{k_1}}{p_{k_2}}\right) + o(\epsilon^3)$$

As $\mu_{\ell_1} > \mu_{\ell_2}$ by hypothesis and we have assumed $p_{k_1} < p_{k_2}$, for small ϵ we have $\mathrm{TSC}(S') < \mathrm{TSC}(S)$. Therefore, as $d(S, S') \leq \epsilon$, there are configurations arbitrarily close to S with lower TSC. This contradicts the fact that TSC has a local minimum at S and therefore we conclude $p_{k_1} \geq p_{k_2}$. □

Lemma 15. *Let* TSC *have a local minimum at* $S \in \mathcal{S}$ *and consider the characterization of lemma 6. Then given* $\ell_1, \ell_2 \in \{1, \ldots, L\}$ *with* $\mathcal{J}_{\ell_1} \neq \emptyset$, $\mu_{\ell_1} > \mu_{\ell_2}$, $n_1 \in \mathcal{I}_{\ell_1}$ *and* $n_2 \in \mathcal{I}_{\ell_2}$ *we must have* $w_{n_1} \leq w_{n_2}$.

Proof. Assume not, i.e. $w_{n_1} > w_{n_2}$. Define S' as follows. For $k \notin \mathcal{J}_{\ell_1} \cup \mathcal{J}_{\ell_2}$ let $s'_k = s_k$. Let α_1, α_2 be real numbers with $|\alpha_1| \leq 1$ and $|\alpha_2| \leq 1$. For $k \in \mathcal{J}_{\ell_1}$, we can write $s_k = a_k q_{n_1} + v_k$, where $a_k = q_{n_1}^T s_k$ and $v_k = (I - q_{n_1} q_{n_1}^T) s_k$; and we define $s'_k = \sqrt{1 - \alpha_1^2} a_k q_{n_1} + \alpha_1 a_k q_{n_2} + v_k$. Note that this is valid because $\|s'_k\| = 1$. Similarly, for $k \in \mathcal{J}_{\ell_2}$, we write $s_k = a_k q_{n_2} + v_k$ where $a_k = q_{n_2}^T s_k$ and $v_k = (I - q_{n_2} q_{n_2}^T) s_k$; and define $s'_k = \sqrt{1 - \alpha_2^2} a_k q_{n_2} + \alpha_2 a_k q_{n_1} + v_k$.

For $k \in \mathcal{J}_{\ell_1}$ we obtain:

$$\begin{aligned} s'_k s'^T_k - s_k s_k^T &= \alpha_1^2 a_k^2 \left(q_{n_2} q_{n_2}^T - q_{n_1} q_{n_1}^T\right) + \alpha_1 \sqrt{1 - \alpha_1^2} a_k^2 \left(q_{n_1} q_{n_2}^T + q_{n_2} q_{n_1}^T\right) \\ &\quad + \left(\sqrt{1 - \alpha_1^2} - 1\right) a_k \left(q_{n_1} v_k^T + v_k q_{n_1}^T\right) + \alpha_1 a_k \left(q_{n_2} v_k^T + v_k q_{n_2}^T\right) \end{aligned}$$

and similarly for $k \in \mathcal{J}_{\ell_2}$:

$$\begin{aligned} s'_k s'^T_k - s_k s_k^T &= \alpha_2^2 a_k^2 \left(q_{n_1} q_{n_1}^T - q_{n_2} q_{n_2}^T\right) + \alpha_2 \sqrt{1 - \alpha_2^2} a_k^2 \left(q_{n_1} q_{n_2}^T + q_{n_2} q_{n_1}^T\right) \\ &\quad + \left(\sqrt{1 - \alpha_2^2} - 1\right) a_k \left(q_{n_2} v_k^T + v_k q_{n_2}^T\right) + \alpha_2 a_k \left(q_{n_1} v_k^T + v_k q_{n_1}^T\right) \end{aligned}$$

