Ensuring Convergence of the MMSE Iteration for Interference Avoidance to the Global Optimum

Pablo Anigstein and Venkat Anantharam, *Fellow, IEEE*

***Abstract*—Viswanath and Anantharam [1] characterize the sum capacity of multiaccess vector channels. For a given number of users, received powers, spreading gain, and noise covariance matrix in a code-division multiple-access (CDMA) system, the authors of [1] present a combinatorial algorithm to generate a set of signature sequences that achieves the maximum sum capacity. These sets also minimize a performance measure called generalized total square correlation (TSC_g).**

Ulukus and Yates [2] 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_g at each step. The MMSE iteration has fixed points not only at the optimal configurations which attain the global minimum TSC_g but also at other configurations which are suboptimal. The authors of [2] claim that simulations show that when starting with random sequences, the algorithm converges to optimum sets of sequences, but they give no formal proof.

We show that the TSC_g function has no local minima, in the sense that given any suboptimal set of sequences, there exist arbitrarily close sets with lower TSC_g. 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_g 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.

***Index Terms*—Code-division multiple access (CDMA), interference avoidance, iterative construction of signature sequences, minimum mean-square error (MMSE) receiver, Welch bound equality (WBE) sequences.**

I. Introduction and Previous Work

We consider the uplink of a symbol-synchronous code-division multiple-access (CDMA) system. An important performance measure of such a 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 [3]. Later, Rupf and Massey [4] characterized the maximum sum capacity of a CDMA channel with white noise and equal user received powers. In [5], the case of different user received powers was solved using majorization theory. Viswanath and Anantharam [1] also consider the case of asymmetric received powers with colored noise, and give a recursive algorithm to construct an optimal configuration of signature sequences.

Another performance measure of the CDMA channel is the generalized total square correlation (TSC_g). An iterative procedure called minimum mean-square error (MMSE) iteration, in which at each step one signature sequence is modified in a way such that TSC_g is nonincreasing, was proposed in [2], [6]. Another iterative procedure with the same property is proposed in [7]. 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_g is nonincreasing, but there is no guarantee that the TSC_g 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_g. A modification of the algorithm of [7] is proposed in [8] in order to guarantee convergence to the optimum TSC_g 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_g to the global minimum. A short version of the results herein was presented in [9].

II. Outline

The rest of this paper is organized as follows. In Section III, we present the CDMA channel model used and some notation. In Section IV, we define the majorization partial order on $\mathbb{R}^n$ and state some results that will be used later. In Section V, the two performance measures used, sum capacity and TSC_g, are defined and basic properties of these are listed. Section VI presents

Manuscript received August 29, 2001; revised July 24, 2002. This work was supported by EPRI/DOD Complex Interactive Networks under Contract EPRI-W08333-04 and the National Science Foundation under Contracts ANI 9872764, ECS 9873086, and IIS 9941569. The material in this paper was presented in part at the 38th Allerton Conference on Communications, Control and Computing.

P. Anigstein was with the Electrical Engineering and Computer Science Department, University of California, Berkeley. He is now with Flarion Technologies, Inc., Bedminster, NJ USA (e-mail: panigstein@flarion.com).

V. Anantharam is with the Electrical Engineering and Computer Science Department, University of California, Berkeley, Berkeley, CA 94720 USA (e-mail: ananth@eecs.berkeley.edu).

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

Digital Object Identifier 10.1109/TIT.2003.809595

the MMSE iteration proposed in [2], [6]. The fixed configurations of this iteration are characterized, and we prove that the MMSE iteration asymptotically approaches the set of fixed configurations. In Section VII, we state the recursive algorithm of [1] 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 [1]. In the process, we provide a characterization of the optimal configurations which is useful later. In Section VIII, we observe and prove that TSC_g has no minima other than the global minima. Motivated by this result, in Section IX, 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_g almost surely regardless of the initial configuration.

III. Model

Consider a symbol-synchronous CDMA system with K users. Let T be the duration of the symbol interval and let s_k: $[0, T] \rightarrow \mathbb{R}$ represent the signature waveform assigned to user k, assumed to be of unit norm. The received signal at the base station in one symbol interval can then be expressed as

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

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 to be a zero-mean Gaussian process independent of the user symbols $x_1, \ldots, x_K$.

Let the processing gain be N. The signature waveform of user k can therefore be represented as an N-dimensional vector s_k. Let $S = [s_1 \cdots s_K]$, $D = \mathrm{diag}(p_1, \ldots, p_K)$, and $x = [x_1 \cdots x_K]^T$. We can write

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

where y and z are N-dimensional vectors representing received signal and noise, respectively. 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 (2) 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 stations.

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

$$\mathcal{S} = \left\{ [s_1 \quad \cdots \quad s_K] : s_k \in \mathbb{S}^{N-1} \; \forall k \in \{1, \ldots, K\} \right\} \tag{3}$$

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

We will denote the MMSE linear filter for user k as v_k, 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. The following formulas are well known [10]:

$$v_k = \sqrt{p_k} (SDS^T + W)^{-1} s_k \tag{4}$$

$$= \alpha_k (S_k D_k S_k^T + W)^{-1} s_k \tag{5}$$

where

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

An 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 [10].

IV. 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 $a_{[1]} \geq \cdots \geq a_{[n]}$.

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

$$\sum_{i=1}^{m} a_{[i]} \geq \sum_{i=1}^{m} b_{[i]}, \qquad \forall m \in \{1, \ldots, n-1\}$$

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i.$$

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 $h_1, \ldots, h_n$ and eigenvalues $\lambda_1, \ldots, \lambda_n$. Then λ majorizes h.

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

Proof: See [11, Theorems 9.B.1 and 9.B.2]. □

In the sequel, given a symmetric matrix $H \in \mathbb{R}^{n \times n}$ we will denote by $\lambda(H)$ the vector whose components are the eigenvalues of H in nonincreasing order.

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 [12] or [13]. □

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

Proof: See [11, Theorem 3.C.1]. □

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 $f\colon A \to \mathbb{R}$ is Schur-convex (Schur-concave) and $b \in A$ is a Schur-minimum of A, then f attains a global minimum (maximum) at b.

