Ensuring convergence of the MMSE iteration for interference avoidance to the global optimum *

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

Abstract

Viswanath and Anantharam [8] characterize the sum capacity of multiaccess vector channels. For given number of users, received powers, spreading gain and noise covariance matrix in a Direct-Sequence (DS) CDMA system, [8] presents a combinatorial algorithm to generate a set of signature sequences that achieves the maximum sum capacity. The optimum sequences sets are generalized Welch bound equality (WBE) sets. These sets also minimize the total square correlation (TSC).

Ulukus and Yates [5] 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 in the TSC at each step. The MMSE iteration has fixed points not only at the optimum generalized WBE sets but also at other sets which are suboptimal. [5] claims that simulations show that when starting with randon sequences, the algorithm converges to optimum sets of sequences, but gives no formal proof.

We show that the TSC 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 optimum 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 an optimum generalized WBE set with probability one for any initial set of sequences.

1 Introduction

We consider a synchronous code-division multiple-access (CDMA) system with K users and processing gain N. The received signal at the base station over one symbol interval will be represented by an N-dimensional column vector y:

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

*This research is supported by EPRI-DoD Center for Complex Interactive Networks, NSF ANI 9872764, ECS 9873086 and IIS 9941569

Here s_k is an N-dimensional column vector corresponding to the signature sequence of user k, assumed unit-energy (i.e. $s_k^T s_k = 1$). The power received from user k is p_k. The information transmitted by user k is modeled by the random variable x_k having zero mean and unit variance. The noise is assumed Gaussian zero-mean with covariance $E[nn^T] = W$, a $K \times K$ symmetric positive definite matrix.

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 (1) can be rewritten as

$$y = SD^{1/2}x + n \tag{2}$$

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 $S \in \mathcal{S}$ where

$$\mathcal{S} \triangleq \left\{ \begin{bmatrix} a_1 & \dots & a_K \end{bmatrix} : a_k \in \mathbb{S}^{N-1}\ \forall k \in \{1, \dots, K\} \right\} \tag{3}$$

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

The sum capacity of channel 2 is [8]

$$C_{sum}(S) = \frac{1}{2}\log\det\left(I + W^{-1}SDS^T\right) = \frac{1}{2}\log\det\left(SDS^T + W\right) - \frac{1}{2}\log\det(W) \tag{4}$$

The problem of maximizing $C_{sum}(S)$ over all $S \in \mathcal{S}$ and finding an optimum signature sequence configuration is solved in [8], where majorization theory is used [2]. The sum capacity is a Schur-concave function of the eigenvalues of $SDS^T + W$. The set of vectors of eigenvalues of $SDS^T + W$ as S varies in $\mathcal{S}$ has a Schur-minimal element [9]. Thus this element corresponds to configurations S that achieve the maximum of C_{sum}.

We define a generalized total square correlation (TSC) function as [4]

$$\text{TSC}(S) = \text{Trace}\left[\left(SDS^T + W\right)^2\right] \tag{5}$$

The TSC is a Schur-convex function of the eigenvalues of $SDS^T + W$. Thus the configurations which achieve the maximum of C_{sum} and those which achieve the minimum of TSC are the same.

Iterative algorithms aimed to minimize the TSC are proposed in [5] and [3]. These algorithms update one signature sequence at a time and are amenable to distributed implementation. The TSC at each iteration is non-increasing, but no formal proof of convergence to the minimum of TSC in the general case is given. A modified algorithm is proposed in [4] to prove convergence, but it has increased complexity and is not suitable for distributed implementation.

We consider the MMSE update algorithm [5]. In the next section we define the MMSE update, state some properties and characterize the fixed configurations. In section 3 we observe and sketch a proof that TSC has no minima other than the global minima. Motivated by this result, in section 4 we define a modified version of the MMSE update adding noise. We prove that if the noise bound is chosen adequately, the noisy MMSE update converges to the optimum TSC almost surely regardless of the initial configuration.

