CDMA Codeword Optimization: interference avoidance and convergence via class warfare

Christopher Rose

June 26, 2001

WINLAB, Dept. of Electrical and Computer Engineering
Rutgers University
94 Brett Road
Piscataway NJ 08854-8060
email: ***crose@winlab.rutgers.edu***

Abstract

Interference avoidance has been shown to reduce total square correlation (TSC) for given ensembles of user signature waveforms (codewords) in a synchronous CDMA system. In all experiments we have conducted, sequential application of interference avoidance produces an optimal codeword set when starting from randomly chosen initial codewords. Here we provide the first formal proof of convergence to optimal codeword ensembles for greedy interference avoidance algorithms augmented by a technique called "class warfare" whereby users which reside in more heavily loaded areas of the signal space purposely interfere with (attack) the reception of users in less crowded areas. Coordination of deliberate interference by a complete class of aggrieved user is also sometimes necessary. Such "attacks" and subsequent codeword adjustment by attacked users are shown to strictly decrease TSC. Along the way we also show, using linear algebra and a variant of stochastic ordering, equivalence between minimization of total square correlation (TSC) and maximization of sum capacity.

KEYWORDS: Codeword optimization, codeword adaptation, adaptive modulation, interference avoidance

Contents

1 Introduction

Interference avoidance has been identified as a method to iteratively obtain optimal signature waveforms (codewords) in multiple access systems [1–5]. The notion behind interference avoidance is that each user, with feedback from the receiver, is allowed to adjust its waveform to minimize interference. Empirically speaking, sequential iterative updates by all users always results in a set of codewords which maximize various measures of system capacity [6–8] and minimize a measure of mutual interference called the total square correlation (TSC). There are even analytic hints that interference avoidance should always converge to an optimal ensemble [3]. Unfortunately, even copious empirical evidence where *never* has a suboptimal set been obtained starting from random codewords [1–5] and strong analytic hints are unsettling and hamper work which depends on provable convergence of greedy interference avoidance [2, 9].[1]

Therefore, in this paper we modify the basic greedy interference avoidance procedure to (a) maximally reduce TSC at each iteration and (b) to include an aggressive component whereby users who reside in a more crowded portion of the signal space deliberately interfere with (attack) other users in more sparsely populated dimensions. The procedure, dubbed *class warfare*, allows escape from suboptimal minima, and coupled to the finite number of possible fixed point TSC values for greedy interference avoidance, forces convergence to codeword sets which minimize the average mutual interference between codewords.

Along the way it will also be shown directly via elementary linear algebra and a variant of stochastic ordering methods that TSC minimization is equivalent to sum capacity maximization, and thereby, that greedy interference avoidance with class warfare achieves signal sets which meet sum capacity bounds. We note that a slightly more compact statement of this equivalence can be obtained by an appeal to matrix majorization theory [8, 11, 12]. However, the direct approach presented here is useful in that it can be wholly understood from first principles without collateral references.

2 Greedy Interference Avoidance: a brief review

We assume that user signals can be represented by $N \times 1$ vectors $\mathbf{s}_m$ in some arbitrary signal space. The power of each signal vector is defined as $p_m = |\mathbf{s}_m|^2$. Information is carried by corresponding independent zero mean, unit mean square b_m and the superposition bathed in some noise process, represented by a $N \times 1$ noise vector $\mathbf{w}$. The result is that the $N \times 1$ received signal vector $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{S}\mathbf{b} + \mathbf{w} \tag{1}$$

[1] Recently, Pablo Anigstein at UC Berkeley [10] has proven stochastic convergence for MMSE interference avoidance [3]. Unfortunately, this simple and elegant proof does not apply to greedy interference avoidance.

where $\mathbf{b}$ is the $M \times 1$ vector of modulations corresponding to each codeword and $\mathbf{S}$ has M columns composed of the $\mathbf{s}_m$. As a concrete example, the $\mathbf{s}_m$ might be the ± 1 chip sequences of standard CDMA codewords. However, more generally, they could also be coefficient sequences for any convenient orthonormal signal representation.

The total interference power experienced by user m assuming a simple matched filter receiver is

$$E\left[\left(\frac{\mathbf{s}_m^\top}{|\mathbf{s}_m|}\left[\mathbf{S}\mathbf{b} - \mathbf{s}_m b_m + \mathbf{w}\right]\right)^2\right] = \frac{\mathbf{s}_m^\top(\mathbf{S}\mathbf{S}^\top + \mathbf{W} - \mathbf{s}_m\mathbf{s}_m^\top)\mathbf{s}_m}{|\mathbf{s}_m|^2} \tag{2}$$

where

$$\mathbf{S}\mathbf{S}^\top = \sum_{m=1}^{M} \mathbf{s}_m\mathbf{s}_m^\top \tag{3}$$

and $\mathbf{W}$ is the covariance matrix of the noise vector $\mathbf{w}$.

Equation (2) suggests that user m could minimize the interference seen at the receiver by choosing $\mathbf{s}_m$ proportional to the eigenvector associated with the minimum eigenvalue of $\mathbf{S}\mathbf{S}^\top + \mathbf{W} - \mathbf{s}_m\mathbf{s}_m^\top$. This simple intuitively pleasing greedy procedure when applied sequentially by all users, results at each iteration in the reduction of a quantity called the total square correlation (TSC) – a measure of the average interference seen by all users [1–5]:

$$\mathrm{TSC} = \mathrm{Trace}[(\mathbf{S}\mathbf{S}^\top + \mathbf{W})^2] \tag{4}$$

Furthermore, empirically speaking, the fixed point reached by this iterative procedure invariably has the absolute minimum TSC attainable [5]. Other procedures have also been shown to reduce TSC [1,3], but here we consider only greedy procedures which reduce interference for each user if at all possible. A formal statement of the greedy algorithm used here is provided in Appendix A.

At any fixed point of the algorithm, each $\mathbf{s}_m$ must obviously be an eigenvector of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$. Equally obvious, the $\mathbf{s}_m$ associated with different eigenvalues must be orthogonal ([13], pp. 212). This orthogonality leads to the observation that if λ_I is the eigenvalue for $\mathbf{s}_k$ and λ_{II} the eigenvalue for $\mathbf{s}_\ell$, $k \neq \ell$, then we must also have $\lambda_I - p_k \leq \lambda_{II}$ since otherwise user k could reduce its interference by setting $\mathbf{s}_k = \sqrt{p_k}\mathbf{s}_\ell/|\mathbf{s}_\ell|$.

3 Convergence Via Class Warfare

It is easily seen that there may exist many such fixed points with differing TSC values. Thus, if we seek to show that greedy interference avoidance can absolutely minimize TSC, we must first provide some means of escape from local minima. Then, since interference avoidance monotonically reduces TSC toward fixed points, if we show that the possible number of fixed point TSC values is finite, eventual convergence to the absolute minimum TSC is unavoidable. To these ends we provide formal escape methods in which users "attack" other users or signal dimensions by deliberate interference, or artful encroachment on less crowded portions of the signal space.

It should be noted that since we assume all user signals are collected at common antennas and it is the reception point which feeds optimal codewords back to users, the notion of unilateral "attack" is more a useful analogy than an operational principle. That is, class warfare is more an analytic method to escape suboptimal minima than a distributed algorithmic method of guaranteeing minimum TSC "in the wild". Of course, as software radios become more powerful and sophisticated, a useful systems-level paradigm might have semi-autonomous software agents responsible for decoding each user's signal as opposed to the centralized multiuser architectures of today [14–19]. In such a case, where agents may or may not collaborate directly, a warfare analogy might be more accurate. However, since all previous numerical experiments indicate that interference avoidance algorithms seem to converge without the help of escape methods or which codeword is chosen for replacement at a given step, even here the point is somewhat moot, and we reiterate that the term class warfare is a conceptual crutch for an analytic method of TSC minimization.

3.1 Some Preliminaries

Assume with no loss of generality that the equilibrium codeword ensemble consists of three known sets $\{\mathbf{a}_k, \mathbf{b}_\ell, \mathbf{c}_j\}$ such that

$$\mathbf{S}\mathbf{S}^\top = \sum_{k=1}^{K_a} \mathbf{a}_k \mathbf{a}_k^\top + \sum_{\ell=1}^{K_b} \mathbf{b}_\ell \mathbf{b}_\ell^\top + \sum_{j=1}^{M-K_a-K_b} \mathbf{c}_j \mathbf{c}_j^\top \tag{5}$$

and $(\mathbf{W}+\mathbf{S}\mathbf{S}^\top)\mathbf{a}_k = \lambda_I \mathbf{a}_k$, $(\mathbf{W}+\mathbf{S}\mathbf{S}^\top)\mathbf{b}_\ell = \lambda_{II} \mathbf{b}_\ell$ and $(\mathbf{W}+\mathbf{S}\mathbf{S}^\top)\mathbf{c}_j = \lambda_j \mathbf{c}_j$ with $\lambda_I > \lambda_{II}$ and $\lambda_j \neq \lambda_I, \lambda_{II}$. By virtue of their different eigenvalues, codewords of different classes, $\{\mathbf{a}_k, \mathbf{b}_\ell, \mathbf{c}_j\}$, are mutually orthogonal. Finally, since it is possible that the codewords might not span $\Re^N$, we also admit the possibility of a set $\{\mathbf{d}_m\}$ of eigenvectors for $(\mathbf{W}+\mathbf{S}\mathbf{S}^\top)$ with cardinality K_d and associated eigenvalues λ_m which are themselves mutually orthogonal as well as orthogonal to the codeword sets $\{\mathbf{a}_k, \mathbf{b}_\ell, \mathbf{c}_j\}$.