V. Sum Capacity and TSC_g

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, $C_{\mathrm{sum}}\colon \mathcal{S} \to \mathbb{R}$. It can be shown that [1]

$$C_{\mathrm{sum}}(S) = \frac{1}{2}\log\det\left(SDS^T + W\right) - \frac{1}{2}\log\det(W). \quad (6)$$

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

We define the generalized total square correlation (TSC_g) as a function $\mathrm{TSC}_g\colon \mathcal{S} \to \mathbb{R}$ with [8]

$$\mathrm{TSC}_g(S) = \mathrm{tr}\left[\left(SDS^T + W\right)^2\right] \quad (7)$$

a weighted sum of the interference-plus-noise power seen by the users. For the case of white noise and equal powers, use of TSC_g as a performance measure is motivated by the work of Massey and Mittelholzer [14] showing that minimizing TSC_g is equivalent to minimizing the worst case interference seen by any user.

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

From now on, we will focus on TSC_g. It is known [1], [13] that the set $\{\lambda(SDS^T + W)\colon S \in \mathcal{S}\}$ has a Schur-minimum element. Therefore, as C_{sum} is Schur-concave and TSC_g is Schur-convex, the configurations attaining this Schur-minimum element will achieve the maximum C_{sum} and the minimum TSC_g. Hence, the optimal configurations are the same whether we use C_{sum} or TSC_g as performance measure.

VI. MMSE Iteration

Ulukus and Yates [2], [6] propose 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_g. In what follows, we state this algorithm and summarize some known properties. Although the authors of [2] consider the case of white noise and equal received powers, the results hold for arbitrary noise covariance and received user powers.

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

$$\Phi_k(S) = [\,s_1 \quad \cdots \quad s_{k-1} \quad c_k(S) \quad s_{k+1} \quad \cdots \quad s_K\,] \quad (8)$$

which replaces the signature sequence for user k by the corresponding normalized linear MMSE filter. This update strictly decreases TSC_g except when the signature sequence for user k coincides with the MMSE filter.

Lemma 3:

$$\forall\, S \in \mathcal{S}\colon \mathrm{TSC}_g(\Phi_k(S)) \leq \mathrm{TSC}_g(S), \quad (9)$$

with equality iff $s_k = c_k(S)$.

Proof: See [2], [6], □

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

$$S^{(t+1)} = \Phi_{t+1}(S^{(t)}) \quad (10)$$

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[1] implementation. The linear MMSE filter for a user can be implemented blindly [15], without needing knowledge of received powers or signature sequences of other users.

Given any initial configuration $S^{(0)} \in \mathcal{S}$, the sequence $\mathrm{TSC}_g(S^{(t)})$ defined by (10) converges because it is nonincreasing by Lemma 3 and bounded below.

The MMSE update function is defined as

$$\Phi(S) = \Phi_K(\Phi_{K-1}(\cdots \Phi_1(S))). \quad (11)$$

Let F_Φ be the set of fixed configurations of Φ

$$F_\Phi = \{S \in \mathcal{S}\colon \Phi(S) = S\}. \quad (12)$$

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

$$\mathrm{TSC}_g(\Phi(S)) \leq \mathrm{TSC}_g(S), \quad \text{with equality iff } S \in F_\Phi. \quad (13)$$

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

Proof: See [2], [6]. □

The following lemma and theorem (proved in [2] for white noise and equal powers) provide a characterization of the fixed configurations.

Lemma 5: Let $S = [s_1 \cdots s_K] \in \mathcal{S}$. Then $S \in F_\Phi$ if and only if for all $k \in \{1, \ldots, K\}$, s_k is an eigenvector of $SDS^T + W$.

[1]Here, distributed means that it can be implemented in parallel modules with no interaction. The user receivers are in the base station, hence colocated.

Proof: See [2]. The proof there carries over with straightforward modification to the case of possibly different received powers and colored noise. □

Theorem 2: Let $S \in F_\Phi$. Then we have the following.

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

$$(SDS^T + W)s_k = \mu_\ell s_k, \qquad \forall\, k \in \mathcal{J}_\ell \tag{14}$$
$$(SDS^T + W)q_n = \mu_\ell q_n, \qquad \forall\, n \in \mathcal{I}_\ell \tag{15}$$
$$\lambda\left(SDS^T + W\right) = (\underbrace{\mu_1, \ldots, \mu_1}_{|\mathcal{I}_1|}, \ldots, \underbrace{\mu_L, \ldots, \mu_L}_{|\mathcal{I}_L|}) \tag{16}$$
$$\mu_\ell = \frac{1}{|\mathcal{I}_\ell|}\left(\sum_{k \in \mathcal{J}_\ell} p_k + \sum_{n \in \mathcal{I}_\ell} w_n\right) \tag{17}$$
$$s_{k_1}^T s_{k_2} = 0, \qquad \forall\, k_1 \in \mathcal{J}_\ell,\ k_2 \notin \mathcal{J}_\ell \tag{18}$$
$$\{s_k \colon k \in \mathcal{J}_\ell\} \subset \operatorname{span}\{q_n \colon n \in \mathcal{I}_\ell\} \tag{19}$$
$$\mathrm{TSC}_g(S) = \sum_{\ell=1}^{L} \frac{1}{|\mathcal{I}_\ell|}\left(\sum_{k \in \mathcal{J}_\ell} p_k + \sum_{n \in \mathcal{I}_\ell} w_n\right)^2 \tag{20}$$

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 $\mu_1 > \cdots > \mu_L$ be such eigenvalues. From Lemma 5, all s_k are eigenvectors of $SDS^T + W$, so we can partition the set $\{1, \ldots, K\}$ grouping the signatures associated to the same eigenvalues

$$\mathcal{J}_\ell = \{k \in \{1, \ldots, K\} \colon (SDS^T + W)s_k = \mu_\ell s_k\}. \tag{21}$$