2 MMSE update

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

The MMSE unit-norm linear filter for user k is

$$c_k(S) = \frac{1}{\sqrt{s_k^T \left(S_k D_k S_k^T + W\right)^{-2} s_k}} \left(S_k D_k S_k^T + W\right)^{-1} s_k \tag{6}$$

or equivalently

$$c_k(S) = \frac{1}{\sqrt{s_k^T \left(SDS^T + W\right)^{-2} s_k}} \left(SDS^T + W\right)^{-1} s_k \tag{7}$$

The filter $c_k(S)$ can be shown to minimize the signal-to-interference ratio (SIR) of user k over all linear receivers (for a derivation in the case of white noise $W = wI$ see section 6.2 of [6]).

We define the MMSE user k update function as

$$\Phi_k(S) = \left[\begin{array}{ccccccc} s_1 & \dots & s_{k-1} & c_k(S) & s_{k+1} & \dots & s_K \end{array}\right] \tag{8}$$

which replaces the signature sequence for user k by the corresponding normalized linear MMSE filter. The following lemma [5] states that this update strictly decreases the TSC except when the signature sequence for user k coincides with the MMSE filter.

Lemma 1

$$\forall S \in \mathcal{S}: \ \mathrm{TSC}(\Phi_k(S)) \le \mathrm{TSC}(S), \qquad \textit{with equality iff } s_k = c_k(S) \tag{9}$$

Proof : See [5] ([5] considers the case of white noise ($W = wI$) and equal received powers, but the proof holds for arbitrary W and powers).

□

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

$$S^{(n+1)} = \Phi_{n+1}(S^{(n)}) \tag{10}$$

where we define Φ_n for $n > K$ setting $\Phi_n = \Phi_{n-K}$. This corresponds to replacing each signature sequence using the MMSE update, one at a time. The signature sequences are updated in a deterministic order. This assumption could be relaxed, and with some more work the results could be extended under suitable conditions to partially asynchronous updates [1].

Note that given any initial configuration $S^{(0)} \in \mathcal{S}$, the sequence $\mathrm{TSC}(S^{(n)})$ defined by equation (10) converges because it is non-increasing by lemma 1 and bounded below.

The MMSE update function is defined as

$$\Phi(S) = \Phi_K(\Phi_{K-1}(\dots\Phi_1(S))) \tag{11}$$

By lemma 1 we obtain that

$$\mathrm{TSC}(\Phi(S)) \le \mathrm{TSC}(S), \qquad \text{with equality iff } \Phi(S) = S \tag{12}$$

Let F_Φ be the set of fixed configurations of Φ:

$$F_\Phi = \{S \in \mathcal{S} : \Phi(S) = S\} \tag{13}$$

We summarize some properties of the fixed configurations of the MMSE update.

Lemma 2 *If* $S = \left[\begin{array}{ccc} s_1 & \dots & s_K \end{array}\right] \in F_\Phi$ *then*

(a) For all $k \in \{1,\dots,K\}$*,* s_k *is an eigenvector of* $SDS^T + W$*.*

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

(c) Let n *be the number of distinct eigenvalues* $\mu_1,\dots,\mu_n$ *of* SDS^T+W*. There exist a partition* $J_1,\dots,J_n$ *(with possibly some of the* J_i *empty) of the set* $\{1,\dots,K\}$ *and a partition* $I_1,\dots,I_n$ *of the set* $\{1,\dots,N\}$ *such that for all* $i \in \{1,\dots,n\}$*:*

$$\begin{aligned}
(SDS^T + W)s_k &= \mu_i s_k \quad \forall k \in J_i && (14)\\
s_k^T s_l &= 0 \quad \forall k \in J_i,\ l \in J_j,\ j \neq i && (15)\\
\mu_i &= \frac{\sum_{k\in J_i} p_k + \sum_{m \in I_i} w_m}{|I_i|} && (16)\\
\mathrm{TSC}(S) &= \sum_{i=1}^{n} \frac{\left(\sum_{k\in J_i} p_k + \sum_{m\in I_i} w_m\right)^2}{|I_i|} && (17)
\end{aligned}$$

where $|I_i|$ *is the number of elements in* I_i *and* $w_1,\dots,w_N$ *are the eigenvalues of* W*. Furthermore,* $|I_i|$ *is equal to the multiplicity of* μ_i*.*