For equal power codewords, each codeword from the set $\{\mathbf{a}_k\}$ suffers greater interference than each codeword from the set $\{\mathbf{b}_\ell\}$. For unequal power, we can only say that set $\{\mathbf{a}_k\}$ is in a more energetic (including the set energy) portion of the signal space than set $\{\mathbf{b}_\ell\}$. The assumption that the codeword constellation be a fixed point of a greedy interference avoidance algorithm requires $\lambda_I - \lambda_i \leq \min_k |\mathbf{a}_k|^2$ where λ_i is any other eigenvalue of $\mathbf{W}+\mathbf{S}\mathbf{S}^\top$.

We assume a basis set $\{\boldsymbol{\phi}_i\}$, $i = 1, 2, \cdots, n_1$, which spans $\{\mathbf{a}_k\}$ and is orthogonal to $\{\mathbf{b}_\ell\}$ and $\{\mathbf{c}_j\}$. Likewise we assume a basis set $\{\boldsymbol{\psi}_j\}$, $j = 1, 2, \cdots, n_2$, $(n_1 + n_2 \leq N)$ which spans $\{\mathbf{b}_\ell\}$ and is orthogonal to $\{\mathbf{a}_k\}$ and $\{\mathbf{c}_j\}$, and a basis set $\{\boldsymbol{\theta}_m\}$, $m = 1, 2, \cdots, N - n_1 - n_2$

which spans both the $\{\mathbf{c}_j\}$ and $\{\mathbf{d}_m\}$ and is therefore orthogonal to $\{\mathbf{a}_k\}$ and $\{\mathbf{b}_\ell\}$. We will later show (Theorem 2) that we can always choose these eigenvector sets as appropriate partitions of the eigenvector set of $\mathbf{W}$. But for this case where we assume all codewords are known, we leave $\{\boldsymbol{\phi}_i\}$, $\{\boldsymbol{\psi}_j\}$ and $\{\theta_m\}$ as convenient bases for their respective codeword sets. As a specific example, since all the elements of $\{\mathbf{a}_k\}$ have the same eigenvalue, λ_I (and similarly for $\{\mathbf{b}_\ell\}$), we can assign with no loss of generality, $\boldsymbol{\phi}_1 = \mathbf{a}_1/|\mathbf{a}_1|$ and $\boldsymbol{\psi}_1 = \mathbf{b}_1/|\mathbf{b}_1|$.

3.2 Warfare Algorithm Overview

An "attack" is the placement of energy from (aggrieved) users residing in a more crowded portion of the signal space into a dimension occupied by (offending) users in a more sparsely populated portion of the signal space. Such an attack adds interference to the offending users and will elicit an avoidance response which overall reduces TSC. The specific method of attack is a simple rotation of the aggrieved codeword (or set of codewords) toward the offending user codeword (or dimension).

We will analyze three different attack methods. The first is an attack by one user directly on another. The second is an attack by a group of users in the same class on a dimension in the offending class. In both these cases, TSC reduction is dependent upon a reaction (codeword movement) by the attacked users and will be effective only if certain requirements are met. The final method of warfare is used when the first two are ineffective, and is actually more of a "migration" than an attack since it does not depend on the reaction of the attacked dimension but only on the coordination of the attack by aggrieved users. If no warfare method will reduce TSC, then it will be shown that TSC is minimized.

Loosely stated, the warfare-augmented greedy interference avoidance algorithm is:

1. Apply greedy interference avoidance until convergence to a fixed point codeword ensemble.

2. If warfare can reduce TSC, apply it and then go to 1).

3. If warfare cannot reduce TSC, the ensemble has attained minimum TSC.

This algorithm is "loose" since we assume absolute convergence to fixed point ensembles in finite time for step 1) – which is not true for greedy interference avoidance in general. However, after developing the warfare method based on this fixed point assumption, we will then suggest a more formal "practical" algorithm which uses finite time approximate convergence for step 1.

3.3 Attacking the Devil You Know

Consider a scenario with a single aggrieved user $\mathbf{a}_1$ having eigenvalue λ_I and a known user $\mathbf{b}_1$ with eigenvalue $\lambda_{II} < \lambda_I$. We must have $\lambda_I - \lambda_j \leq |\mathbf{a}_1|^2 = p_1$ where λ_j is any eigenvalue of

$\mathbf{W} + \mathbf{S}\mathbf{S}^\top$ and therefore $\lambda_I - \lambda_{II} \le p_1$ since otherwise, simple greedy interference avoidance could be used to reduce TSC. We assume that $p_1 = 1$ with no loss of generality[2] so that $\boldsymbol{\phi}_1 = \mathbf{a}_1$. Likewise we assume $\mathbf{b}_1 = \beta\boldsymbol{\psi}_1$ where $|\mathbf{b}_1|^2 = \beta^2$. The aggrieved user $\mathbf{a}_1$ "attacks" user $\mathbf{b}_1$ by adjusting its codeword to

$$\mathbf{a}_1' = (\cos\omega)\boldsymbol{\phi}_1 + (\sin\omega)\boldsymbol{\psi}_1 \tag{6}$$

so that

$$\mathbf{a}_1'(\mathbf{a}_1')^\top = (\cos^2\omega)\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top + (\sin^2\omega)\boldsymbol{\psi}_1\boldsymbol{\psi}_1^\top + \frac{1}{2}(\sin 2\omega)(\boldsymbol{\psi}_1\boldsymbol{\phi}_1^\top + \boldsymbol{\phi}_1\boldsymbol{\psi}_1^\top) \tag{7}$$

for some non-zero ω.

User $\mathbf{b}_1$ will react to this challenge by replacing $\mathbf{b}_1$ with $\mathbf{b}_1'$, the new minimum eigenvalue eigenvector. That is,

$$(\mathbf{W} + \mathbf{S}\mathbf{S}^\top - \mathbf{a}_1\mathbf{a}_1^\top + \mathbf{a}_1'(\mathbf{a}_1')^\top - \mathbf{b}_1\mathbf{b}_1^\top)\mathbf{b}_1' = \lambda^*\mathbf{b}_1' \tag{8}$$

where λ^* is the minimum eigenvalue of

$$\mathbf{G} = \mathbf{W} + \mathbf{S}\mathbf{S}^\top - \mathbf{a}_1\mathbf{a}_1^\top + \mathbf{a}_1'(\mathbf{a}_1')^\top - \mathbf{b}_1\mathbf{b}_1^\top \tag{9}$$

We then note that

$$\mathbf{G}\boldsymbol{\phi}_i = \lambda_I\boldsymbol{\phi}_i \tag{10}$$

for $i = 2, 3, \cdots, n_1$

$$\mathbf{G}\boldsymbol{\psi}_j = \lambda_{II}\boldsymbol{\psi}_j \tag{11}$$

for $j = 2, 3, \cdots, n_2$, and

$$\mathbf{G}\boldsymbol{\theta}_m = \lambda_m\boldsymbol{\theta}_m \tag{12}$$

for $m = 1, 2, \cdots, N - n_1 - n_2$. We therefore have $N-2$ eigenvectors for $\mathbf{G}$. Since $\boldsymbol{\phi}_1$ and $\boldsymbol{\psi}_1$ are orthogonal to these, the remaining two (new) eigenvectors must be linear superpositions of $\boldsymbol{\phi}_1$ and $\boldsymbol{\psi}_1$, $\mathbf{x} = x_1\boldsymbol{\phi}_1 + x_2\boldsymbol{\psi}_1$. Depending upon ω, the best new codeword $\mathbf{b}_1'$ might be one of these two new eigenvectors, or could be another one entirely chosen from equations (10), (11) and (12). However, we will restrict our attention to $\mathbf{b}_1'$ of the form

$$\mathbf{b}_1' = \beta\left[(\sin\chi)\boldsymbol{\phi}_1 + (\cos\chi)\boldsymbol{\psi}_1\right] \tag{13}$$

and note that a potentially better choice might exist. That is, by showing suitable selection of χ in equation (13) can strictly reduce TSC, we will have also shown that a better response

[2] Note that we can always effectively set $|\mathbf{a}_1| = 1$ by normalizing $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ with $|\mathbf{a}_1|^2 = p_1$. Thus, any codeword can be used as $\mathbf{a}_1$ with $|\mathbf{a}_1| = 1$.

to attack can only decrease TSC even more.

With $\mathbf{b}_1'$ as in equation (13) we then have

$$\mathbf{b}_1'(\mathbf{b}_1')^\top = \beta^2 \left[(\sin^2\chi)\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top + (\cos^2\chi)\boldsymbol{\psi}_1\boldsymbol{\psi}_1^\top + \frac{1}{2}(\sin 2\chi)(\boldsymbol{\psi}_1\boldsymbol{\phi}_1^\top + \boldsymbol{\phi}_1\boldsymbol{\psi}_1^\top) \right] \tag{14}$$

Now we calculate the difference in TSC after the attack and response,

$$\Delta = \mathrm{Trace}[(\mathbf{S}\mathbf{S}^\top + \mathbf{W})^2] - \mathrm{Trace}[(\mathbf{G} + \mathbf{b}_1'(\mathbf{b}_1')^\top)^2] \tag{15}$$

Making the required substitutions and defining

$$\begin{aligned} \mathbf{Q}(\omega,\chi) &= -[\sin^2\omega - \beta^2\sin^2\chi]\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top + [\sin^2\omega - \beta^2\sin^2\chi]\boldsymbol{\psi}_1\boldsymbol{\psi}_1^\top \\ &+ \tfrac{1}{2}[(\sin 2\omega) + \beta^2(\sin 2\chi)](\boldsymbol{\psi}_1\boldsymbol{\phi}_1^\top + \boldsymbol{\phi}_1\boldsymbol{\psi}_1^\top) \end{aligned} \tag{16}$$

we obtain

$$\Delta(\omega,\chi) = -2\mathrm{Trace}[(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{Q}(\omega,\chi)] - \mathrm{Trace}[\mathbf{Q}(\omega,\chi)\mathbf{Q}(\omega,\chi)] \tag{17}$$

Performing the indicated substitution and remembering that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\phi} = \lambda_I\boldsymbol{\phi}$ and $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\psi} = \lambda_{II}\boldsymbol{\psi}$ yields

$$\tfrac{1}{2}\Delta(\omega,\chi) = \delta(\sin^2\omega - \beta^2\sin^2\chi) - [\sin^2\omega - \beta^2\sin^2\chi]^2 - \tfrac{1}{4}[\sin(2\omega) + \beta^2\sin(2\chi)]^2 \tag{18}$$

where $\delta = \lambda_I - \lambda_{II}$. We note that $0 < \delta \le 1$ since we assumed $p_1 = |a_1|^2 = 1$.