We claim that

$$\sum_{k \in \mathcal{J}_{\ell_1}} p_k a_k v_k = 0$$

To see this use (28) to write

$$\begin{aligned} \sum_{k \in \mathcal{J}_{\ell_1}} p_k a_k v_k &= \sum_{k \in \mathcal{J}_{\ell_1}} p_k (I - q_{n_1} q_{n_1}^T) s_k s_k^T q_{n_1} \\ &= (I - q_{n_1} q_{n_1}^T) S_{\mathcal{J}_{\ell_1}} D_{\mathcal{J}_{\ell_1}} S_{\mathcal{J}_{\ell_1}}^T q_{n_1} = (I - q_{n_1} q_{n_1}^T)(\mu_{\ell_1} - w_{n_1}) q_{n_1} = 0 \end{aligned}$$

Similarly, $\sum_{k \in \mathcal{J}_{\ell_2}} p_k a_k v_k = 0$. Using these identities it is straightforward to obtain:

$$S' D S'^T = S D S^T + \Delta_1 + \Delta_2$$

where

$$\begin{aligned}
\Delta_1 &= \alpha_1^2 P_1 \left(q_{n_2}q_{n_2}^T - q_{n_1}q_{n_1}^T\right) + \alpha_1\sqrt{1-\alpha_1^2}P_1\left(q_{n_1}q_{n_2}^T + q_{n_2}q_{n_1}^T\right) \\
\Delta_2 &= \alpha_2^2 P_2 \left(q_{n_1}q_{n_1}^T - q_{n_2}q_{n_2}^T\right) + \alpha_2\sqrt{1-\alpha_2^2}P_2\left(q_{n_1}q_{n_2}^T + q_{n_2}q_{n_1}^T\right) \\
P_1 &= \sum_{k\in\mathcal{J}_{\ell_1}} p_k a_k^2 \\
P_2 &= \sum_{k\in\mathcal{J}_{\ell_2}} p_k a_k^2
\end{aligned}$$

Now

$$\begin{aligned}
&\mathrm{TSC}(S) - \mathrm{TSC}(S') = \\
&-2\mathrm{tr}\left[\left(SDS^T + W\right)(\Delta_1+\Delta_2)\right] - \mathrm{tr}\left(\Delta_1^2\right) - \mathrm{tr}\left(\Delta_2^2\right) - 2\mathrm{tr}\left(\Delta_1\Delta_2\right)
\end{aligned}$$

and after some manipulation we get:

$$\begin{aligned}
\mathrm{TSC}(S) - \mathrm{TSC}(S') &= 2(\mu_{\ell_1}-\mu_{\ell_2})\left(\alpha_1^2P_1 - \alpha_2^2P_2\right) - 2\left(\alpha_1^2P_1^2 + \alpha_2^2P_2^2\right) \\
&\quad + 4\alpha_1^2\alpha_2^2P_1P_2 - 4\alpha_1\alpha_2\sqrt{1-\alpha_1^2}\sqrt{1-\alpha_2^2}P_1P_2
\end{aligned}$$

Hence

$$\mathrm{TSC}(S) - \mathrm{TSC}(S') = 2(\mu_{\ell_1}-\mu_{\ell_2})\left(\alpha_1^2P_1 - \alpha_2^2P_2\right) - 2\left(\alpha_1P_1 + \alpha_2P_2\right)^2 + o(\|\alpha\|^3)$$

where $\|\alpha\| = \sqrt{\alpha_1^2+\alpha_2^2}$.

From (28) follows $\mu_{\ell_1} = P_1 + w_{n_1}$, $\mu_{\ell_2} = P_2 + w_{n_2}$. As we are assuming $w_{n_1} > w_{n_2}$, we have $\mu_{\ell_1} - \mu_{\ell_2} + P_2 = P_1 + w_{n_1} - w_{n_2} > 0$ and thus we can take $\alpha_2 = -\frac{\alpha_1 P_1}{\mu_{\ell_1}-\mu_{\ell_2}+P_2}$. Operating we get:

$$\mathrm{TSC}(S) - \mathrm{TSC}(S') = \frac{2(\mu_{\ell_1}-\mu_{\ell_2})\alpha_1^2P_1\left(w_{n_1}-w_{n_2}\right)}{P_1 + w_{n_1} - w_{n_2}} + o(\alpha_1^3)$$