Proof : Apply lemma 1 to obtain $s_k = c_k(S)$ and for all $k \in \{1,\dots,K\}$ and thus

$$\left(SDS^T + W\right) s_k = \lambda_k s_k \qquad (18)$$

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

To prove (b) and (c) let n be the number of distinct eigenvalues $\mu_1,\dots,\mu_n$ of $SDS^T + W$. From (a) 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

$$J_i \triangleq \{k \in \{1,\dots,K\} :\ (SDS^T + W)s_k = \mu_i s_k\} \qquad (19)$$

the J_i are disjoint and $\bigcup_{i=1}^n J_i = \{1,\dots,K\}$. As SDS^T+W is a symmetric matrix, eigenvectors associated with distinct eigenvalues are orthogonal and (15) is proved. If we write $S_{J_i} = [s_k, k \in J_i]$ and $D_{J_i} = \mathrm{diag}\left(p_k, k \in J_i\right)$ it follows

$$(S_{J_i} D_{J_i} S_{J_i}^T + W)s_k = \mu_i s_k \quad \forall k \in J_i,\ \ \forall i \in \{1,\dots,n\} \qquad (20)$$

Thus we get

$$(S_{J_i} D_{J_i} S_{J_i}^T + W) S_{J_i} D_{J_i} S_{J_i}^T = \mu_i S_{J_i} D_{J_i} S_{J_i}^T \quad \forall i \in \{1,\dots,n\}$$

which implies that W and $S_{J_i} D_{J_i} S_{J_i}^T$ commute. As $SDS^T = \sum_{i=1}^n S_{J_i} D_{J_i} S_{J_i}^T$ we get that W and SDS^T commute. Therefore there exists an orthonormal basis $q_1,\dots,q_N$ of eigenvectors of W and SDS^T. Thus $q_1,\dots,q_N$ are eigenvectors of $SDS^T + W$. W.l.o.g. we assume $Wq_j = w_j q_j$. Then (16) and (17) are obtained by choosing the partition of the set $\{1,\dots,N\}$ as follows:

$$I_i \triangleq \{j \in \{1,\dots,N\} :\ \left(SDS^T + W\right) q_j = \mu_i q_j\}$$

Note that condition (a) in the lemma above is equivalent to S being a fixed configuration because from (7) $c_k(S) = s_k$ iff s_k is an eigenvector of $SDS^T + W$:

$$S \in F_\Phi \iff \forall k \in \{1, \dots, K\} : s_k \text{ is an eigenvector of } SDS^T + W \tag{21}$$

By condition (c), if S is a fixed configuration we can partition the signatures s_k in mutually orthogonal sets each associated with a distinct eigenvalue of $SDS^T + W$.

Given $S^{(0)} \in \mathcal{S}$ we can define the ω-limit set with respect to the dynamics (10) as:

$$\omega_\Phi(S^{(0)}) = \{S \in \mathcal{S} : \exists n_1 < n_2 < \dots \text{ s.t. } \lim_{m \to \infty} S^{(n_m)} = S\} \tag{22}$$

The following lemma shows that for any initial set of signature sequences, the MMSE update (10) converges to the set of fixed configurations.