We need to determine whether $\Delta > 0$ for some choices of ω. However, we first note that for any given ω, the response $\mathbf{b}_1'(\chi)$ to the attack by $\mathbf{a}_1'(\omega)$ will maximize Δ [5]. Therefore, we first seek the maximizing χ for equation (18). Algebraic manipulations and trigonometric identities applied to equation (18) yield,[3]

$$\begin{aligned} \tfrac{1}{2}\Delta(\omega,\chi) &= (\delta - (1-\beta^2))\sin^2\omega - \tfrac{1}{2}\beta^2(\delta + \beta^2) \\ &+ \tfrac{1}{2}\beta^2\left[(\delta - (1-\beta^2))(\cos 2\chi) + \cos(2\chi + 2\omega)\right] \end{aligned} \tag{19}$$

We then note that

$$\max_x |A\cos 2x + B\cos(2x + 2\omega)| = \sqrt{(A + B\cos 2\omega)^2 + B^2\sin^2 2\omega} \tag{20}$$

[3]The calculation is mundane but involved. The interested reader should simply compare equation (18) and equation (19) with a symbolic mathematics program such as Maple©.

Therefore letting $\Gamma = \delta + \beta^2$ we have for $A = \Gamma - 1$ and $B = 1$,

$$\begin{aligned} \max_{\chi} \tfrac{1}{2}\Delta(\omega,\chi) &= (\Gamma - 1)\sin^2\omega \\ &- \tfrac{1}{2}(\Gamma - \delta)\Gamma\left[1 - \sqrt{1 - 4\tfrac{\Gamma-1}{\Gamma^2}\sin^2\omega}\right] \end{aligned} \tag{21}$$

We then note that $\Delta(\omega,\chi) > 0$ relies on having $\Gamma - 1 > 0$ which in turn implies $\Gamma - \delta > 0$ since $\delta \leq 1$. Thus, the positivity of equation (21) rests on whether $\exists\omega$ such that

$$(\Gamma - 1)\sin^2\omega > \frac{1}{2}(\Gamma - \delta)\Gamma\left[1 - \sqrt{1 - 4\frac{\Gamma - 1}{\Gamma^2}\sin^2\omega}\right] \tag{22}$$

and we note that the term inside the radical is non-negative so that the right hand side of equation (22) is non-negative. Rearranging we have

$$1 - 2\frac{\Gamma - 1}{\Gamma - \delta}\frac{1}{\Gamma}\sin^2\omega < \sqrt{1 - 4\frac{\Gamma - 1}{\Gamma^2}\sin^2\omega} \tag{23}$$

and we again note that since the term inside the radical is non-negative ($4(\Gamma - 1)\sin^2\omega/\Gamma^2 \leq 1$) that the left hand side of equation (23) is non-negative as well. So, squaring both sides and simplifying leads to

$$\frac{\Gamma - 1}{\Gamma - \delta}\sin^2\omega < \delta \tag{24}$$

which is true for some small enough ω, so long as $\delta > 0$.

Thus,

$$\delta > 0 \tag{25}$$

and

$$\Gamma - 1 = \delta - (1 - \beta^2) > 0 \tag{26}$$

suffice for the existence of an ω such that $\Delta(\omega,\chi) > 0$. We summarize and generalize the result as a theorem:

Theorem 1 *Suppose a single user with codeword power α^2 and eigenvalue λ_I attacks another user with codeword power β^2 and eigenvalue $\lambda_{II} < \lambda_I$ by adjusting its codeword according to equation (6). Further suppose that the attacked user responds by adjusting its codeword to obtain the maximum possible SINR. Then there exists an attack parameter ω that will be successful in* strictly *reducing TSC if $\delta = \lambda_I - \lambda_{II} > 0$ and $\delta > \alpha^2 - \beta^2$.*

This result is intuitively pleasing since it makes little sense to engage in warfare if at best what you gain ($\lambda_{II} - \beta^2$) is no better than what you already have ($\lambda_I - \alpha^2$). Reworking these two expressions leads directly to $\beta^2 > \alpha^2 - \delta$. Another looser but more memorable condition can be had by noting that since $\delta > 0$, any $\beta^2 \geq \alpha^2$ allows $\Delta(\omega,\chi) > 0$. Since $\beta^2 = \alpha^2$ implies equal power aggrieved and offending users, $\beta^2 > \alpha^2$ implies a weaker user attacking a stronger user. Thus, we may say that TSC can always be strictly reduced whenever a less powerful user with associated eigenvalue λ_I attacks an equal or more powerful user with associated eigenvalue $\lambda_{II} < \lambda_I$.

3.4 Attacking the Devil You Don't Know

Now suppose that user codewords in one class are unknown to those in another so that a directed attack against a particular user is infeasible. However, if the background noise covariance $\mathbf{W}$ is known to all users, then an aggrieved user could measure the signal energy along dimensions (eigenvectors) of the noise covariance and attack specific offending dimensions.

We will find the following two theorems useful. The first theorem pertains to the relationship between equilibrium codeword sets and the eigenspace of the noise covariance matrix $\mathbf{W}$.

Theorem 2 *At equilibrium, codewords with different eigenvalues reside in mutually orthogonal spaces. These spaces are spanned by mutually orthogonal partitions of an eigenvector set for* $\mathbf{W}$*. Furthermore,* $\mathbf{W}$ *shares a complete set of eigenvectors with* $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ *at equilibrium.*

Proof: Theorem 2 Consider a set of codewords $\{\mathbf{a}_i\}$ with cardinality m_1, dimensionality n_1 and associated unique eigenvalue λ_a. Let us then define

$$\mathbf{S}\mathbf{S}^\top = \mathbf{A}\mathbf{A}^\top + \mathbf{Z}\mathbf{Z}^\top \tag{27}$$

where

$$\mathbf{A}\mathbf{A}^\top = \sum_{i=1}^{m_1} \mathbf{a}_i\mathbf{a}_i^\top \tag{28}$$

and $\mathbf{Z}$ is a matrix containing the codewords with eigenvalues other than λ_a. We then have

$$(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{a}_i = (\mathbf{A}\mathbf{A}^\top + \mathbf{Z}\mathbf{Z}^\top + \mathbf{W})\mathbf{a}_i = (\mathbf{A}\mathbf{A}^\top + \mathbf{W})\mathbf{a}_i = \lambda_a\mathbf{a}_i \tag{29}$$

Note that any linear combination of the $\{\mathbf{a}_i\}$ is also an eigenvector of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ since all the $\mathbf{a}_i$ share the same eigenvalue λ_a.

Since $\mathbf{A}\mathbf{A}^\top + \mathbf{W}$ has a full set of eigenvectors and the $\{\mathbf{a}_i\}$ are spanned by exactly n_1 of them, there must be an additional $N - n_1$ eigenvectors $\boldsymbol{\phi}_m$ which lie in the orthogonal complement of $\{\mathbf{a}_i\}$ and therefore satisfy

$$(\mathbf{A}\mathbf{A}^\top + \mathbf{W})\boldsymbol{\phi}_m = \mathbf{W}\boldsymbol{\phi}_m = \lambda_m\boldsymbol{\phi}_m \tag{30}$$

so that they are eigenvectors of $\mathbf{W}$ as well. Since there are exactly $N - n_1$ of these eigenvectors, the remaining n_1 eigenvectors $\boldsymbol{\phi}_i$, $i = 1, 2, \cdots, n_1$ of $\mathbf{W}$ must form the basis for the set $\{\mathbf{a}_i\}$.

Extending this result, we see that for all sets of codewords with the same eigenvalue, there must be an associated set of eigenvectors of $\mathbf{W}$ which exactly spans the space of these codewords. Therefore, equilibrium codewords which share different eigenvalues reside in different partitions of the noise covariance signal space.

Finally, as mentioned previously, each of the eigenvectors of $\mathbf{W}$ which span the codeword space can be expressed as a suitable linear combination of codewords

$$\sum_{i=1}^{m_1} \mu_{ij} \mathbf{a}_i = \boldsymbol{\phi}_j \tag{31}$$

$j = 1, 2, \cdots, n_1$. Therefore, since $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{s}_m = \lambda_a \mathbf{s}_m$ we can also find $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\phi}_j = \lambda_a \boldsymbol{\phi}_j$ with all the $\boldsymbol{\phi}_j$ eigenvectors of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ with eigenvalue λ_a. This completes the proof. •

It is worth mentioning that from this point onward we will express codewords as linear superpositions of the noise covariance eigenvector partition in which they reside as opposed to the arbitrary basis set used in section 3.3. The next theorem establishes a relationship between the mutual correlations of codewords in a given class and their projections onto a noise covariance eigenvector contained in that class.