The $\mathcal{J}_\ell$ are disjoint, $\bigcup_{\ell=1}^{L} \mathcal{J}_\ell = \{1, \ldots, K\}$, and (14) is satisfied. As $SDS^T + W$ is a symmetric matrix, eigenvectors associated with distinct eigenvalues are orthogonal and (18) is proved. Consider any $\ell \in \{1, \ldots, L\}$ 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:

$$(S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T + W)s_k = \mu_\ell s_k, \qquad \forall\, k \in \mathcal{J}_\ell. \tag{22}$$

Multiplying (22) on the right by $p_k s_k^T$, summing over $k \in \mathcal{J}_\ell$, and operating we obtain

$$W S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T = \mu_\ell S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T - (S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T)^2.$$

Hence, $W S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ is a symmetric matrix, which implies that W and $S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T$ commute for all ℓ, and, thus, 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., [16, Corollary 3 of Theorem 3′ in Ch. VIII]). Now choose the partition of the set $\{1, \ldots, N\}$ as follows:

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

Then (15) is satisfied and (16) 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 $\mathcal{J}_\ell = \emptyset$. Then

$$S_{\mathcal{J}_\ell} D_{\mathcal{J}_\ell} S_{\mathcal{J}_\ell}^T q_n = \begin{cases} (\mu_\ell - w_n)q_n, & \text{if } n \in \mathcal{I}_\ell \\ 0, & \text{if } n \notin \mathcal{I}_\ell. \end{cases} \tag{23}$$

Equations (17), (19), and (20) are straightforward to obtain. □

We remark that the characterization obtained in the proof of Theorem 2 may in general not be the only one satisfying (14)–(20). As an example, let $K = 2$, $N = 2$, $p_1 = p_2 = 4$, $W = \operatorname{diag}(1, 9)$, 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 (14)–(20) 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 (14)–(20) 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 (23)

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

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 $\mathcal{J}_\ell$. Let us define $\mathcal{J}_\ell' = \emptyset$, $\mathcal{J}_\ell'' = \mathcal{J}_\ell$, and

$$\mathcal{I}_\ell' = \{n \in \mathcal{I}_\ell \colon \mu_\ell = w_n\}, \quad \mathcal{I}_\ell'' = \{n \in \mathcal{I}_\ell \colon \mu_\ell > w_n\}.$$

Note that $|\mathcal{J}_\ell''| \geq |\mathcal{I}_\ell''|$ because $\{q_n: n \in \mathcal{I}_\ell''\}$ are orthonormal eigenvectors of $S_{\mathcal{J}_\ell''} D_{\mathcal{J}_\ell''} S_{\mathcal{J}_\ell''}^T$ associated with nonzero eigenvalues, and hence, $S_{\mathcal{J}_\ell''}$ has rank $|\mathcal{I}_\ell''|$ and $|\mathcal{J}_\ell''|$ columns.

A new characterization satisfying (14)–(20) (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 $n \in \mathcal{I}_\ell'$, $(\emptyset\{n\})$.

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 $\hat{L}$, $(\hat{\mathcal{J}}_1, \ldots, \hat{\mathcal{J}}_{\hat{L}})$, $(\hat{\mathcal{I}}_1, \ldots, \hat{\mathcal{I}}_{\hat{L}})$, $\hat{\mu}_1 \geq \cdots \geq \hat{\mu}_{\hat{L}}$ 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 [17] with respect to the dynamics (10) as

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

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 (10) converges to the set of fixed configurations.

Lemma 7: Given any $S^{(0)} \in \mathcal{S}$

$$\omega_\Phi(S^{(0)}) \subset F_\Phi. \tag{24}$$

Proof: If $S \in \omega_\Phi(S^{(0)})$ then $\exists t_1 < t_2 < \cdots$ such that $\lim_{m\to\infty} S^{(t_m)} = S$. For some $k \in \{1, \ldots, K\}$, $t_m - k$ 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

$$\mathrm{TSC}_g(\Phi_{k+1}(S)) < \mathrm{TSC}_g(S).$$

Let $\Delta = \mathrm{TSC}_g(S) - \mathrm{TSC}_g(\Phi_{k+1}(S))$. As TSC_g is continuous, there exists p such that for all $m > p$ it holds that

$$\mathrm{TSC}_g(S^{(t_m'+1)}) < \mathrm{TSC}_g(S^{(t_m')}) - \frac{\Delta}{2}.$$

Thus,

$$\mathrm{TSC}_g(S^{(t_{m+1}')}) < \mathrm{TSC}_g(S^{(t_m')}) - \frac{\Delta}{2}, \qquad \text{for } m > p$$

and, therefore, $\mathrm{TSC}_g(S^{(t_m')}) \to -\infty$ as $m \to \infty$. This is a contradiction because TSC_g is positive, and, thus, we obtain $\Phi_{k+1}(S) = S$.

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_g is a continuous function, this implies that

$$\lim_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) \in T_F$$

where

$$T_F = \{\mathrm{TSC}_g(S): S \in F_\Phi\}. \tag{25}$$

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 $\{1, \ldots, K\}$ in L subsets. Hence, as $L \in \{1, \ldots, N\}$

$$|T_F| \leq \sum_{L=1}^{N} L^{K+N}. \tag{26}$$

Let τ be the minimum of TSC_g

$$\tau = \min_{S\in\mathcal{S}} \mathrm{TSC}_g(S). \tag{27}$$

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

$$\Omega = \{S \in \mathcal{S}: \mathrm{TSC}_g(S) = \tau\}. \tag{28}$$

Clearly, $\Omega \subset F_\Phi$: For any $S \in \mathcal{S}$, $\tau \leq \mathrm{TSC}_g(\Phi(S))$, and by Lemma 4, $\mathrm{TSC}_g(S) \geq \mathrm{TSC}_g(\Phi(S))$. If $S \in \Omega$ then $\tau = \mathrm{TSC}_g(S)$ and, therefore, $\mathrm{TSC}_g(\Phi(S)) = \mathrm{TSC}_g(S)$, which again by Lemma 4 implies $S \in F_\Phi$. But it is easy to see that F_Φ contains nonoptimal configurations, that is, $F_\Phi \neq \Omega$ except for the trivial case $N = 1$. As an example, take $N \geq 2$ and let $w_1 \leq \cdots \leq w_N$ be the ordered eigenvalues of W, and $q_1, \ldots, q_N$ be an orthogonal basis of associated eigenvectors. Then, if we take $s_k = q_N$ for all $k \in \{1, \ldots, K\}$, 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 $k \in \{2, \ldots, K\}$, the new configuration S' attains a lower TSC_g value: $\mathrm{TSC}_g(S') < \mathrm{TSC}_g(S)$. Hence, $S \notin \Omega$. Actually, S attains the global maximum of the TSC_g over $\mathcal{S}$.