Lemma 3 *Given any* $S^{(0)} \in \mathcal{S}$,

$$\omega_\Phi(S^{(0)}) \subset F_\Phi \tag{23}$$

Proof : If $S \in \omega_\Phi(S^{(0)})$ then $\exists m_1 < m_2 < \dots$ s.t. $\lim_{l \to \infty} S^{(m_l)} = S$. For some $k \in \{1, \dots, K\}$, m_l is a multiple of k for infinitely many l, let n_l be the corresponding subsequence. Then $S^{(n_l)} \to S$ as $l \to \infty$. By continuity of Φ_{k+1}, $\Phi_{k+1}(S^{(n_l)}) \to \Phi_{k+1}(S)$ as $l \to \infty$. Now assume $\Phi_{k+1}(S) \neq S$. Then by lemma 1, $\mathrm{TSC}(\Phi_{k+1}(S)) < \mathrm{TSC}(S)$. Let $\Delta = \mathrm{TSC}(S) - \mathrm{TSC}(\Phi_{k+1}(S))$. Then, as TSC is continuous, there exists p such that $\forall l > p$ it is $\mathrm{TSC}(S^{(n_l+1)}) < \mathrm{TSC}(S^{(n_l)}) - \frac{\Delta}{2}$. Thus $\mathrm{TSC}(S^{(n_{l+1})}) < \mathrm{TSC}(S^{(n_l)}) - \frac{\Delta}{2}$ for $p > l$ and therefore $\mathrm{TSC}(S^{(n_l)}) \to -\infty$ as $l \to \infty$. This is a contradiction because TSC is positive, and thus $\Phi_{k+1}(S) = S$. But then $\Phi_{k+1}(S^{(n_l)}) = S^{(n_l+1)} \to \Phi_{k+1}(S) = S$ as $l \to \infty$. 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 update approaches the set of fixed configurations as $n \to \infty$. As TSC is a continuous function, this implies that $\lim_{n \to \infty} \mathrm{TSC}(S^{(n)}) \in T_F$ where

$$T_F = \{\mathrm{TSC}(S) : S \in F_\Phi\} \tag{24}$$

Note that from lemma 2(c), T_F has a finite number of elements because there is a finite number of ways of partitioning the sets $\{1, \dots, K\}$ and $\{1, \dots, N\}$. Let τ be the minimum of the TSC:

$$\tau \triangleq \min_{S \in \mathcal{S}} \mathrm{TSC}(S) \tag{25}$$

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

$$\Omega \triangleq \{S \in \mathcal{S} : \mathrm{TSC}(S) = \tau\} \tag{26}$$

Clearly we have $\Omega \subset F_\Phi$, because if $S \in \Omega$ then $\tau = \mathrm{TSC}(S) \geq \mathrm{TSC}(\Phi(S)) \geq \mathrm{TSC}(S)$ and therefore $\mathrm{TSC}(\Phi(S)) = \tau$ which implies by (12) that $\Phi(S) = S$. But it is easy to see that F_Φ contains non-optimal configurations, i.e. $F_\Phi \neq \Omega$ except for the trivial case $N = 1$. Therefore T_F has more than one element and we cannot conclude that $\lim_{n \to \infty} \mathrm{TSC}(S^{(n)}) = \tau$ as we would like. Simulations suggest that if the initial condition $S^{(0)}$ is chosen randomly, then $\mathrm{TSC}(S^{(n)})$ converges to τ with probability one [5], but no formal proof has been given.

3 Minima of TSC

An important property of the TSC function is 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 $S, S' \in \mathcal{S}$, let

$$d(S,S') = \max_{k=1\dots K} \arccos(s_k^T s'_k) \tag{27}$$

Note that the triangle inequality holds:

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

and d 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:

$$B[S,\theta] = \{S' \in \mathcal{S} : d(S',S) \leq \theta\} \tag{28}$$

Lemma 4 *Local minima of TSC are global, i.e. if $S \notin \Omega$, then for all $\epsilon \in (0,\pi]$ there exists $S' \in B[S,\epsilon]$ with* $\mathrm{TSC}(S') < \mathrm{TSC}(S)$.

Proof : Here we sketch a proof for the case of white noise, i.e. $W = wI$. This can be extended for the case of colored noise [1].

Take any $S \in \mathcal{S}$. Consider two cases:

1. Assume $\exists k \in \{1,\dots,K\}$ 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. Take S' with $s'_j = s_j$ for $j \neq k$ and $s'_k = \alpha s_k + \beta v$, where $\alpha = \cos\left(\frac{\epsilon}{2}\right)$ and $\beta = -\alpha s_k^T v + \mathrm{sign}(s_k^T v)\sqrt{\alpha^2 \left(s_k^T v\right)^2 + 1 - \alpha^2}$. 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 $d(S,S') = \arccos\left(s_k^T s'_k\right) \leq \frac{\epsilon}{2}$. Direct computation shows:

   $$\mathrm{TSC}(S) - \mathrm{TSC}(S') = 2p_k(1-\alpha^2)\left[s_k^T \left(S_k D_k S_k^T + W\right) s_k - \lambda\right] \tag{29}$$

   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 $1 - \alpha^2 = \sin^2\left(\frac{\epsilon}{2}\right) > 0$ and therefore $\mathrm{TSC}(S) > \mathrm{TSC}(S')$.