Theorem 3 *Let $\boldsymbol{\phi}_i$, $i = 1, 2, \cdots, N$ be a common complete set of eigenvectors for $\mathbf{W}$ and $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ (Theorem 2) with $\mathbf{W}\boldsymbol{\phi}_i = \sigma_i \boldsymbol{\phi}_i$. Assume $\{\boldsymbol{\phi}_i\}$, $i = 1, 2, \cdots, L$ exactly spans the set of codewords $\mathbf{s}_k$, $k = 1, 2, \cdots, K$ which share the same eigenvalue λ, $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{s}_k = \lambda \mathbf{s}_k$.*

If we define $\alpha_{k\ell} = \mathbf{s}_k^\top \boldsymbol{\phi}_\ell$, then for $\rho_{ij} = \mathbf{s}_i^\top \mathbf{s}_j$ we must have

$$\sum_{j=1}^{K} \alpha_{i\ell}\alpha_{j\ell} \left(\rho_{ij} - \alpha_{i\ell}\alpha_{j\ell} \right) = 0 \tag{32}$$

Proof: Theorem 3 By Theorem 2 and Mercer's theorem [20] we have

$$\mathbf{S}\mathbf{S}^\top + \mathbf{W} = \lambda \sum_{i=1}^{L} \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top + \sum_{i=L+1}^{N} \lambda_i \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top \tag{33}$$

where the $\{\boldsymbol{\phi}_i\}$ is the complete eigenvector set for $\mathbf{W}$. Since the $\{\mathbf{s}_k\}$, $k = 1, 2, \cdots, K$ are exactly spanned by the $\boldsymbol{\phi}_i$, $i = 1, 2, \cdots, L$ we then have

$$\lambda \sum_{i=1}^{L} \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top = \sum_{k=1}^{K} \mathbf{s}_k \mathbf{s}_k^\top + \sum_{i=1}^{L} \sigma_i \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top \tag{34}$$

which implies

$$\sum_{k=1}^{K} \mathbf{s}_k \mathbf{s}_k^\top = \sum_{i=1}^{L} (\lambda - \sigma_i) \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top \tag{35}$$

We define $E_i = \lambda - \sigma_i$, the codeword energy along eigenvector $\boldsymbol{\phi}_i$ and note that for $\alpha_{ki} = \boldsymbol{\phi}_i^\top \mathbf{s}_k$ we have $E_i = \sum_{k=1}^K \alpha_{ki}^2$.

We then have

$$\sum_{j=1}^{K} \alpha_{i\ell}\alpha_{j\ell}\left(\rho_{ij} - \alpha_{i\ell}\alpha_{j\ell}\right) = \sum_{j=1}^{K} \alpha_{i\ell}\rho_{ij}\alpha_{j\ell} - \alpha_{i\ell}^2 E_\ell \tag{36}$$

and note that

$$\sum_{j=1}^{K} \alpha_{i\ell}\rho_{ij}\alpha_{j\ell} = \sum_{j=1}^{K} \boldsymbol{\phi}_\ell^\top(\mathbf{s}_i\mathbf{s}_i^\top)(\mathbf{s}_j\mathbf{s}_j^\top)\boldsymbol{\phi}_\ell = \boldsymbol{\phi}_\ell^\top(\mathbf{s}_i\mathbf{s}_i^\top)\left(\sum_{i=1}^{L} E_i\boldsymbol{\phi}_i\boldsymbol{\phi}_i^\top\right)\boldsymbol{\phi}_\ell^\top = \alpha_{i\ell}^2 E_\ell \tag{37}$$

which completes the proof. •

Now, let $\mathcal{A}$ be the aggrieved set of codewords with elements $\mathbf{a}_i$ such that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{a}_i = \lambda_I \mathbf{a}_i$. We can write

$$\mathbf{a}_i = \alpha_i\boldsymbol{\phi}_1 + \mathbf{x}_i \tag{38}$$

where $\boldsymbol{\phi}_1$ is one of the eigenvectors of $\mathbf{W}$ which spans the aggrieved codeword set. We note that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\phi}_1 = \lambda_I\boldsymbol{\phi}_1$, that $\mathbf{W}\boldsymbol{\phi}_1 = \sigma_1\boldsymbol{\phi}_1$ and that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{x}_i = \lambda_I\mathbf{x}_i$.

Let $\mathcal{S}$ be the (non-empty) set of offending codewords with eigenvalue $\lambda_{II} < \lambda_I$. And by Theorem 2 let the noise covariance dimension $\boldsymbol{\phi}_N$ be contained in the span of $\mathcal{S}$ with $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\phi}_N = \lambda_{II}\boldsymbol{\phi}_N$. For the codewords $\{\mathbf{s}_i\}$, $i \in \mathcal{S}$ in this set we write

$$\mathbf{s}_i = \beta_i\boldsymbol{\phi}_N + \mathbf{u}_i \tag{39}$$

with $\beta_i = \boldsymbol{\phi}_N^\top\mathbf{s}_i$. As for $\mathbf{a}$ we note that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\phi}_N = \lambda_{II}\boldsymbol{\phi}_N$, that $\mathbf{W}\boldsymbol{\phi}_N = \sigma_N\boldsymbol{\phi}_N$ and that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{u}_i = \lambda_{II}\mathbf{u}_i$.

To attack, we rotate each $\mathbf{a}_i$ into $\boldsymbol{\phi}_N$ as

$$\mathbf{a}_i' = \alpha_i(\cos\omega\boldsymbol{\phi}_1 + \sin\omega\boldsymbol{\phi}_N) + \mathbf{x}_i \tag{40}$$

and then script the evasion response by the offending set as

$$\mathbf{s}_i' = \beta_i(\cos\chi\boldsymbol{\phi}_N + \sin\chi\boldsymbol{\phi}_1) + \mathbf{u}_i \tag{41}$$

To calculate the change in TSC we write

$$\begin{aligned} \mathbf{a}_i'(\mathbf{a}_i')^\top &= \mathbf{a}_i\mathbf{a}_i^\top + \alpha_i^2\left[\sin^2\omega\boldsymbol{\phi}_N\boldsymbol{\phi}_N^\top - \sin^2\omega\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top + \cos\omega\sin\omega(\boldsymbol{\phi}_1\boldsymbol{\phi}_N^\top + \boldsymbol{\phi}_N\boldsymbol{\phi}_1^\top)\right] \\ &+ \alpha_i(\cos\omega - 1)(\boldsymbol{\phi}_1\mathbf{x}_i^\top + \mathbf{x}_i\boldsymbol{\phi}_1^\top) + \alpha_i\sin\omega(\boldsymbol{\phi}_N\mathbf{x}_i^\top + \mathbf{x}_i\boldsymbol{\phi}_N^\top) \end{aligned} \tag{42}$$

and

$$\begin{aligned} \mathbf{s}_i'(\mathbf{s}_i')^\top &= \mathbf{s}_i\mathbf{s}_i^\top + \beta_i^2\left[\sin^2\chi\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top - \sin^2\chi\boldsymbol{\phi}_N\boldsymbol{\phi}_N^\top + \cos\chi\sin\chi(\boldsymbol{\phi}_1\boldsymbol{\phi}_N^\top + \boldsymbol{\phi}_N\boldsymbol{\phi}_1^\top)\right] \\ &+ \beta_i(\cos\chi - 1)(\boldsymbol{\phi}_N\mathbf{u}_i^\top + \mathbf{u}_i\boldsymbol{\phi}_N^\top) + \beta_i\sin\chi(\boldsymbol{\phi}_1\mathbf{u}_i^\top + \mathbf{u}_i\boldsymbol{\phi}_1^\top) \end{aligned} \tag{43}$$

Therefore,

$$\begin{aligned}
\mathbf{S}'(\mathbf{S}')^\top &= \mathbf{S}\mathbf{S}^\top + \sum_{i\in\mathcal{A}} \alpha_i^2 \left[\sin^2\omega\boldsymbol{\phi}_N\boldsymbol{\phi}_N^\top - \sin^2\omega\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top + \cos\omega\sin\omega(\boldsymbol{\phi}_1\boldsymbol{\phi}_N^\top + \boldsymbol{\phi}_N\boldsymbol{\phi}_1^\top)\right] \\
&+ \sum_{i\in\mathcal{A}} \alpha_i(\cos\omega - 1)(\boldsymbol{\phi}_1\mathbf{x}_i^\top + \mathbf{x}_i\boldsymbol{\phi}_1^\top) + \sum_{i\in\mathcal{A}} \alpha_i\sin\omega(\boldsymbol{\phi}_N\mathbf{x}_i^\top + \mathbf{x}_i\boldsymbol{\phi}_N^\top) \\
&+ \sum_{i\in\mathcal{S}} \beta_i^2 \left[\sin^2\chi\boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top - \sin^2\chi\boldsymbol{\phi}_N\boldsymbol{\phi}_N^\top + \cos\chi\sin\chi(\boldsymbol{\phi}_1\boldsymbol{\phi}_N^\top + \boldsymbol{\phi}_N\boldsymbol{\phi}_1^\top)\right] \\
&+ \sum_{i\in\mathcal{S}} \beta_i(\cos\chi - 1)(\boldsymbol{\phi}_N\mathbf{u}_i^\top + \mathbf{u}_i\boldsymbol{\phi}_N^\top) + \sum_{i\in\mathcal{S}} \beta_i\sin\chi(\boldsymbol{\phi}_1\mathbf{u}_i^\top + \mathbf{u}_i\boldsymbol{\phi}_1^\top)
\end{aligned} \tag{44}$$

We then define the total aggrieved codeword energy in $\boldsymbol{\phi}_1$ as $\alpha^2 = \sum_{i\in\mathcal{A}}\alpha_i^2$, the total offending codeword energy in dimension $\boldsymbol{\phi}_N$ as $\beta^2 = \sum_{i\in\mathcal{S}}\beta_i^2$ and form $Q(\omega,\chi)$ as in section 3.3;

$$\begin{aligned}
\mathbf{Q}(\omega,\chi) &= (\alpha^2\sin^2\omega - \beta^2\sin^2\chi)\left(\boldsymbol{\phi}_N\boldsymbol{\phi}_N^\top - \boldsymbol{\phi}_1\boldsymbol{\phi}_1^\top\right) \\
&+ (\alpha^2\cos\omega\sin\omega + \beta^2\cos\chi\sin\chi)(\boldsymbol{\phi}_1\boldsymbol{\phi}_N^\top + \boldsymbol{\phi}_N\boldsymbol{\phi}_1^\top) \\
&+ \sum_{i\in\mathcal{A}} \alpha_i(\cos\omega - 1)(\boldsymbol{\phi}_1\mathbf{x}_i^\top + \mathbf{x}_i\boldsymbol{\phi}_1^\top) + \sum_{i\in\mathcal{A}} \alpha_i\sin\omega(\boldsymbol{\phi}_N\mathbf{x}_i^\top + \mathbf{x}_i\boldsymbol{\phi}_N^\top) \\
&+ \sum_{i\in\mathcal{S}} \beta_i(\cos\chi - 1)(\boldsymbol{\phi}_N\mathbf{u}_i^\top + \mathbf{u}_i\boldsymbol{\phi}_N^\top) + \sum_{i\in\mathcal{S}} \beta_i\sin\chi(\boldsymbol{\phi}_1\mathbf{u}_i^\top + \mathbf{u}_i\boldsymbol{\phi}_1^\top)
\end{aligned} \tag{45}$$