By hypothesis $\mathcal{J}_{\ell_1} \neq \emptyset$ which implies (lemma 6) that $\mu_{\ell_1} > w_{n_1}$, i.e. $P_1 > 0$. Also by hypothesis $\mu_{\ell_1} > \mu_{\ell_2}$. Thus for α_1 small enough we get $\mathrm{TSC}(S) - \mathrm{TSC}(S') > 0$. Hence, as $d(S, S') \leq |\arcsin(\alpha_1)|$, there are configurations arbitrarily close to S with lower TSC. This contradicts the hypothesis that TSC has a local minimum at S, so we conclude that $w_{n_1} \leq w_{n_2}$. □

Lemma 16. *Let* TSC *have a local minimum at $S \in \mathcal{S}$ and consider the characterization of lemma 6. Let $\ell \in \{1, \ldots, L\}$ with $\mu_\ell > \min_{\ell'\in\{1,\ldots,L\}} \mu_{\ell'}$. Then $|\mathcal{J}_\ell| \leq |\mathcal{I}_\ell|$.*

Proof. Assume not. Then there exist $\ell_1, \ell_2 \in \{1, \ldots, L\}$ with $\mu_{\ell_1} > \mu_{\ell_2}$ and $|\mathcal{J}_{\ell_1}| > |\mathcal{I}_{\ell_1}|$. Take any $n \in \mathcal{I}_{\ell_2}$. As $\mathrm{rank}(S_{\mathcal{J}_{\ell_1}}) = |\mathcal{I}_{\ell_1}| < |\mathcal{J}_{\ell_1}|$, we can find a column vector $v \in \mathbb{R}^{|\mathcal{J}_{\ell_1}|}$ such that $\|v\| = 1$ and $S_{\mathcal{J}_{\ell_1}} D_{\mathcal{J}_{\ell_1}} v = 0$. Consider any $\epsilon > 0$ and define S' with $s'_k = s_k$ for $k \notin \mathcal{J}_{\ell_1}$ and $s'_k = \cos(\alpha_k)s_k + \sin(\alpha_k)q_n$ for $k \in \mathcal{J}_{\ell_1}$, where $\alpha_k = \epsilon v_k$. With this choice, after some manipulation we get:

$$\mathrm{TSC}(S) - \mathrm{TSC}(S') = 2\epsilon^2(\mu_{\ell_1}-\mu_{\ell_2})\|D_{\mathcal{J}_{\ell_1}} v\|^2 + o(\epsilon^3) \qquad (56)$$

So it suffices to make ϵ small enough to get $\mathrm{TSC}(S') < \mathrm{TSC}(S)$ and $d(S,S') = \epsilon \max_{k\in\mathcal{J}_{\ell_1}} |v_k| \leq \epsilon$. This contradicts the fact that TSC has a local minimum at S. □

Theorem 6. *Let* TSC *have a local minimum at* $S \in \mathcal{S}$. *Then* S *has an* efficient *characterization.*

Proof. Consider the characterization of lemma 6.

Let $\ell \in \{1,\ldots,L-1\}$. If $\mu_\ell > \mu_L$, by lemma 14 $|\mathcal{J}_\ell| \leq |\mathcal{I}_\ell|$. If $\mu_\ell = \mu_L$, by condition *3* of lemma 6 we have $|\mathcal{J}_\ell| = 0 < |\mathcal{I}_\ell|$. Therefore condition *1* of definition 1 is satisfied.

Now let $\ell_1 < \ell_2 \in \{1,\ldots,L\}$ with $\mathcal{J}_{\ell_1} \neq \emptyset$. If it were $\mu_{\ell_1} = \mu_{\ell_2}$, condition *3* of lemma 6 would imply $\mathcal{J}_{\ell_1} = \emptyset$. Hence $\mu_{\ell_1} > \mu_{\ell_2}$. Then by lemmas 14 and 15, conditions *2* and *3* of definition 1 are satisfied. □

Theorem 7. *Local minima of* TSC *are global. I.e. if* TSC *has a local minimum at* $S \in \mathcal{S}$, *then* $S \in \Omega$.