2. Assume $\forall k \in \{1,\dots,K\}$, s_k is an eigenvector associated with the minimum eigenvalue of $(S_k D_k S_k^T + W)$. Then S is a fixed configuration, and we can partition the K users and the N degrees of freedom as in lemma 2(c). W.l.o.g. assume the eigenvalues of $SDS^T + W$ are ordered $\mu_1 > \mu_2 > \dots > \mu_n$. Let v be an eigenvenctor of $SDS^T + W$ associated with μ_n.

   - Assume there exists $i \in \{1,\dots,n-1\}$ with $|J_i| > |I_i|$. Take S' with $s'_k = s_k$ for $k \notin J_i$ and $s'_k = \cos(\alpha_k)s_k + \sin(\alpha_k)v$ for $k \in J_i$, where α_k are chosen so that the column vector $\alpha \in \mathbb{R}^{|J_i|}$ with elements $(\alpha_k,\ k \in J_i)$ satisfies $S_{J_i} D_{J_i} \alpha = 0$ (this can be done because S_{J_i} has rank $|I_i|$). With this choice, after some manipulation we get:

     $$\mathrm{TSC}(S) - \mathrm{TSC}(S') = 2(\mu_i - \mu_n)\|D_{J_i}\alpha\|^2 + o(\|\alpha\|^4) \tag{30}$$

     So it suffices to make α small enough to get $\mathrm{TSC}(S') < \mathrm{TSC}(S)$ and $d(S,S') = \max_{k \in J_i} \alpha_k \leq \epsilon$.

- If $|J_i| \le |I_i|$ for all $i \in \{1, \ldots, n-1\}$ it can be proved that, in the sense defined in [7], all the users in $\bigcup_{i=1}^{n-1} J_i$ are oversized, and in fact S attains the optimum TSC, i.e. $\mathrm{TSC}(S) = \tau$. To see this note that from (16) we have

$$\mu_i = \frac{\sum_{k \in J_i} p_k}{|I_i|} + w \tag{31}$$

From now on consider $i \in \{1, \ldots, n-1\}$. As $\mu_i > \mu_n$ it can be shown $|J_i| = |I_i|$ (because, when $W = wI$, $|J_i| < |I_i|$ is only possible if $|J_i| = 0$. But then it would be $\mu_i = w \le \mu_n$). Using (20), for all $k \in J_i$,

$$s_k^T \left(S_{J_i} D_{J_i} S_{J_i}^T\right) s_k = p_k + \sum_{l \in J_i \setminus \{k\}} p_l \left(s_k^T s_l\right)^2 = \mu_i - w \ge p_k \tag{32}$$

Then from (31), for all $k \in J_i$ we have $\sum_{l \in J_i} p_l \ge |J_i| p_k$ and therefore $p_k = p_l$ and $s_k^T s_l = 0$ for all $k, l \in J_i$ with $k \ne l$. So we can define p_{J_i} by $p_{J_i} = p_k$ for $k \in J_i$ and using (31) $\mu_i = p_{J_i} + w$. We get $p_{J_1} > p_{J_2} > \ldots > p_{J_{n-1}} > \frac{\sum_{k \in J_n} p_k}{|I_n|}$. This implies that the users in $\bigcup_{i=1}^{n-1} J_i$ are oversized. But each of this users has a signature that is orthogonal to all the other signatures, and thus it is easy to see that S is an optimal signature allocation [7], i.e. $\mathrm{TSC}(S) = \tau$.