We then define $\Delta(\omega,\chi)$ as in equation (17) and reduce it to

$$\begin{aligned}
\tfrac{1}{2}\Delta(\omega,\chi) &= \delta(\alpha^2\sin^2\omega - \beta^2\sin^2\chi) - (\alpha^2\sin^2\omega - \beta^2\sin^2\chi)^2 - \tfrac{1}{4}(\beta^2\sin 2\chi + \alpha^2\sin 2\omega)^2 \\
&- ((\cos\omega - 1)^2 + \sin^2\omega)\sum_{i,j\in\mathcal{A}} \alpha_i\alpha_j(\kappa_{ij} - \alpha_i\alpha_j) \\
&- ((\cos\chi - 1)^2 + \sin^2\chi)\sum_{i,j\in\mathcal{S}} \beta_i\beta_j(\rho_{ij} - \beta_i\beta_j)
\end{aligned} \tag{46}$$

where $\kappa_{ij} = \mathbf{a}_i^\top\mathbf{a}_j$, $i,j\in\mathcal{A}$, $\rho_{ij} = \mathbf{s}_i^\top\mathbf{s}_j$, $i,j\in\mathcal{S}$ and $\delta = \lambda_I - \lambda_{II}$ as before.

By Theorem 3 the terms in $(\cos\omega - 1)^2 + \sin^2\omega$ and $(\cos\chi - 1)^2 + \sin^2\chi$ must be identically zero so we have

$$\begin{aligned}
\tfrac{1}{2}\Delta(\omega,\chi) &= \delta(\alpha^2\sin^2\omega - \beta^2\sin^2\chi) - (\alpha^2\sin^2\omega - \beta^2\sin^2\chi)^2 - \tfrac{1}{4}(\beta^2\sin 2\chi + \alpha^2\sin 2\omega)^2 \\
&= \alpha^4\left[\tfrac{\delta}{\alpha^2}(\sin^2\omega - \tfrac{\beta^2}{\alpha^2}\sin^2\chi) - (\sin^2\omega - \tfrac{\beta^2}{\alpha^2}\sin^2\chi)^2 - \tfrac{1}{4}(\sin 2\omega + \tfrac{\beta^2}{\alpha^2}\sin 2\chi)^2\right]
\end{aligned} \tag{47}$$

Equation (47) is identical to equation (18) within a factor of α^4 if we define $\tilde{\delta} = \delta/\alpha^2$ and $\tilde{\beta} = \beta/\alpha^2$. Therefore the result is identical to that in section 3.3: TSC will be strictly reduced by dimensional attack so long as $\tilde{\delta} > 0$ and $\tilde{\delta} > 1 - \tilde{\beta}^2$ which reduces to $\delta > 0$ and $\delta > \alpha^2 - \beta^2$. The only difference is that we have launched a coordinated attack and have "scripted" an ensemble response (specifically, a rotation by all offending users from the attacked dimension $\boldsymbol{\phi}_N$ toward the attacking signal dimension $\boldsymbol{\phi}_1$) as opposed to assuming individual greedy responses by each codeword.

3.5 Leveling the Playing Field Through Cooperation

Now suppose that application of greedy interference avoidance has resulted in a fixed point such that there exists an eigenvector $\boldsymbol{\psi}$ of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ with eigenvalue $\lambda_{II} < \lambda_I$ where λ_I is the eigenvalue of one of the user codewords. Further, suppose that no sufficiently powerful codeword (or none at all) resides in the offending dimension $\boldsymbol{\psi}$ so that attack and response warfare will be ineffective.[4] We will show that in such a case a *coordinated migration* into the offending dimension will strictly reduce TSC unless a simple set of conditions is violated.

As with dimensional attack assume there exist codewords $\{\mathbf{s}_i\}$, $i = 1, 2, \cdots, K$ such that $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{s}_i = \lambda_I \mathbf{s}_i$ and we assume the codewords occupy $n < N$ dimensions. Each codeword "attacks" $\boldsymbol{\psi}$ by forming

$$\mathbf{s}'_i = \cos\omega_i \mathbf{s}_i + \sin\omega_i \sqrt{p_i}\boldsymbol{\psi} \tag{48}$$

so that

$$\begin{aligned} \mathbf{s}'_i(\mathbf{s}'_i)^\top &= \mathbf{s}_i\mathbf{s}_i^\top - \sin^2\omega_i \mathbf{s}_i\mathbf{s}_i^\top + p_i \sin^2\omega_i \boldsymbol{\psi}\boldsymbol{\psi}^\top \\ &+ \tfrac{1}{2}\sqrt{p_i}\sin 2\omega_i(\mathbf{s}_i\boldsymbol{\psi}^\top + \boldsymbol{\psi}\mathbf{s}_i^\top) \end{aligned} \tag{49}$$

Now we form the new $\mathbf{S}'(\mathbf{S}')^\top + \mathbf{W}$ as

$$\mathbf{S}'(\mathbf{S}')^\top + \mathbf{W} = \mathbf{S}\mathbf{S}^\top + \mathbf{W} + \mathbf{Q} \tag{50}$$

where $\mathbf{Q} = \sum_{i=1}^{K} [\mathbf{s}'_i(\mathbf{s}'_i)^\top - \mathbf{s}_i\mathbf{s}_i^\top]$. Using equation (49) we have

$$\begin{aligned} \mathbf{Q} &= -\sum_{i=1}^{K} \sin^2\omega_i(\mathbf{s}_i\mathbf{s}_i^\top - p_i\boldsymbol{\psi}\boldsymbol{\psi}^\top) \\ &+ \sum_{i=1}^{K} \frac{1}{2}\sqrt{p_i}\sin 2\omega_i(\mathbf{s}_i\boldsymbol{\psi}^\top + \boldsymbol{\psi}\mathbf{s}_i^\top) \end{aligned} \tag{51}$$

As found previously (see equation (17)), the difference in TSC before and after migration is $\Delta = -2\mathrm{Trace}[(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{Q}] - \mathrm{Trace}[\mathbf{Q}^2]$. So defining $\delta = \lambda_I - \lambda_{II}$ as before and $\rho_{ij} = \mathbf{s}_i^\top \mathbf{s}_j$ we have

$$\tfrac{1}{2}\Delta(\Omega) = \delta\sum_{i=1}^{K} p_i \sin^2\omega_i - \sum_{i,j=1}^{K} \sin^2\omega_i \sin^2\omega_j(\rho_{ij}^2 + p_i p_j) - \frac{1}{4}\sum_{i,j=1}^{K}\sqrt{p_i}\sqrt{p_j}\rho_{ij}\sin 2\omega_i \sin 2\omega_j \tag{52}$$

[4]See Appendix B for an example.

Now, for small enough ω_i, terms on the order of $\sin^4(\cdot)$ become insignificant and we have

$$\tfrac{1}{2}\Delta(\Omega) \approx \delta\sum_{i=1}^{K} p_i\omega_i^2 - \sum_{i,j=1}^{K} \sqrt{p_i}\omega_i\rho_{ij}\sqrt{p_j}\omega_j \tag{53}$$

If we define

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_K \end{bmatrix} \tag{54}$$

where $v_i = \sqrt{p_i}\omega_i$ we see that

$$\frac{1}{2}\Delta(\Omega) \approx \delta\sum_{i=1}^{K} p_i\omega_i^2 - \mathbf{v}^\top\mathbf{Z}\mathbf{v} \tag{55}$$

where $\mathbf{Z}$ is a matrix whose entries are $z_{ij} = \rho_{ij}$.

We then note that the $K \times K$ matrix

$$\mathbf{Z} = \begin{bmatrix} -\mathbf{s}_1- \\ -\mathbf{s}_2- \\ -\vdots- \\ -\mathbf{s}_K- \end{bmatrix} \begin{bmatrix} | & | & | & | \\ \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_K \\ | & | & | & | \end{bmatrix} \tag{56}$$

will be singular if the $\mathbf{s}_i$ are linear combinations of only $n < K$ eigenvectors $\boldsymbol{\phi}_j$. In such a case there exists a non-zero vector $\mathbf{v}$ for which $\mathbf{v}^\top\mathbf{Z}\mathbf{v} = 0$. And since $\delta\sum_{i=1}^{K} p_i\omega_i^2 > 0$ for any of the ω_i nonzero, we can always find a set of suitably small $\{\omega_i\}$ for which $\Delta(\Omega) > 0$.

We summarize the result as a theorem:

Theorem 4 *Let $\{\mathbf{s}_i\}$ be the set of all codewords which share a common eigenvalue λ_I. Let $\boldsymbol{\psi}$ be an eigenvector of $\mathbf{W}$ orthogonal to all the $\{\mathbf{s}_i\}$ and for which $(\mathbf{SS}^\top + \mathbf{W})\boldsymbol{\psi} = \lambda_{II}\boldsymbol{\psi}$ where $\lambda_I - \lambda_{II} = \delta > 0$. TSC can always be strictly reduced by coordinated attack on dimension $\boldsymbol{\psi}$ if the number of codewords is at least one more than the number of dimensions they span.*

4 Properties of Eigenvalue Sets for $\mathbf{SS}^\top + \mathbf{W}$

Since TSC is the trace of $(\mathbf{SS}^\top + \mathbf{W})^2$ we see that TSC depends only on the eigenvalues of $\mathbf{SS}^\top + \mathbf{W}$, $\{\lambda_i\}$. That is, Trace$[(\mathbf{SS}^\top + \mathbf{W})^2] = \sum_i \lambda_i^2$. To more easily distinguish one eigenvalue set from another, we assume that each set is ordered from largest to smallest and define such ordered sets as λ-constellations. We will then derive order bounds which must be obeyed by all possible λ-constellations and show that when the bounds are met with equality, TSC is minimized and that almost incidentally, the sum capacity (see Appendix C) of the codeword set is maximized.