Proof. Assume TSC has a local minimum at $S \in \mathcal{S}$. By theorem 6, S has an *efficient* characterization. Hence we can apply theorems 3 and 4 to obtain that for all $S' \in \mathcal{S}$,

$$\lambda(S'DS'^T + W) \text{ majorizes } \lambda(SDS^T + W)$$

Thus as TSC is Schur-convex $\mathrm{TSC}(S) \leq \mathrm{TSC}(S')$ for all $S' \in \mathcal{S}$, i.e. $S \in \Omega$. □

Theorem 7 can be rephrased saying that if $S \in \mathcal{S}$ is not a global optimal configuration, then TSC cannot have a local minimum at S. I.e. given any $S \in \mathcal{S} \setminus \Omega$, for all $\epsilon \in (0,\pi]$ there exists $S' \in B[S,\epsilon]$ with $\mathrm{TSC}(S') < \mathrm{TSC}(S)$.

Hence theorem 7 implies that all the non-optimal fixed configurations are unstable equilibria of the MMSE update. If a fixed configuration S does not achieve the minimum of TSC, then there exist arbitrarily small perturbations such that if the MMSE iteration is started from these perturbed configurations, the TSC converges as $t \to \infty$ to a value strictly smaller than $\mathrm{TSC}(S)$. We state this formally in the following lemma.

Lemma 17. *Given* $S \in F_\Phi \backslash \Omega$, *for all* $\epsilon > 0$ *there exists* $S' \in B[S,\epsilon]$ *such that for the MMSE iteration with* $S^{(0)} = S'$ *we have* $\lim_{t\to\infty} \mathrm{TSC}(S^{(t)}) < \mathrm{TSC}(S)$.

Proof. As $S \in F_\Phi \backslash \Omega$, TSC does not have a global minimum at S. Hence by theorem 7 given any $\epsilon > 0$ there exists $S' \in B[S,\epsilon]$ such that $\mathrm{TSC}(S') < \mathrm{TSC}(S)$. If we start the MMSE iteration with $S^{(0)} = S'$, as $\mathrm{TSC}(S^{(t)})$ is non-increasing, we get $\lim_{t\to\infty} \mathrm{TSC}(S^{(t)}) \leq \mathrm{TSC}(S') < \mathrm{TSC}(S)$. □

On the other hand, if a configuration S achieves the minimum of TSC, then if we start the MMSE iteration from any configuration close enough to S, the TSC converges to $\mathrm{TSC}(S)$ as $t \to \infty$.

Lemma 18. *Given* $S \in \Omega$ *there exists* $\epsilon > 0$ *such that for all* $S' \in B[S,\epsilon]$ *the MMSE iteration with* $S^{(0)} = S'$ *satisfies* $\lim_{t\to\infty} \mathrm{TSC}(S^{(t)}) = \tau$.

Proof. Follows from the fact that T_F is finite and TSC is continuous. □

Hence the only stable equilibria of the MMSE update are the optimal configurations.

10 Noisy MMSE iteration

Our last observation on the TSC is key to understand the convergence of the MMSE iteration. We will next slightly modify the MMSE update algorithm adding noise. To this end we first make some definitions. Given two unit-norm orthogonal vectors v_1, v_2 ($v_1, v_2 \in \mathbb{S}^{N-1}$ with $v_1^T v_2 = 0$) and an angle θ, let $h(v_1, v_2, \theta)$ denote the rotation of v_1 of angle θ towards v_2:

$$h(v_1, v_2, \theta) = \cos\theta v_1 + \sin\theta v_2 \tag{57}$$

Analogously, given $S, R \in \mathcal{S}$ with $s_k^T r_k = 0$ for all $k \in \{1, \dots, K\}$, and $\theta \in \mathbb{R}^K$ let

$$h(S, R, \theta) = \left[\begin{array}{ccc} h(s_1, r_1, \theta_1) & \dots & h(s_K, r_K, \theta_K) \end{array}\right]$$

Given a sequence of angles $\{\theta_{max}^{(t)}\}_{t=1}^{\infty} \subset (0, 2\pi)$, we define the MMSE noisy iteration as:

$$S^{(t+1)} = h(\Phi(S^{(t)}), R^{(t+1)}, \theta^{(t+1)}) \tag{58}$$

where $r_k^{(t)}, \theta_k^{(t)}$ ($k \in \{1, \dots, K\}$, $t \in \mathbb{N}$) are independent random variables, $\theta_k^{(t)}$ is uniform $(0, \theta_{max}^{(t)})$ and $r_k^{(t)}$ is a random unit-norm vector uniformly distributed orthogonal to the k-th column of $\Phi(S^{(t-1)})$. In words, the MMSE noisy update consists of applying the MMSE update (13) to all the signatures one at a time, and then adding a random bounded independent noise to each signature.