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

VII. Global Optimal Configurations

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

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

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 (which implies $SDS^T + W$ is 2.6 times the 2×2 identity matrix). But, being that SDS^T and W are 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 $\mu_1 \geq \cdots \geq \mu_L$, 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 (14)–(20).

2) For each $\ell \in \{1, \ldots, L\}$

$$\mu_\ell \geq \max_{n\in\mathcal{I}_\ell} w_n \tag{29}$$

$$\mu_\ell \geq \max_{1\leq M\leq \min(|\mathcal{I}_\ell|, |\mathcal{J}_\ell|)} \frac{1}{M}\sum_{m=1}^{M}\left(p_m^\ell + w_m^\ell\right) \tag{30}$$

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

Proof: See [18]. □

Hence, the problem of minimizing $\mathrm{TSC}_g(S)$ over $S \in \mathcal{S}$ is equivalent to minimizing (20) over all partition pairs that satisfy (29), (30). Next we present an algorithm proposed in [5], [12] that solves this optimization problem.

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

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

Syntax:

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

Update:

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

2) Let

$$\mu_1 = \max\left(w_N, \frac{1}{N}\left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right), \max_{1\leq M\leq \overline{M}} \frac{1}{M}\sum_{m=1}^{M}(p_m + w_m)\right)$$

where $\overline{M} = \min(N-1, K)$.

3) a) If $\mu_1 = w_N$ then
- Let $\mathcal{J}_1 = \emptyset$ $\mathcal{I}_1 = \{N\}$, $\hat{M} = 0$.
- Call

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

b) Else if

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

then let $\mathcal{J}_1 = \{1, \ldots, K\}$, $\mathcal{I}_1 = \{1, \ldots, N\}$, $L' = 0$ and $\hat{M} = 0$.

c) Else if

$$\mu_1 = \frac{1}{M}\sum_{m=1}^{M}(p_m + w_m)$$

for some $M \in \{1, \ldots, \overline{M}\}$, let $\hat{M}$ be the maximum such M, and
- Let $\mathcal{J}_1 = \{1, \ldots, \hat{M}\}, \mathcal{I}_1 = \{1, \ldots, \hat{M}\}$.
- 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 $\mu_\ell = \mu'_{\ell-1}$, $\mathcal{J}_\ell = \mathcal{J}'_{\ell-1} + \hat{M}$, $\mathcal{I}_\ell = \mathcal{I}'_{\ell-1} + \hat{M}$, 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 some simple facts 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 $\mu_1 \geq \cdots \geq \mu_L$.

Proof: See [1] or [12]. □

As proved in the following lemma, the partitions output by Algorithm 1 satisfy conditions (29), (30) 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 (14)–(20) are satisfied. In particular

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

Proof: See [1], [18]. □

The optimality of Algorithm 1 has been proved in [1], [12]. The rest of this section presents an alternative proof. The results will be useful in the next section when we analyze the local minima of TSC_g.

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) $|\mathcal{J}_{\ell_1}| \le |\mathcal{I}_{\ell_1}|$;
2) $p_{k_1} \ge p_{k_2}$, for all $k_1 \in \mathcal{J}_{\ell_1}$, $k_2 \in \mathcal{J}_{\ell_2}$;
3) If $\mathcal{J}_{\ell_1} \ne \emptyset$ then $w_{n_1} \le w_{n_2}$, for all $n_1 \in \mathcal{I}_{\ell_1}$, $n_2 \in \mathcal{I}_{\ell_2}$.

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

$$\sum_{\ell=1}^{\ell'} \mu_\ell |\mathcal{I}_\ell| = \sum_{m=1}^{M} (p_m + w_m) + \sum_{r=0}^{R-1} w_{N-r} \qquad (31)$$

and

$$M + R = \sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell|. \qquad (32)$$

Proof: Consider any $\ell' \in \{1, \ldots, L-1\}$. From (17)

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

Define $\mathcal{J} = \bigcup_{\ell=1}^{\ell'} \mathcal{J}_\ell$ and $\mathcal{I} = \bigcup_{\ell=1}^{\ell'} \mathcal{I}_\ell$. 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 $\mathcal{L} = \{\ell \in \{1, \ldots, \ell'\} : \mathcal{J}_\ell \ne \emptyset\}$, $\mathcal{I}' = \bigcup_{\ell \in \mathcal{L}} \mathcal{I}_\ell$, and $\mathcal{I}'' = \mathcal{I} \setminus \mathcal{I}'$. Hence,

$$\sum_{\ell=1}^{\ell'} \mu_\ell |\mathcal{I}_\ell| = \sum_{k \in \mathcal{J}} p_k + \sum_{n \in \mathcal{I}'} w_n + \sum_{n \in \mathcal{I}''} w_n. \qquad (33)$$

Consider $\ell \in \mathcal{L}$. As $\mathcal{J}_\ell \ne \emptyset$, $|\mathcal{J}_\ell| \ge |\mathcal{I}_\ell|$ (see Condition 1 in Lemma 6). As $\ell \le \ell' < L$, by Condition 1 in Definition 1, $|\mathcal{J}_\ell| \le |\mathcal{I}_\ell|$. 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 (32) is verified.