4.1 Bounds On Partial Sums of Ordered Eigenvalues

We define the eigenvalue and eigenvector set of the noise covariance matrix $\mathbf{W}$ as $\{\sigma_i\}$ and $\{\boldsymbol{\phi}_i\}$ respectively, $i = 1, 2, \cdots, N$. With no loss of generality, we assume that the $\{\sigma_i\}$ and the signal energies $\{p_k\}$, $k = 1, 2, \cdots, M$ are ordered as $\sigma_i \geq \sigma_{i+1}$ and $p_k \geq p_{k+1}$. For convenience, we also define $P = \sum_{k=1}^{M} p_k$ and $U = \sum_{i=1}^{N} \sigma_i$. We will assume at least as many codewords as dimensions ($M \geq N$) since if not, optimality dictates that the codewords be contained in the space spanned by the M least noisy dimensions – those with energies $\sigma_N, \sigma_{N-1}, \cdots, \sigma_{N-M+1}$. Our approach is to find lower bounds to sums of ordered eigenvalues of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$, $\lambda_i \geq \lambda_{i+1}$.

Now, for any matrix $\mathbf{Q}$ with eigenvalues $\{\mu_i\}$ ordered from largest to smallest we have ([13], pp. 253),

$$\max_{\mathbf{x}_i^\top \mathbf{x}_j = \delta_{ij}} \sum_{i=1}^{k} \mathbf{x}_i^\top \mathbf{Q} \mathbf{x}_i = \sum_{i=1}^{k} \mu_i \tag{57}$$

Consider then

$$\mathbf{S}\mathbf{S}^\top + \mathbf{W} = \sum_{i=1}^{M} \mathbf{s}_i \mathbf{s}_i^\top + \sum_{j=1}^{N} \sigma_j \boldsymbol{\phi}_j \boldsymbol{\phi}_j^\top \tag{58}$$

and that

$$\max_{|\mathbf{x}|=1} \mathbf{x}^\top (\mathbf{S}\mathbf{S}^\top + \mathbf{W}) \mathbf{x} = \lambda_1 \tag{59}$$

It follows immediately that $\lambda_1 \geq \sigma_1$ and $\lambda_1 \geq |\mathbf{s}_1|^2 + \sigma_N = p_1 + \sigma_N$. Equally obvious is that $\lambda_1 \geq (P + U)/N$ [5,6,8]. Slightly less obvious is that

$$\lambda_1 \geq \frac{1}{k} \sum_{i=1}^{k} (p_i + \sigma_{N-i+1}) \tag{60}$$

for $k = 1, 2, \cdots, N-1$. This may be shown by noting that $\sum_{i=1}^{k} \lambda_i \geq \sum_{i=1}^{k} (p_i + \sigma_{N-i+1})$ and that the minimum maximum λ_1 must then be at least this quantity divided by k.

Therefore, we must have in total

$$\lambda_1 \geq \max \left[\sigma_1, \max_{0<k<N} \frac{1}{k} \sum_{i=1}^{k} (p_i + \sigma_{N-i+1}), \frac{P+U}{N} \right] \tag{61}$$

and we will show that equation (61) may be used recursively to bound the quantity $\sum_{i=1}^{n} \lambda_i$. In what follows we denote the eigenvector of $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})$ associated with λ_i as $\boldsymbol{\psi}_i$.

Since we seek the smallest possible λ_1 we must consider in turn each possibility expressed in equation (61). So to start, suppose σ_1 is the minmax λ_1 bound. To meet the bound, no codeword can have energy along $\boldsymbol{\phi}_1$ since otherwise we must have $\lambda_1 > \sigma_1$. So, if $\boldsymbol{\psi}_1 = \alpha\mathbf{x} + \sqrt{1-\alpha^2}\boldsymbol{\phi}_1$ where $\mathbf{x} \perp \boldsymbol{\phi}_1$ is the eigenvector with largest eigenvalue σ_1, then we must also have

$$(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{x} = \sigma_1\mathbf{x} \tag{62}$$

which implies that $\exists\mathbf{u} \perp \boldsymbol{\psi}_1$ also with eigenvalue σ_1. The overall implication is that we must have $\lambda_2 \geq \sigma_1$ as well.

Now consider that with $\boldsymbol{\psi}_1 = \boldsymbol{\phi}_1$ we have

$$\lambda_2 \geq \max\left[\sigma_2, \max_{0<k<N-1} \frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U-\sigma_1}{N-1}\right] \tag{63}$$

and we note that since we have

$$\max\left[\sigma_1, \max_{0<k<N} \frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U}{N}\right] = \sigma_1 \tag{64}$$

we must also have

$$\max\left[\sigma_2, \max_{0<k<N-1} \frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U-\sigma_1}{N-1}\right] \leq \sigma_1 \tag{65}$$

a bound smaller than that if $\boldsymbol{\psi}_1 \neq \boldsymbol{\phi}_1$. Thus, equation (63) is the lowest lower bound on the minmax λ_2 for $\lambda_1 = \sigma_1$ and can also be seen as a recursive application of equation (61) after $\boldsymbol{\phi}_1$ has been expurgated from the ensemble and the dimensionality of the problem decreased by 1.

Continuing sequentially with the terms in equation (61), now suppose that instead of σ_1, we have $p_1 + \sigma_N$ as the minmax λ_1 bound. To meet this bound we must have $\boldsymbol{\psi}_1 = \boldsymbol{\phi}_\ell$ and $\mathbf{s}_1 = \sqrt{p_1}\boldsymbol{\phi}_\ell$ for any ℓ such that $\sigma_\ell = \sigma_N$. If not then λ_1 must be strictly greater than $p_1 + \sigma_N$. We then have

$$\lambda_2 \geq \max\left[\sigma_1, \max_{1<k<N} \frac{1}{k-1}\sum_{i=2}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U-p_1-\sigma_N}{N-1}\right] \tag{66}$$

As before, we note that equation (66) is a recursive application of equation (61) with $\mathbf{s}_1$ and $\boldsymbol{\phi}_N$ expurgated.

Finally we consider the last term in equation (61) and suppose that now $\frac{p_1+p_2+\sigma_N+\sigma_{N-1}}{2}$ is the minmax λ_1 bound. Since we must have $\lambda_1 + \lambda_2 \geq p_1 + p_2 + \sigma_N + \sigma_{N-1}$, if the minmax bound for λ_1 is met with equality, we must have $\lambda_2 = \lambda_1$ and we are left to bound λ_3 as

$$\lambda_3 \geq \max\left[\sigma_1, \max_{2<k<N} \frac{1}{k-2}\sum_{i=3}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U-p_1-p_2-\sigma_N-\sigma_{N-1}}{N-2}\right] \tag{67}$$

which is an application of equation (61) with codewords $\mathbf{s}_1$ and $\mathbf{s}_2$ and dimensions $\boldsymbol{\phi}_N$ and $\boldsymbol{\phi}_{N-1}$ expurgated.

In general, if

$$\max\left[\sigma_1, \max_{0<k<N} \frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U}{N}\right] = \frac{1}{n}\sum_{i=1}^{n}(p_i + \sigma_{N-i+1}) \tag{68}$$

and $n < N$ we will have

$$\sum_{j=1}^{\ell}\lambda_j \geq \frac{\ell}{n}\sum_{i=1}^{n}(p_i + \sigma_{N-i+1}) \tag{69}$$

for $\ell \leq n$ and

$$\lambda_{n+1} \geq \max\left[\sigma_1, \max_{n<k<N} \frac{1}{k-n}\sum_{i=n+1}^{k}(p_i + \sigma_{N-i+1}), \frac{P+U-\sum_{i=1}^{n}(p_i+\sigma_{N-i+1})}{N-n}\right] \tag{70}$$

And once again we note that equation (70) is simply equation (61) applied after $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_n$ and $\boldsymbol{\phi}_N, \boldsymbol{\phi}_{N-1}, \cdots, \boldsymbol{\phi}_{N-n+1}$ have been expurgated. Therefore, we see that equation (61) can be used as a recursive kernel to generate the least lower bound for partial sums of ordered eigenvalues λ_i.

As a specific example consider, $\underline{\sigma} = (26, 6, 4, 4)$ and $\underline{p} = (4, 3, 2, 1)$. Application of equation (61) has $\lambda_1 \geq \sigma_1 = 26$. Expurgation of the associated dimension leaves us with $\underline{\sigma}' = (6, 4, 4)$ and $\lambda_1 + \lambda_2 \geq 34 = 26 + p_1 + \sigma_3'$. After expurgating p_1 and σ_3' we are left with $\underline{\sigma}'' = (6, 4)$ and $\underline{p}'' = (3, 2, 1)$. Applying the bound again we have $\lambda_3 + \lambda_2 + \lambda_1 \geq 42 = 34 + (p_1 + p_3'' + p_3'' + \sigma_1'' + \sigma_2'')/2$. As another example[5], consider $\underline{\sigma} = (1.1, 1, 0)$ and $\underline{p} = (1, 1, 1/5)$. In this case $\lambda_1 \geq 3/2$ and $\lambda_1 + \lambda_2 \geq 3$.

In general, is it clear that when the bound for $\sum_{i=1}^{n}\lambda_i$ is plotted against n, the result must be a concave segmented arc above the line $n(P+U)/N$ as illustrated in FIGURE 1 for the first example.