We now present an intuitive argument to be formalized in the next theorem. We have proved in section 7 that the (noiseless) MMSE iteration approaches the set of fixed configurations as $t \to \infty$. In section 9 we have seen that TSC has no other local minima than the global one. Hence, if we start with any configuration that does not attain the global minimum of TSC and perturb it a little, there will be a nonzero probability of getting a new configuration with a lower TSC. This observation suggests that if we fix a sufficiently small noise upper bound in the noisy iteration, $S^{(t)}$ can be made to converge to an arbitrary small neighborhood of the optimal set with probability one regardless of the initial configuration.

Theorem 8. *Given any $\delta > 0$ there exists $\theta_{max} > 0$ such that for any initial condition $S^{(0)}$ the MMSE noisy iteration defined by (58) with $\theta_{max}^{(t)} = \theta_{max}$ for all t, satisfies*

$$\limsup_{t \to \infty} \mathrm{TSC}(S^{(t)}) \leq_{\text{a.s.}} \tau + \delta \tag{59}$$

Proof. Without loss of generality assume δ is small enough so that if $S \in F_\Phi$ and $\mathrm{TSC}(S) \leq \tau + \delta$ then $\mathrm{TSC}(S) = \tau$. This can be done because, by theorem 2 the set T_F has a finite number of elements (recall equation (33)). Define the sets

$V_1 = \{S \in \mathcal{S} : \text{TSC}(S) \geq \tau + \delta\}$ and $V_2 = \{S \in \mathcal{S} : \text{TSC}(S) \leq \tau + \delta\}$. As TSC$(\cdot)$ is continuous, V_1 and V_2 are compact sets. If $V_1 = \emptyset$, then (59) is trivially satisfied. Hence in what follows we assume $V_1 \neq \emptyset$. Let

$$\theta_{max} = \min\{d(S, S') : S \in V_1 \ , \ S' \in \Phi(V_2)\}$$

Note that θ_{max} is well-defined: $d(\cdot,\cdot)$ is a continuous function, V_1 is a compact set, V_2 is compact and thus $\Phi(V_2)$ is compact because $\Phi(\cdot)$ is continuous.

We claim $\theta_{max} > 0$. To prove this by contradiction, assume $\theta_{max} = 0$. Then there exist $S \in V_1$ and $S' \in \Phi(V_2)$ with $d(S, S') = 0$. So $S = S'$ and hence $\text{TSC}(S) \geq \tau+\delta$ and $S = \Phi(S'')$ for some $S'' \in V_2$. Therefore $\text{TSC}(S'') \leq \tau + \delta$ and we get

$$\tau + \delta \leq \text{TSC}(S) \leq \text{TSC}(S'') \leq \tau + \delta$$

and so $\text{TSC}(S) = \text{TSC}(S'') = \tau + \delta$. By (16) this implies $S = S''$ and thus $S \in F_\Phi$. But then, by our assumption that δ was small enough, we must have $\text{TSC}(S) = \tau$ which contradicts $\text{TSC}(S) = \tau + \delta$.

Because of our choice of θ_{max}, if $S^{(t)} \in V_2$ then $S^{(t+1)} \in V_2$ and thus $S^{(t+m)} \in V_2$ for all $m \geq 0$.

For each $S \in \mathcal{S}$ define

$$\beta(S) = \min\{\text{TSC}(S') : S' \in B[S, \theta_{max}]\}$$

Note that $\beta(S)$ is well-defined because TSC is continuous and $B[S, \theta_{max}]$ is compact. Also $\beta(S)$ is a continuous function of S because TSC$(\cdot)$ is continuous and the set $B[S, \theta_{max}]$ depends continuously on S. Now define

$$\gamma = \min\{\text{TSC}(S) - \beta(S) : S \in V_1\}$$

which is well-defined because $(\text{TSC} - \beta)(\cdot)$ is continuous and V_1 is compact.