As $|\mathcal{J}| = M$ (recall $p_1 \ge \cdots \ge p_K$)

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

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 $\ell_1 \le \ell'$ and $m \in \mathcal{J}_{\ell_2}$ for some $\ell_2 > \ell'$, which contradicts Condition 2 of Definition 1. Therefore,

$$\sum_{k \in \mathcal{J}} p_k = \sum_{m=1}^{M} p_m. \qquad (34)$$

As $|\mathcal{I}''| = R$ (recall $w_1 \le \cdots \le w_N$)

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

Assume the preceding inequality is strict. This implies that there exist $n \in \mathcal{I}''$ and $m \in \{N-R+1, \ldots, N\} \setminus \mathcal{I}''$ such that $w_n < w_m$. 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 $\ell_2 \le \ell'$, $\mathcal{J}_{\ell_2} = \emptyset$ and $\mathcal{I}_{\ell_2} = \{n\}$ (see Lemma 6). Then $\mu_{\ell_2} = w_n$. As $m \in \mathcal{I}_{\ell_1}$, $\mu_{\ell_1} \ge w_m$. Therefore, $\mu_{\ell_1} \ge w_m > w_n = \mu_{\ell_2}$. Hence (recall $\mu_1 \ge \cdots \ge \mu_L$) $\ell_1 < \ell_2$. So $\ell_1 < \ell'$ and as $m \notin \mathcal{I}''$ we must have $\ell_1 \in \mathcal{L}$, that is, $\mathcal{J}_{\ell_1} \ne \emptyset$. But then, by Condition 3 of Definition 1 we should have $w_m \le w_n$, which is a contradiction. Therefore,

$$\sum_{n \in \mathcal{I}''} w_n = \sum_{r=0}^{R-1} w_{N-r}. \qquad (35)$$

As $|\mathcal{I}'| = M$ (recall $w_1 \le \cdots \le w_N$)

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

Assume the preceding 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 $\ell_2 \le \ell'$. Then, as $m \notin \mathcal{I}'$, we have $\mathcal{J}_{\ell_2} = \emptyset$ and so $\mu_{\ell_2} = w_m < w_n \le \mu_{\ell_1}$. Hence, $\ell_1 < \ell_2$. Now assume $\ell_2 \le \ell'$. As $n \in \mathcal{I}'$ we have $\ell_1 \in \mathcal{L}$, so $\mathcal{J}_{\ell_1} \ne 0$. But then, by Condition 3 of Definition 1 we should have $w_n \le w_m$, which is a contradiction. Therefore,

$$\sum_{n \in \mathcal{I}'} w_n = \sum_{m=1}^{M} w_m. \qquad (36)$$

Now (31) follows from (33)–(36). □

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

Then for all $S \in F_\Phi$

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

Proof: Let

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

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

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

We want to prove that λ majorizes λ^*. Suppose the statement is not true. 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

$$\sum_{\ell=1}^{\ell'-1} |\mathcal{I}_\ell^*| < V \quad \text{and} \quad \sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell^*| \ge V.$$

[2]Note that the components of λ^* are ordered nonincreasing, but the components of λ are ordered according to the noise eigenvalues.

Define

$$\hat{V} = \sum_{\ell=1}^{\ell'} |\mathcal{I}_\ell^*|.$$

For all $m \in \{V+1, \ldots, \hat{V}\}$ we have $\lambda_m^* = \mu_{\ell'}^* = \lambda_V^* > \lambda_{[V]} \geq \lambda_{[m]}$. Therefore,

$$\sum_{m=1}^{\hat{V}} \lambda_{[m]} < \sum_{m=1}^{\hat{V}} \lambda_m^*. \tag{37}$$

Clearly, $\hat{V} < N$ because

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

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

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

with $1 \leq M \leq \min(N-1, K)$, $0 \leq R \leq N-M-1$, and $M + R = \hat{V}$. Hence, by (37)

$$\sum_{m=1}^{\hat{V}} \lambda_{[m]} < \sum_{m=1}^{M} (p_m + w_m) + \sum_{r=0}^{R-1} w_{N-r}. \tag{39}$$

Now for $n \in \{1, \ldots, N\}$ let γ_n be the eigenvalue of SDS^T associated with q_n, that is, $\gamma_n = \lambda_n - w_n$. 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

$$(\gamma_{[1]}, \ldots, \gamma_{[\min(K,N)]}, \underbrace{0, \ldots, 0}_{K-\min(K,N)}) \text{ majorizes } (p_1, \ldots, p_K). \tag{40}$$

Let $A_M \subset \{1, \ldots, N\}$with $|A_M| = M$ and

$$\sum_{n \in A_M} \gamma_n = \sum_{m=1}^{M} \gamma_{[m]}.$$

Define $B_M = \{N-R+1, \ldots, N\} \setminus A_M$. Take any subset $C_M \subset \{1, \ldots, N\} \setminus (A_M \cup B_M)$ with $|C_M| = R - |B_M|$ (note that $|B_M| \leq R$). This is always possible because

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

as $\hat{V} < N$.

Now from the definition of A_M and using (40) we get

$$\begin{aligned} \sum_{m \in A_M} \lambda_m &= \sum_{m \in A_M} (\gamma_m + w_m) \\ &= \sum_{m=1}^{M} \gamma_{[m]} + \sum_{m \in A_M} w_m \\ &\geq \sum_{m=1}^{M} p_m + \sum_{m \in A_M} w_m. \end{aligned} \tag{41}$$

As SDS^T is nonnegative definite, γ_n is nonnegative 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 (41)

$$\sum_{m \in A_M \cup B_M \cup C_M} \lambda_m \geq \sum_{m=1}^{M} p_m + \sum_{m \in A_M \cup B_M \cup C_M} w_m. \tag{42}$$

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,

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

Introducing this inequality in (42) we obtain

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

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 (37). 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 $k \in \{1, \ldots, K\}$, 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

$$\hat{S}^{(t+1,k)} = \left[s_1^{(t)} \quad \cdots \quad s_{k-1}^{(t)} \quad v_k \quad s_{k+1}^{(t)} \quad \cdots \quad s_K^{(t)} \right].$$

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

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

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

$$\lambda(S^{(t)} D (S^{(t)})^T + W) \text{ majorizes } \lambda(S^{(t+1)} D (S^{(t+1)})^T + W). \tag{43}$$