[5]This example was provided by P. Anigstein at UCB.

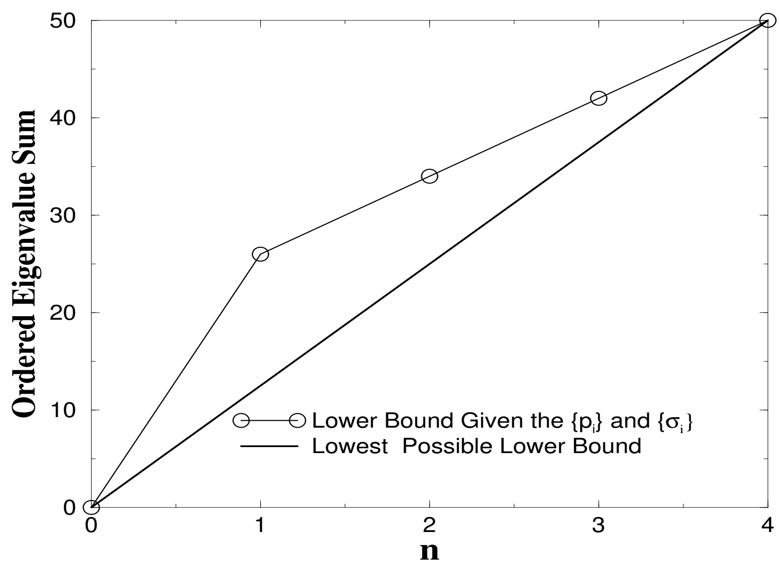

Figure 1: Illustration of $\sum_{i=1}^{n} \lambda_i$ bounds in n for text example.

4.2 The Lower Bound Is the Optimal λ-Constellation

The definition of the constraint set in the previous section specifies ordered eigenvalues and partial sums of eigenvalues. To this end let us define

$$F_X(k) = \sum_{j=1}^{k} x_j \tag{71}$$

$k = 1, 2, \cdots, N$ and where the $x_j \geq 0$ are ordered $x_j \geq x_{j+1}$. The reader will note that $F_X(k)/F_X(N)$ is exactly the cumulative distribution function of an ordered probability distribution $x_j/F_X(N)$, $i = 1, 2, \cdots, N$ – a variation on stochastic ordering [21–23]. To avoid confusion we note that the term "stochastic" is irrelevant in our context since we do not consider random variables, but make mention of it because the mathematics of ordered distributions is useful for our purposes. [6] The necessary result is now stated as a theorem.

Theorem 5 *Suppose two non-negative non-increasing sequences $\{x_j\}$ and $\{r_j\}$ have $F_X(k) \geq F_R(k)$ $k = 1, 2, \cdots, N-1$ and $F_X(N) = F_R(N) = 1$. For any convex function $g()$ we must have $G(\mathbf{x}) = \sum_{j=1}^{N} g(x_j) \geq \sum_{j=1}^{N} g(r_j) = G(\mathbf{r})$. If $g()$ is concave, then $G(\mathbf{x}) \leq G(\mathbf{r})$.*

A proof of this (known) result is provided in Appendix D for convenience.

[6]It should also be noted that these results, when applied to ordered distributions of eigenvalues as we do here, also go under the heading of majorization [11]. However, here we follow the ordered distribution approach since it is self-contained, compact and possibly more familiar than majorization to many readers.

We then note that for any two ordered eigenvalue sets $\{\lambda_j^{(1)}\}$ and $\{\lambda_j^{(2)}\}$ with $\sum_{j=1}^N \lambda_j^{(i)} = (P+U)$ and $\sum_{j=1}^N \lambda_j^{(1)} \geq \sum_{j=1}^N \lambda_j^{(2)}$, we may normalize by $P+U$ to obtain $\mathbf{x}$ and $\mathbf{r}$ respectively as defined in Theorem 5. We further note that

$$\mathrm{TSC}(\mathbf{x}) = (P+U)^2 \sum_{j=1}^N x_j^2 \tag{72}$$

and (see Appendix C)

$$C_{\mathbf{s}}(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^N \log x_i + \frac{1}{2}\sum_{i=1}^N \log(P+U) - \frac{1}{2}\sum_{i=1}^N \log \sigma_i \tag{73}$$

Since $g(x_j) = x_j^2$ is convex, $\mathrm{TSC}(\mathbf{x}) \geq \mathrm{TSC}(\mathbf{r})$. Likewise since $g(x_j) = \log x_j$ is concave we have $C_{\mathbf{s}}(\mathbf{x}) \leq C_{\mathbf{s}}(\mathbf{r})$.

Therefore, since the bounds generated by recursive application of equation (61) describe the ordered set of possible eigenvalues with smallest "stochastic" order, TSC is minimized and $C_{\mathbf{s}}$ is maximized by eigenvalue sets which attain the bound. For emphasis, we state this result as a theorem.

Theorem 6 *For a set of codewords $\{\mathbf{s}_i\}$ with ordered powers $p_i \geq p_{i+1}$, $i = 1, 2, \cdots, M$ in a space of dimension N with noise covariance $\mathbf{W}$, sum capacity is maximized and TSC minimized when the partial sums of non-increasingly ordered eigenvalues, $\{\lambda_i\}$, of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ attain the lower bounds generated by recursive application of equation (61).*

The eigenvalue structure of such sets is illustrated in FIGURE 2 We prove their existence in the next section.

5 Warfare Minimizes TSC (Maximizes Sum Capacity)

Through a progression of theorems we examine fixed points for warfare-augmented greedy interference avoidance and show equivalence between the implied eigenvalues of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ and the eigenvalues given by the lower bounds derived from equation (61). We start with some preliminary theorems which will be used extensively throughout. We follow these with theorems which parallel the three terms in equation (61) and then a theorem which states that the only true equilibrium for greedy interference avoidance with class warfare is an optimum codeword ensemble. We close the section by showing that warfare-augmented greedy interference avoidance eventually achieves an optimum ensemble after traversing at most a finite number of suboptimal minima.

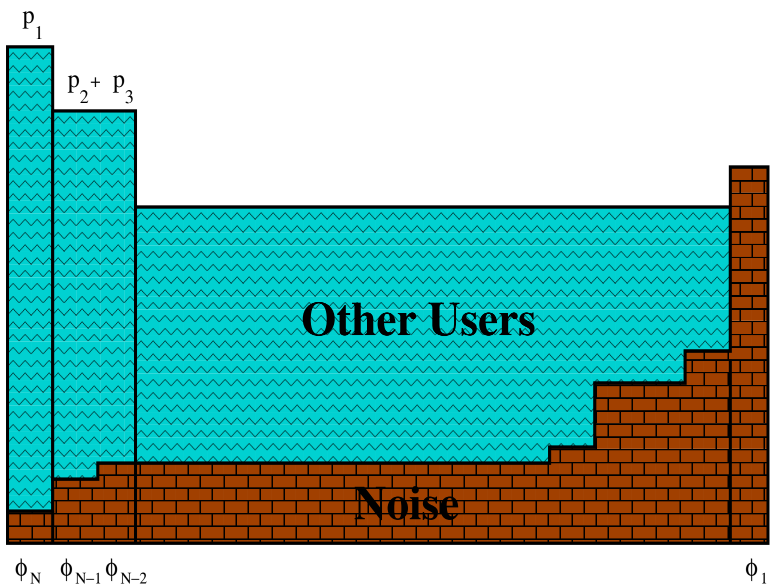

Figure 2: Illustration of optimal fixed point for a codeword ensemble with different powers. Vertical bars denote noise covariance dimension partitions.

5.1 Preliminaries

Consider FIGURE 3 which depicts the power distribution for a general suboptimal codeword ensemble. We would like to drive this distribution toward the optimal distribution given in FIGURE 2. To this end, consider the following set of assertions which establish the codeword occupancy of the various noise covariance dimensions.

As a corollary of Theorem 1 we have

Corollary 1 *Suppose single user warfare cannot reduce TSC for a given codeword set* $\{\mathbf{s}_i\}$ *with associated eigenvalue* λ_I*. Then, the codeword energies* $p_i = |\mathbf{s}_i|^2$ *for all elements of this set must at least as large as that for all other codewords outside the set with eigenvalues* $\lambda_{II} < \lambda_I$*.*

Theorem 7 *Let* $\{\mathbf{s}_k\}$ *be the set of all codewords with eigenvalue* λ_I*. Suppose there exists a signal dimension* $\boldsymbol{\psi}$*, an eigenvector of* $\mathbf{W}$ *with* $(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\boldsymbol{\psi} = \lambda_{II}\boldsymbol{\psi}$*,* $\mathbf{W}\boldsymbol{\psi} = \gamma\boldsymbol{\psi}$ *and* $\lambda_{II} < \lambda_I$ *but TSC cannot be reduced by dimensional attack. Then the partition of noise covariance eigenvectors* $\boldsymbol{\phi}_i$ *which spans* $\{\mathbf{s}_k\}$ *must all have eigenvalues* $\{\sigma_i\}$ *no larger than* γ*, the noise covariance eigenvalue associated with* $\boldsymbol{\psi}$*.*

Proof: Theorem 7 Let E_ϕ be the signal energy along dimension $\boldsymbol{\phi}$. We then have $\lambda_I = E_\phi + \sigma_\phi$. Likewise define the signal energy E_ψ as the signal energy along $\boldsymbol{\psi}$ so that $\lambda_{II} = E_\psi + \gamma$. For dimensional attack to be ineffective, we must have $\delta \leq E_\phi - E_\psi$ which implies

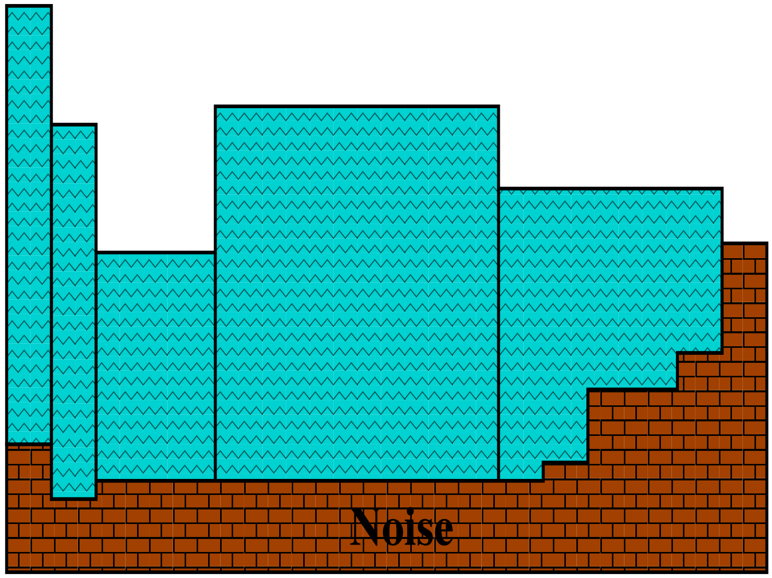

Figure 3: Illustration of suboptimal fixed point for a codeword ensemble with different powers.