We claim $\gamma > 0$. To prove this by contradiction assume $\gamma = 0$. Then for some $S \in V_1$ it is $\beta(S) = \text{TSC}(S)$. But this means that S is a local minimum of TSC$(\cdot)$. Thus, by theorem 7, S must be a global minimum of TSC$(\cdot)$ and therefore $\text{TSC}(S) = \tau$ which contradicts $S \in V_1$.

We will write $\Pr(\cdot)$ for probabilities. For $S \in \mathcal{S}$ define

$$P(S) = \Pr\left(\text{TSC}(h(\Phi(S), R, \theta)) \leq \max\left\{\text{TSC}(S) - \frac{\gamma}{2}, \tau + \delta\right\}\right)$$

where $r_k, \theta_k,\ k \in \{1, \ldots, K\}$ are independent random variables, θ_k is uniform $(0, \theta_{max})$ and r_k is a random unit-norm vector uniformly distributed orthogonal to the k-th column of $\Phi(S)$. Note that $P(S)$ is a continuous function of S because TSC$(\cdot)$, $\Phi(\cdot)$ and $h(\cdot, R, \theta)$ are continuous and the probability distributions involved are continuous. Let

$$p = \min_{S \in V_1} P(S)$$

We claim $p > 0$. To prove this by contradiction assume $p = 0$. Then there exists $S \in V_1$ such that $P(S) = 0$. Consider two cases:

- Assume $\Phi(S) \in V_1$. By definition of γ, there exists $S' \in B[\Phi(S), \theta_{max}]$ such that $\text{TSC}(\Phi(S)) - \text{TSC}(S') > \frac{\gamma}{2}$. By continuity of $\text{TSC}(\cdot)$ and as the probability density of $h(\Phi(S), R, \theta)$ is not identically zero in any open subset of $B[\Phi(S), \theta_{max}]$, this implies

$$\begin{aligned} P(S) &\geq \Pr\left(\text{TSC}(h(\Phi(S), R, \theta)) \leq \text{TSC}(S) - \frac{\gamma}{2}\right) \\ &\geq \Pr\left(\text{TSC}(h(\Phi(S), R, \theta)) \leq \text{TSC}(\Phi(S)) - \frac{\gamma}{2}\right) > 0 \end{aligned}$$

which contradicts $P(S) = 0$.

- Assume $\Phi(S) \notin V_1$. Then $\text{TSC}(\Phi(S)) < \tau + \delta$ and thus by continuity of $\text{TSC}(\cdot)$ and as the probability density of $h(\Phi(S), R, \theta)$ is not identically zero in any open subset of $B[\Phi(S), \theta_{max}]$, we have

$$P(S) \geq \Pr\left(\text{TSC}(h(\Phi(S), R, \theta)) \leq \tau + \delta\right) > 0$$

which contradicts $P(S) = 0$.

Define $M = \left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right)^2$. Note that $\forall S \in \mathcal{S}$, $\text{TSC}(S) \leq M$. Let $Q = \lceil \frac{2(M-\tau-\delta)}{\gamma} \rceil$. Let E_t denote the event that $\text{TSC}(S^{(t)}) \in V_2$ (i.e. that $\text{TSC}(S^{(t)}) \leq \tau + \delta$). Write $z_m = \Pr(E_{Qm})$. Then

$$z_{m+1} = z_m \Pr\left(E_{Q(m+1)} \middle| E_{Qm}\right) + (1 - z_m) \Pr\left(E_{Q(m+1)} \middle| E_{Qm}^c\right)$$

We have argued before that $E_t \subset E_{t+1}$. Therefore $\Pr\left(E_{Q(m+1)} \middle| E_{Qm}\right) = 1$ and

$$z_{m+1} = z_m + (1 - z_m) \Pr\left(E_{Q(m+1)} \middle| E_{Qm}^c\right)$$

Let F_t denote the event that $\text{TSC}(S^{(t)}) \leq \text{TSC}(S^{(t-1)}) - \frac{\gamma}{2}$, and let $G_t = E_t \cup F_t$ (i.e. G_t is the event $\text{TSC}(S^{(t)}) \leq \max\left\{\text{TSC}(S^{(t-1)}) - \frac{\gamma}{2}, \tau + \delta\right\}$).