Also, for any $k \in \{1, \ldots, K\}$, we can apply Lemma 1 with $H = S_k^{(t)} D_k (S_k^{(t)})^T + W$, $v = v_k$, and $s = c_k(S)$ 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, $\mathrm{TSC}_g(\hat{S}^{(t+1,k)}) \leq \mathrm{TSC}_g(\Phi_k(S^{(t)}))$ due to the Schur-convexity of $\mathrm{TSC}_g(\cdot)$. Hence, for all $k \in \{1, \ldots, K\}$

$$\mathrm{TSC}_g(S^{(t+1)}) \leq \mathrm{TSC}_g(\Phi_k(S^{(t)})). \tag{44}$$

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, (43) implies

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

Take any $k \in \{1, \ldots, K\}$. Using (43), (44), and Lemma 3 we can write

$$\begin{aligned} \mathrm{TSC}_g(S^{(t_{m+1})}) &\leq \mathrm{TSC}_g(S^{(t_m+1)}) \\ &\leq \mathrm{TSC}_g(\Phi_k(S^{(t_m)})) \leq \mathrm{TSC}_g(S^{(t_m)}) \end{aligned} \tag{45}$$

where the first inequality follows from (43) because $\mathrm{TSC}_g(\cdot)$ is Schur-convex and the last one from Lemma 3. Letting $m \to \infty$ in (45), by continuity of $\mathrm{TSC}_g(\cdot)$ and $\Phi_k(\cdot)$ we obtain

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

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

Theorem 5: Let

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

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

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

Proof: Use Theorems 4 and 3 and Lemma 11. □

Corollary 1: Let

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

Then

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

is a Schur-minimal element of $\{\lambda(SDS^T + W) \colon S \in \mathcal{S}\}$, and

$$\min_{S \in \mathcal{S}} \mathrm{TSC}_g(S) = \sum_{\ell=1}^{L^*} |\mathcal{I}_\ell^*| (\mu_\ell^*)^2$$

$$\max_{S \in \mathcal{S}} C_{\mathrm{sum}}(S) = \frac{1}{2} \sum_{\ell=1}^{L^*} |\mathcal{I}_\ell^*| \log(\mu_\ell^*) - \frac{1}{2} \log\det(W).$$

Proof: Follows from Theorem 5 and Lemma 10 because TSC_g is Schur-convex and C_{sum} is Schur-concave. □

VIII. Local Minima of TSC_g

In this section, we will prove an important property of the TSC_g 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

$$d(S, S') = \max_{k=1 \cdots K} \arccos(s_k^T s_k'). \tag{46}$$

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

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

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

$$B[S, \theta] = \{S' \in \mathcal{S} \colon d(S', S) \leq \theta\}. \tag{47}$$

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

Lemma 13: If TSC_g 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: See [7]. □

Corollary 2: If TSC_g 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_g are fixed configurations of the MMSE update. Hence, in what follows, we can associate with each local minimum of TSC_g the characterization of Lemma 6. The next three lemmas, which use the same ideas as in [8], present necessary conditions on this characterization for a configuration to be a local minimum of TSC_g.

Lemma 14: Let TSC_g 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: Suppose the statement of the lemma does not hold, that is, $p_{k_1} < p_{k_2}$. 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

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

where

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

Hence,

$$\mathrm{TSC}_g(S') = \mathrm{TSC}_g(S) + 2\mathrm{tr}((SDS^T + W)\Delta) + \mathrm{tr}(\Delta^2).$$

Using (14) and (18) we obtain

$$\begin{aligned}\mathrm{TSC}_g(S') = &\mathrm{TSC}_g(S) - 2(\mu_{\ell_1} - \mu_{\ell_2})\left(\alpha^2 p_{k_1} - \beta^2 p_{k_2}\right)\\ &+ 2\left(\alpha^2 p_{k_1} - \beta^2 p_{k_2}\right)^2\\ &+ 2\left(p_{k_1}\alpha\sqrt{1-\alpha^2} + p_{k_2}\beta\sqrt{1-\beta^2}\right)^2.\end{aligned}$$

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

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

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

Lemma 15: Let TSC_g 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 $w_{n_1} \le w_{n_2}$.

Proof: Suppose the statement of the lemma does not hold, that is, $w_{n_1} > w_{n_2}$. 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 $|\alpha_1| \le 1$ and $|\alpha_2| \le 1$. For $k \in \mathcal{J}_{\ell_1}$, we 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

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

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

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

We claim that $\sum_{k\in\mathcal{J}_{\ell_1}} p_k a_k v_k = 0$. To see this, use (23) 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 $S'DS'^T = SDS^T + \Delta_1 + \Delta_2$ where

$$\begin{aligned}\Delta_1 = &\alpha_1^2 P_1\left(q_{n_2} q_{n_2}^T - q_{n_1} q_{n_1}^T\right)\\ &+ \alpha_1\sqrt{1-\alpha_1^2}\, P_1\left(q_{n_1} q_{n_2}^T + q_{n_2} q_{n_1}^T\right)\\ \Delta_2 = &\alpha_2^2 P_2\left(q_{n_1} q_{n_1}^T - q_{n_2} q_{n_2}^T\right)\\ &+ \alpha_2\sqrt{1-\alpha_2^2}\, P_2\left(q_{n_1} q_{n_2}^T + q_{n_2} q_{n_1}^T\right)\end{aligned}$$

with $P_1 = \sum_{k\in\mathcal{J}_{\ell_1}} p_k a_k^2$, $P_2 = \sum_{k\in\mathcal{J}_{\ell_2}} p_k a_k^2$. Now

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

and after some manipulation we get

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

Hence,

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

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

From (23) follows that $\mu_{\ell_1} = P_1 + w_{n_1}$, $\mu_{\ell_2} = P_2 + w_{n_2}$. 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

$$\begin{aligned}&\mathrm{TSC}_g(S) - \mathrm{TSC}_g(S')\\ &\quad = \frac{2(\mu_{\ell_1} - \mu_{\ell_2})\alpha_1^2 P_1 (w_{n_1} - w_{n_2})}{P_1 + w_{n_1} - w_{n_2}} + o(\alpha_1^3).\end{aligned}$$

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

Lemma 16: Let TSC_g 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 $|\mathcal{J}_\ell| \le |\mathcal{I}_\ell|$.