$\square$

We observe that lemma 4 implies that all the non-optimal fixed configurations are unstable, in the sense that given $S \in F_\Phi \setminus \Omega$, for all $\epsilon > 0$ there exists $S' \in B[S, \epsilon]$ such that for the MMSE update with $S^{(0)} = S'$ we have $\lim_{n\to\infty} \mathrm{TSC}(S^{(n)}) \le \mathrm{TSC}(S') < \mathrm{TSC}(S)$. On the other hand, for $S \in \Omega$ there exists $\epsilon > 0$ such that for all $S' \in B[S, \epsilon]$ the MMSE update with $S^{(0)} = S'$ satisfies $\lim_{n\to\infty} \mathrm{TSC}(S^{(n)}) = \tau$. This last assertion follows from the fact that T_F is finite and TSC is continuous.

4 Noisy MMSE update

Our last observation on the TSC is key to understand the convergence of the MMSE update. 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:

$$h(v_1, v_2, \theta) \triangleq \cos\theta v_1 + \sin\theta v_2 \tag{33}$$

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

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

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

$$S^{(n+1)} = h(\Phi(S^{(n)}), R^{(n+1)}, \theta^{(n+1)}) \tag{34}$$

where $r_k^{(n)}, \theta_k^{(n)}$, $k \in \{1, \ldots, K\}$, $n \ge 1$ are independent random variables, $\theta_k^{(n)}$ is uniform $(0, \theta_{max}^{(n)})$ and $r_k^{(n)}$ is a random (i.e. "uniform") unit-norm vector orthogonal to

$\left[\Phi(S^{(n-1)})\right]_k$. 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 can now prove that fixing a sufficiently small noise upper bound in the noisy iteration, $S^{(n)}$ can be made to converge to an arbitrary small neighborhood of the optimal set with probability one regardless of the initial configuration.

Theorem 1 *Given any $\delta > 0$ there exists $\theta_{max} > 0$ such that for any initial condition $S^{(0)}$ the MMSE noisy iteration defined by (34) with $\theta_{max}^{(n)} = \theta_{max}$ for all n, satisfies*

$$\limsup \mathrm{TSC}(S^{(n)}) \leq_{\text{a.s.}} \tau + \delta \tag{35}$$

Proof : We present a sketch. For details see [1]. W.l.o.g. 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 lemma 2(c) the set T_F has a finite number of elements. Define the set $V \triangleq \{S \in \mathcal{S} : \mathrm{TSC}(S) \geq \tau + \delta\}$. Let

$$\theta_{max} = \min\{d(S, S') : S \in V \ , \ S' \in \Phi(\overline{V^c})\}$$

where $\overline{V^c}$ is the closure of the complement of V. Note that θ_{max} is well-defined: $d(\cdot, \cdot)$ is a continuous function, V is a compact set, $\overline{V^c}$ is compact and thus $\Phi(\overline{V^c})$ is compact because $\Phi(\cdot)$ is continuous.

We claim $\theta_{max} > 0$ (for details see [1]).

Because of our choice of θ_{max}, if $S^{(n)} \in V^c$ then $S^{(n+1)} \in V^c$ and thus $S^{(n+k)} \in V^c$ for all $k \geq 0$.

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

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

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 $\mathrm{TSC}(\cdot)$ is continuous and the set $B[S, \theta_{max}]$ depends continuously on S. Now define

$$\gamma \triangleq \min\{\mathrm{TSC}(S) - \beta(S) : S \in V\}$$

which is well-defined because $(\mathrm{TSC} - \beta)(\cdot)$ is continuous and $B[S, \theta_{max}]$ is compact.

We claim $\gamma > 0$ ([1]).

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

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

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 (i.e. "uniform") unit-norm vector orthogonal to $[\Phi(S)]_k$. Note that $P(S)$ is a continuous function of S because $\mathrm{TSC}(\cdot)$ is continuous and the probability distributions involved are continuous. Let

$$p \triangleq \min_{S \in V} P(S)$$

We claim $p > 0$ ([1]).