We claim that $\bigcap_{q=1}^{Q} G_{Qm+q} \subset E_{Q(m+1)}$. To see this, note that

$$\begin{aligned} E_{Q(m+1)}^c \cap \bigcap_{q=1}^{Q} G_{Qm+q} &= \bigcap_{q=1}^{Q} \left[\left(E_{Q(m+1)}^c \cap E_{Qm+q}\right) \cup \left(E_{Q(m+1)}^c \cap F_{Qm+q}\right)\right] \\ &= \bigcap_{q=1}^{Q} \left(E_{Q(m+1)}^c \cap F_{Qm+q}\right) \\ &= E_{Q(m+1)}^c \cap \bigcap_{q=1}^{Q} F_{Qm+q} \\ &= \emptyset \end{aligned}$$

where the last equality follows from the fact that if $\text{TSC}(S^{Qm+q}) \leq \text{TSC}(S^{Qm+q-1}) - \frac{\gamma}{2}$ for all $q \in \{1, \ldots, Q\}$, then $\text{TSC}(S^{Q(m+1)}) \leq \text{TSC}(S^{Qm}) - Q\frac{\gamma}{2} \leq \tau + \delta$ (i.e. $\bigcap_{q=1}^{Q} F_{Qm+q} \subset E_{Q(m+1)}$).

Therefore

$$\Pr\left(E_{Q(m+1)} \middle| E_{Qm}^c\right) \geq \Pr\left(\bigcap_{q=1}^{Q} G_{Qm+q} \middle| E_{Qm}^c\right)$$
$$= \prod_{q=1}^{Q} \Pr\left(G_{Qm+q} \,\middle|\, G_{Qm+1}, \ldots, G_{Qm+q-1}, E_{Qm}^c\right)$$

By the definition of p, for all q we have $\Pr\left(G_{Qm+q} \middle| G_{Qm+1}, \ldots, G_{Qm+q-1}, E_{Qm}^c\right) \geq p$. Hence

$$\Pr\left(E_{Q(m+1)} \middle| E_{Qm}^c\right) \geq p^Q$$

and

$$z_{m+1} \geq z_m + (1 - z_m)p^Q$$

Therefore $1 - z_{m+1} \leq (1 - z_m)(1 - p^Q)$ and by induction $1 - z_m \leq (1 - z_0)\left(1 - p^Q\right)^m \leq (1 - p^Q)^m$. I.e. $z_m \geq 1 - \left(1 - p^Q\right)^m$. Now

$$\Pr\left(\bigcup_{m=0}^{\infty} E_{Qm}\right) = \lim_{m\to\infty} z_m \geq 1 - \lim_{m\to\infty} \left(1 - p^Q\right)^m = 1$$

because $E_{Qm} \subset E_{Q(m+1)}$ and $p > 0$. This implies that with probability 1 for some finite t_0, $S^{(t_0)} \in V_2$. Hence $S^{(t)} \in V_2$ for all $t \geq t_0$, and (59) follows. □

The next theorem shows that if $\theta_{max}^{(t)}$ is chosen suitably with $\theta_{max}^{(t)} \to 0$ as $t \to 0$, then $S^{(t)}$ approaches the optimal set Ω as $t \to \infty$ with probability 1.

Theorem 9. *There exists a sequence $\theta_{max}^{(t)}$ such that for any initial condition $S^{(0)}$ the MMSE noisy iteration defined by (58) satisfies*