Proof: Suppose the statement of the lemma does not hold. 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

$$\mathrm{TSC}_g(S) - \mathrm{TSC}_g(S') = 2\epsilon^2(\mu_{\ell_1} - \mu_{\ell_2})\|D_{\mathcal{J}_{\ell_1}} v\|^2 + o(\epsilon^3). \tag{48}$$

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

Theorem 6: Let TSC_g 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 $\mu_\ell > \mu_L$, Lemma 14 implies $|\mathcal{J}_\ell| \le |\mathcal{I}_\ell|$. 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: The local minima of TSC_g are global, i.e., if TSC_g has a local minimum at $S \in \mathcal{S}$, then $S \in \Omega$.

Proof: Assume TSC_g 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_g is Schur-convex $\mathrm{TSC}_g(S) \le \mathrm{TSC}_g(S')$ for all $S' \in \mathcal{S}$, that is, $\mathcal{S} \in \Omega$. □

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

Hence, Theorem 7 implies that all the nonoptimal fixed configurations are unstable equilibria of the MMSE update. If a fixed configuration S does not achieve the minimum of TSC_g, then there exist arbitrarily small perturbations such that if the MMSE iteration is started from these perturbed configurations, the TSC_g converges as $t \to \infty$ to a value strictly smaller than $\mathrm{TSC}_g(S)$. 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 $\lim_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) < \mathrm{TSC}_g(S)$.

Proof: As $S \in F_\Phi \backslash \Omega$, TSC_g 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 $\mathrm{TSC}_g(S') < \mathrm{TSC}_g(S)$. If we start the MMSE iteration with $S^{(0)} = S'$, as $\mathrm{TSC}_g(S^{(t)})$ is nonincreasing, we get

$$\lim_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) \le \mathrm{TSC}_g(S') \qquad \square$$

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

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

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

IX. Noisy MMSE Iteration

Our last observation on the TSC_g 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 toward v_2

$$h(v_1, v_2, \theta) = \cos\theta v_1 + \sin\theta v_2. \tag{49}$$

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

$$h(S, R, \theta) = [\, h(s_1, r_1, \theta_1) \quad \cdots \quad h(s_K, r_K, \theta_K)\,].$$

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

$$S^{(t+1)} = h(\Phi(S^{(t)}), R^{(t+1)}, \theta^{(t+1)}) \tag{50}$$

where $r_k^{(t)}, \theta_k^{(t)}$ ($k \in \{1, \ldots, K\}$, $t \in \mathbb{N}$) 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 kth column of $\Phi(S^{(t-1)})$. In words, the MMSE noisy update consists of applying the MMSE update (10) 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 VI that the (noiseless) MMSE iteration approaches the set of fixed configurations as $t \to \infty$. In Section VIII, we have seen that TSC_g has no other local minima than the global ones. Hence, if we start with any configuration that does not attain the global minimum of TSC_g and perturb it a little, there will be a nonzero probability of getting a new configuration with a lower TSC_g. 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 (50) with $\theta_{\max}^{(t)} = \theta_{\max}$ for all t, satisfies

$$\limsup_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) \le_{\text{a.s.}} \tau + \delta. \tag{51}$$

Proof: Without loss of generality, assume δ is small enough so that if $S \in F_\Phi$ and $\mathrm{TSC}_g(S) \le \tau + \delta$ then $\mathrm{TSC}_g(S) = \tau$. This can be done because, by Theorem 2, the set T_F has a finite number of elements (recall (26)). Define the sets

$$V_1 = \{S \in \mathcal{S}\colon \mathrm{TSC}_g(S) \ge \tau + \delta\}$$

and

$$V_2 = \{S \in \mathcal{S}\colon \mathrm{TSC}_g(S) \le \tau + \delta\}.$$

As $\mathrm{TSC}_g(\cdot)$ is continuous, V_1 and V_2 are compact sets. If $V_1 = \emptyset$, then (51) is trivially satisfied. Hence, in what follows we assume $V_1 \neq \emptyset$. Let

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

Note that $\theta_{\max}$ 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 $\mathrm{TSC}_g(S) \geq \tau + \delta$ and $S = \Phi(S'')$ for some $S'' \in V_2$. Therefore, $\mathrm{TSC}_g(S'') \leq \tau + \delta$ and we get

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

and so

$$\mathrm{TSC}_g(S) = \mathrm{TSC}_g(S'') = \tau + \delta.$$

By (13), this implies $S = S''$ and, thus, $S \in F_\Phi$. But then, by our assumption that δ was small enough, we must have $\mathrm{TSC}_g(S) = \tau$ which contradicts $\mathrm{TSC}_g(S) = \tau + \delta$.

Due to our choice of $\theta_{\max}$, 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

$$\beta(S) = \min\{\mathrm{TSC}_g(S') \colon S' \in B[S, \theta_{\max}]\}.$$

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

$$\gamma = \min\{\mathrm{TSC}_g(S) - \beta(S) \colon S \in V_1\}$$

which is well defined because $(\mathrm{TSC}_g - \beta)(\cdot)$ 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 $\beta(S) = \mathrm{TSC}_g(S)$. But this means that S is a local minimum of $\mathrm{TSC}_g(\cdot)$. Thus, by Theorem 7, S must be a global minimum of $\mathrm{TSC}_g(\cdot)$ and, therefore, $\mathrm{TSC}_g(S) = \tau$ which contradicts $S \in V_1$.

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

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

where r_k, θ_k $(k \in \{1, \dots, K\})$ 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 kth column of $\Phi(S)$. Note that $P(S)$ is a continuous function of S because $\mathrm{TSC}_g(\cdot)$, $\Phi(\cdot)$, 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 the following two cases.

- Assume $\Phi(S) \in V_1$. By definition of γ, there exists $S' \in B[\Phi(S), \theta_{\max}]$ such that $\mathrm{TSC}_g(\Phi(S)) - \mathrm{TSC}_g(S') > \frac{\gamma}{2}$. By continuity of $\mathrm{TSC}_g(\cdot)$ 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

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

  which contradicts $P(S) = 0$.