Define $M \triangleq \left(\sum_{k=1}^{K} p_k + \sum_{j=1}^{N} w_j\right)^2$. Note that $\forall S \in \mathcal{S}$, $\text{TSC}(S) \leq M$. Let $L \triangleq \lceil \frac{2(M-\tau-\delta)}{\gamma} \rceil$. Write $x_n = \mathbb{P}\left(\text{TSC}(S^{(nL)}) \leq \tau + \delta\right)$. It is straightforward to see

$$x_{n+1} \geq x_n + \mathbb{P}\left(\text{TSC}(S^{((n+1)L)}) \leq \tau + \delta \mid \text{TSC}(S^{(nL)}) > \tau + \delta\right)(1 - x_n)$$

Now

$$\mathbb{P}\left(\text{TSC}(S^{((n+1)L)}) \leq \tau + \delta \mid \text{TSC}(S^{(nL)}) > \tau + \delta\right) \geq p^L$$

and therefore $x_n \geq 1 - \left(1 - p^L\right)^n$. This implies that with probability 1 for some finite n_0, $S^{(n_0)} \in V^c$ and the theorem statement follows.

□

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

Theorem 2 *There exists a sequence $\theta_{max}^{(n)}$ such that for any initial condition $S^{(0)}$ the MMSE noisy iteration defined by (34) satisfies*

$$\limsup \text{TSC}(S^{(n)}) \to_{\text{a.s.}} \tau \tag{36}$$

Proof : Take a decreasing sequence δ_m with $lim_{m\to\infty}\delta_m = 0$, and take any $q \in (0,1)$. By the proof of theorem 1 we can find $\hat{\theta}_m$ such that, if we fix some m, the noisy MMSE update (34) with $\theta_{max}^{(n)} = \hat{\theta}_m$ satisfies $\mathbb{P}\left(\text{TSC}(S^{(n)}) \leq \tau + \delta_m\right) \to 1$ as $n \to \infty$ uniformly in the initial condition $S^{(0)}$. Thus there exists l_m such that for all $S^{(0)}$ and all $n \geq l_m$, $\mathbb{P}\left(\text{TSC}(S^{(n)}) \leq \tau + \delta_m\right) > q$. Let $L_m = \sum_{i=1}^{m} l_i$. It follows that if we choose $\theta_{max}^{(n)} = \hat{\theta}_m$ for all $n = (1+L_{m-1}), \ldots, L_m$, we obtain that for all $z \geq 0$ it holds $\mathbb{P}\left(\text{TSC}(S^{(L_m+z)}) \leq \tau + \delta_m\right) > 1 - (1-q)^z$. This implies $\limsup \text{TSC}(S^{(n)}) \leq_{\text{a.s.}} \tau + \delta_m$ for all m. Making $m \to \infty$ we get the desired result.

5 Conclusion

The MMSE update is known to decrease the total square correlation. However, it does not guarantee convergence to the minimum TSC. We have observed 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 update converges with probability one to the optimal.

References

[1] P. Anigstein and V. Anantharam, “Ensuring Convergence of the MMSE Iterarion for Interference Avoidance to the Global Optimum,” in preparation.

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

[3] C. Rose, S. Ulukus, and R.D. Yates, “Interference Avoidance for Wireless Systems,” In *Vehicular Technology Conference*, May 2000.

[4] C. Rose, "Interference Avoidance, Sum Capacity and Convergence via Class Warfare," submitted to *IEEE Transactions on Information Theory.* Available at http://www.winlab.rutgers.edu/∼crose.

[5] S. Ulukus and R.D. Yates, "Iterative Construction of Optimal Sequence Sets to Maximize the Capacity of CDMA Systems," submitted to *IEEE Transactions on Information Theory.* Available at http://www.winlab.rutgers.edu/∼ryates.

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

[7] P. Viswanath and V. Anantharam, "Optimal Sequences and Sum Capacity of Synchronous CDMA Systems," *IEEE Transactions on Information Theory*, Vol. 45(6), pp. 1984-1991, Sept. 1999.

[8] P. Viswanath and V. Anantharam, "Total Capacity of Multiaccess Vector Channels," UC Berkeley Electronics Research Laboratory, Memorandum No. UCB/ERL M99/47, May 1999.

[9] P. Viswanath, "Capacity of Vector Multiple Access Channels," PhD. Thesis, UC Berkeley, July 2000.