$$\lim_{t\to\infty} \mathrm{TSC}(S^{(t)}) =_{\text{a.s.}} \tau \tag{60}$$

Proof. Take a decreasing sequence δ_m with $\lim_{m\to\infty} \delta_m = 0$, and take any $q \in (0,1)$. Fix any m. By the proof of theorem 8 we can find $\hat{\theta}_m$ such that the noisy MMSE iteration (58) with $\theta_{max}^{(t)} = \hat{\theta}_m$ satisfies $\Pr\left(\mathrm{TSC}(S^{(t)}) \leq \tau + \delta_m\right) \to 1$ as $t \to \infty$ uniformly in the initial condition $S^{(0)}$. Thus there exists l_m such that for all $S^{(0)}$ and all $t \geq l_m$, $\Pr\left(\mathrm{TSC}(S^{(t)}) \leq \tau + \delta_m\right) > q$. Let $L_m = \sum_{i=1}^{m} l_i$. It follows that if we choose $\theta_{max}^{(t)} = \hat{\theta}_m$ for all $t = (1 + L_{m-1}), \ldots, L_m$, we obtain that for all $z \geq 0$ it holds $\Pr\left(\mathrm{TSC}(S^{(L_{m+z})}) \leq \tau + \delta_m\right) > 1 - (1 - q)^z$. This implies $\limsup_{t\to\infty} \mathrm{TSC}(S^{(t)}) \leq_{\text{a.s.}} \tau + \delta_m$ for all m. Making $m \to \infty$ we get $\limsup_{t\to\infty} \mathrm{TSC}(S^{(t)}) \leq_{\text{a.s.}} \tau$. As $\mathrm{TSC}(S^{(t)}) \geq \tau$ for all t, we get the desired result. □

11 Conclusions

Given a symbol-synchronous CDMA system with fixed number of users, processing gain, received powers and noise covariance, we considered the problem of assigning signature sequences to the users. Two performance measures where proposed, sum capacity and TSC, and we proved that the optimal configurations for both are the same. The MMSE iteration is an iterative procedure amenable to distributed implementation that decreases the total square correlation at each iteration. However, it does not guarantee convergence to the minimum TSC. We have shown that the TSC has no local minima other than the global, and therefore the fixed configurations of the MMSE update that are not optimal are unstable. Using this fact we have proved that a modified noisy version of the MMSE iteration asymptotically approaches the set of optimal configurations with probability one .

References

[1] P. Anigstein and V. Anantharam. Ensuring convergence of the mmse iteration for interference avoidance to the global optimum. In *38th Allerton Conference on Communications, Control and Computing*, 2000.

[2] F. Gantmacher. *The Theory of Matrices.* Chelsea Publishing Company, 1960.

[3] M. Honig, U. Madhow, and S. Verdú. Blind adaptive multiuser detection. *IEEE Transactions on Information Theory*, 41(4):944–960, July 1995.

[4] A. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications.* Academic Press, 1979.

[5] C. Rose. Cdma codeword optimization: Interference avoidance and convergence via class warfare. *submitted to IEEE Transactions on Information Theory*, 2000. Available at http://www.winlab.rutgers.edu/~crose.

[6] C. Rose, S. Ulukus, and R. Yates. Interference avoidance for wireless systems. In *Vehicular Technology Conference*, 2000.

[7] M. Rupf and J. Massey. Optimum sequence multisets for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 40(4):1261–1266, July 1994.

[8] S. Sastry. *Nonlinear Systems: Analysis, Stability and Control.* Springer, 1999.

[9] S. Ulukus and R. Yates. Iterative signature adaptation for capacity maximization of cdma systems. In *36th Allerton Conference on Communications, Control and Computing*, September 1998.

[10] S. Ulukus and R. Yates. Iterative construction of optimal sequence sets to maximize the capacity of cdma systems. *submitted to IEEE Transactions on Information Theory*, 1999. Available at http://www.winlab.rutgers.edu/~ryates.

[11] S. Verdú. Capacity region of gaussian cdma channels: The symbol-synchronous case. In *24th Allerton Conference on Communications, Control and Computing*, October 1986.

[12] S. Verdú. *Multiuser Detection.* Cambridge University Press, 1998.

[13] P. Viswanath. *Capacity of Vector Multiple Access Channels.* PhD thesis, UC Berkeley, Department of Electrical and Computer Sciences, 2000.

[14] P. Viswanath and V. Anantharam. Optimal sequences and sum capacity of synchronous cdma systems. *IEEE Transactions on Information Theory*, 45(6):1984–1991, September 1999.

[15] P. Viswanath and V. Anantharam. Total capacity of multiaccess vector channels. Memorandum UCB/ERL M99/47, UC Berkeley Electronics Research Laboratory, May 1999.

[16] P. Viswanath and V. Anantharam. Colored additive noise and capacity of cdma. *submitted to IEEE Transactions on Information Theory*, 2000.