- Assume $\Phi(S) \notin V_1$. Then, $\mathrm{TSC}_g(\Phi(S)) < \tau + \delta$ and thus by continuity of $\mathrm{TSC}_g(\cdot)$ 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

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

  which contradicts $P(S) = 0$.

Define

$$M = \left(\sum_{k=1}^{K} p_k + \sum_{n=1}^{N} w_n\right)^2.$$

Note that for all $S \in \mathcal{S}$, $\mathrm{TSC}_g(S) \leq M$. Let $Q = \lceil \frac{2(M-\tau-\delta)}{\gamma} \rceil$. Let E_t denote the event that $S^{(t)} \in V_2$. Write $z_m = \Pr(E_{Qm})$. Then

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

Because of our choice of $\theta_{\max}$, $E_t \subset E_{t+1}$. Therefore, $\Pr\left(E_{Q(m+1)} | E_{Qm}\right) = 1$ and

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

Let F_t be the event $\mathrm{TSC}_g(S^{(t)}) \leq \mathrm{TSC}_g(S^{(t-1)}) - \frac{\gamma}{2}$, and define $G_t = E_t \cup F_t$ (that is, G_t denotes the event $\mathrm{TSC}_g(S^{(t)}) \leq \max\{\mathrm{TSC}_g(S^{(t-1)}) - \frac{\gamma}{2}, \tau + \delta\}$).

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

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

where the last equality follows from the fact that if the following inequality holds for all $q \in \{1, \dots, Q\}$:

$$\mathrm{TSC}_g(S^{Qm+q}) \leq \mathrm{TSC}_g(S^{Qm+q-1}) - \frac{\gamma}{2}$$

then $\mathrm{TSC}_g(S^{Q(m+1)}) \leq \mathrm{TSC}_g(S^{Qm}) - Q\frac{\gamma}{2} \leq \tau + \delta$.

Therefore,

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

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

$$1 - z_m \leq (1 - z_0)\left(1 - p^Q\right)^m \leq (1 - p^Q)^m.$$

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 (51) 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 (50) satisfies

$$\lim_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) =_{\text{a.s.}} \tau. \tag{52}$$

Proof: Take any decreasing sequence δ_m such that $\lim_{m\to\infty} \delta_m = 0$, and any $q \in (0, 1)$. Fix m. As shown in the proof of Theorem 8, we can find $\hat{\theta}_m$ such that the noisy MMSE iteration (50) with $\theta_{\max}^{(t)} = \hat{\theta}_m$ satisfies

$$\Pr\left(\mathrm{TSC}_g(S^{(t)}) \leq \tau + \delta_m\right) \to 1, \qquad \text{as } t \to \infty$$

uniformly in the initial condition $S^{(0)}$. Thus, there exists l_m such that for all $S^{(0)}$ and all $t \geq l_m$,

$$\Pr\left(\mathrm{TSC}_g(S^{(t)}) \leq \tau + \delta_m\right) > q.$$

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 that

$$\Pr\left(\mathrm{TSC}_g(S^{(L_{m+z})}) \leq \tau + \delta_m\right) > 1 - (1 - q)^z.$$

This implies

$$\limsup_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) \leq_{\text{a.s.}} \tau + \delta_m, \qquad \text{for all } m.$$

Making $m \to \infty$ we get $\limsup_{t\to\infty} \mathrm{TSC}_g(S^{(t)}) \leq_{\text{a.s.}} \tau$. As $\mathrm{TSC}_g(S^{(t)}) \geq \tau$ for all t, we get the desired result. □

X. Conclusion

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 were proposed, sum capacity and TSC_g, and we observed that the optimal configurations for both are the same. The MMSE iteration is an iterative procedure amenable to distributed implementation that decreases the generalized total square correlation at each iteration. However, it does not guarantee convergence to the minimum TSC_g. We have shown that TSC_g has no local minima other than the global ones, 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. Viswanath and V. Anantharam, "Total capacity of multiaccess vector channels," Univ. Calif., Berkeley, Electronics Res. Lab., Berkeley, CA, Memo. UCB/ERL M99/47, May 1999.

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

[3] S. Verdú, "Capacity region of Gaussian CDMA channels: The symbol-synchronous case," in *Proc. 24th Allerton Conf. Communications, Control and Computing*, Monticello, IL, 1986, pp. 1025–1034.

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

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

[6] S. Ulukus and R. Yates, "Iterative signature adaptation for capacity maximization of CDMA systems," in *Proc. 36th Allerton Conf. Communications, Control and Computing*, Monticello, IL, 1998, pp. 506–515.

[7] C. Rose, S. Ulukus, and R. Yates, "Interference avoidance for wireless systems," in *Proc. Vehicular Technology Conf.*, vol. 2, Tokyo, Japan, 2000, pp. 901–906.

[8] C. Rose, "CDMA codeword optimization: Interference avoidance and convergence via class warfare," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2368–2382, Sept. 2001.

[9] P. Anigstein and V. Anantharam, "Ensuring convergence of the MMSE iteration for interference avoidance to the global optimum," in *Proc. 38th Allerton Conf. Communications, Control and Computing*, Monticello, IL, 2000.

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

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

[12] P. Viswanath and V. Anantharam, "Optimal sequences for CDMA with colored noise: A Schur-saddle function property," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1295–1318, June 2002.

[13] P. Viswanath, "Capacity of vector multiple access channels," Ph.D. dissertation, Univ. Calif., Berkeley, Elec. Eng. Comput. Ssi. Dept., Berkeley, CA, 2000.

[14] J. Massey and T. Mittelholzer, "Welch's bound and sequence sets for code-division multiple-access systems," in *Sequences II: Methods in Communication, Security and Computer Science*, R. Capocelli, A. D. Santis, and U. Vaccaro, Eds. New York: Springer-Verlag, 1991.

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

[16] F. Gantmacher, *The Theory of Matrices*. New York: Chelsea , 1960.

[17] S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*. New York: Springer-Verlag, 1999.

[18] P. Anigstein and V. Anantharam, "Iterative contruction of optimal signature sequences for CDMA," Univ. Calif. Berkeley Electron. Res. Lab., Berkeley, CA, Memo. UCB/ERL M01/24, Feb. 2001.