$E_\phi + \sigma_\phi - (E_\psi + \gamma) \leq E_\phi - E_\psi$ which in turn implies $\sigma_\phi - \gamma \leq 0$ for any $\boldsymbol{\phi}$ in the spanning set of $\{\mathbf{s}_k\}$ thus proving the theorem. •

The following theorem is implied directly by Corollary 1 and Theorem 7 and is stated without proof.

Theorem 8 *Let* $\mathbf{s}_i$ *and* $\mathbf{s}_j$ *be codewords with eigenvalues* λ_I *and* λ_{II} *respectively. Let each reside in spaces which contain largest noise covariances* γ_i *and* γ_j *respectively. If* $|s_i| \geq |s_j|$*, then* $\lambda_I \geq \lambda_{II}$ *(Corollary 1) and* $\gamma_i \leq \gamma_j$ *(Theorem 7).*

In words, at equilibrium a higher power codeword has an eigenvalue at least as large as that of a lower power codeword. In addition, a higher power codeword resides in a partition of the noise covariance space with variances at least as small the space in which a lower power codewords resides.

It is useful to note that Theorem 8 implies that higher power codewords command the best noise space partitions and that lower power codewords are relegated to higher power noise space partitions. Although this is not in keeping with the egalitarian spirit of class warfare, it is (unfortunately) in keeping with the physical reality that might often makes right!

5.2 The Bounds of Equation (61)

The next three theorems address the three terms inside the max function of equation (61) and will be used directly to show that the warfare procedure produces eigenvalue sets which meet the lower bound specified by recursive application of equation (61).

Theorem 9 Minmax $[\sigma_1]$**:** *If* σ_1 *is the minmax eigenvalue bound in equation (61) then at equilibrium, all codewords must be orthogonal to* $\boldsymbol{\phi}_1$*.*

Proof: Theorem 9 Suppose at some equilibrium $\exists s_1$ with nonzero projection onto the eigenvector $\boldsymbol{\phi}_1$ associated with σ_1. This codeword must then have eigenvalue $\lambda_I > \sigma_1$. Now we note that since σ_1 is the minmax eigenvalue bound we must also have by equation (61) that $\sigma_1 \geq (P+U)/N$ where P is the total codeword energy and U is the total noise energy. Since the eigenvalues of $\mathbf{SS}^\top + \mathbf{W}$ must sum to $P+U$, if $\exists \lambda_I > \sigma_1 \geq (P+U)/N$ then $\exists \lambda_{II}$ with $\lambda_{II} < \sigma_1 < \lambda_I$ which in turn implies that $\exists \boldsymbol{\phi}_i$ with $(\mathbf{SS}^\top + \mathbf{W})\boldsymbol{\phi}_i = \lambda_{II}\boldsymbol{\phi}_i$ and $\mathbf{W}\boldsymbol{\phi}_i = \sigma_i \boldsymbol{\phi}_i$ with $\sigma_i < \sigma_1$ since $\lambda_{II} \geq \sigma_i$. This contradicts Theorem 7. Thus, no codeword energy can exist along dimension $\boldsymbol{\phi}_1$ at equilibrium if σ_1 is the minmax bound for λ_1. •

Theorem 10 $\left[\mathbf{Minmax}\frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1})\right]$**:** *Define*

$$\gamma = \frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1}) \tag{74}$$

where k *is the largest integer* $< N$ *such that*

$$\frac{1}{k}\sum_{i=1}^{k}(p_i + \sigma_{N-i+1}) \geq \frac{1}{n}\sum_{i=1}^{n}(p_i + \sigma_{N-i+1}) \tag{75}$$

$n = 1, 2, \cdots, N-1$.

If the bound of equation (61) is equal to γ*, then at an equilibrium where warfare cannot reduce TSC, the largest* k *codewords,* $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k$*, will reside in the space spanned by* $\boldsymbol{\phi}_N, \boldsymbol{\phi}_{N-1}, \cdots, \boldsymbol{\phi}_{N-k+1}$ *with the smallest* k *noise covariances,* $\sigma_N, \sigma_{N-1}, \cdots, \sigma_{N-k+1}$*. These codewords will share the same eigenvalue* $\lambda_I = \gamma$ *and will therefore meet the lower bound for the first* k *eigenvalues specified by equation (61).*

Proof: Theorem 10 Let $\mathcal{S}$ be the set of *all* codewords (cardinality L) with the largest eigenvalue λ_I. Assume this set spans an n-dimensional space. By Theorem 8 these L codewords must be the largest L codewords $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_L$ and will span the lowest n noise dimensions.

Now if there exists a codeword $\mathbf{s}_{L+1}$, it must have eigenvalue $\lambda_{II} < \lambda_I$ by the definition of $\mathcal{S}$ as containing all codewords with eigenvalue λ_I. Therefore we must have $L = n$ since otherwise, group migration (Theorem 4) could be applied to reduce TSC. So we must have

$$\lambda_I = \frac{1}{n}\sum_{i=1}^{n}(p_i + \sigma_{N-i+1}) \tag{76}$$

and the first n eigenvalues of $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ are λ_I. However, by equation (61) we must also have $\lambda_1 \geq \gamma$ which is a direct contradiction unless

$$\frac{1}{n}\sum_{i=1}^{n}(p_i + \sigma_{N-i+1}) = \gamma \tag{77}$$

We then note that if equation (77) is true for some $n < k$ then by the definition of γ we must also have

$$\frac{1}{k-n}\sum_{i=n+1}^{k}(p_i + \sigma_{N-i+1}) = \gamma \tag{78}$$

as well. But this is impossible since by construction we require the largest eigenvalue outside $\mathcal{S}$ to be $\lambda_{n+1} < \lambda_I = \gamma$ and recursive application of equation (61) to the remaining codewords $\mathbf{s}_L, \cdots, \mathbf{s}_M$ in noise dimensions $\boldsymbol{\phi}_{N-n}, \boldsymbol{\phi}_{N-n-1}, \cdots, \boldsymbol{\phi}_1$ along with satisfaction of equation (78) requires that $\lambda_{n+1} \geq \gamma$. Thus, n cannot be less than k.

Now suppose $n > k$. Then $\lambda_1 = \lambda_I \leq \gamma$ and $\lambda_1 \geq \gamma$ is a direct contradiction owing to the "largest k" stipulation in the definition of γ. That is, there exists no $n > k$ such that $\lambda_I \geq \gamma$. Thus, we must have $n = k$.

So in summary, if γ as defined in equation (74) is the minmax bound on λ_1, then at any equilibrium where warfare cannot be used to reduce TSC, the k highest power codewords must reside in the k lowest power noise dimensions and share the same eigenvalue γ, unique and the largest over all codewords. $\bullet$

Theorem 11 $\left[\textbf{Minmax}\frac{P+U}{N}\right]$: *If the minmax bound for λ_1 is $(P+U)/N$, then all codewords must share the same eigenvalue $\lambda_I = (P+U)/N$ at an equilibrium where warfare cannot reduce TSC.*

Proof: Theorem 11 As in Theorem 10, let $\mathcal{S}$ be the set of L highest power codewords residing in the n lowest power noise dimensions. By Theorem 4, unless $L = n$, any difference in eigenvalues can be exploited to reduce TSC. Therefore, unless all the eigenvalues are already equal to $(P+U)/N$, we must have

$$\lambda_I = \frac{1}{n}\sum_{i=1}^{n}(p_i + \sigma_{N-i+1}) \tag{79}$$

However, since $(P+U)/N$ is the minmax value of λ_1, the structure of equation (61) requires $\lambda_I \geq (P+U)/N$, a condition which can only be met if $\lambda_I = (P+U)/N$. By the assumed monotonicity of eigenvalues for $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$, if $\lambda_1 = (P+U)/N$, then all the eigenvalues must be $(P+U)/N$ thus completing the proof. $\bullet$

5.3 The λ-Constellation Bound Is the Stopping Condition

Theorem 9, Theorem 10 and Theorem 11 along with the recursive nature of the eigenvalue bound in equation (61) lead directly to the following theorem.

Theorem 12 *The only stable codeword ensembles for warfare-augmented greedy interference avoidance are those which meet with equality the eigenvalue bound generated by recursive application of equation (61).*

Proof: Theorem 12 Suppose the conditions of Theorem 9 are satisfied. Then at equilibrium the bound on λ_1 specified by equation (61) is met with equality since no codeword energy resides in $\boldsymbol{\phi}_1$. All the codewords and remaining dimensions are orthogonal to $\boldsymbol{\phi}_1$.

Alternately, suppose that the conditions of Theorem 10 are satisfied for some $k < N$. Then once again, the bound specified in equation (61) is met with equality and the remaining codewords $\mathbf{s}_{k+1}, \mathbf{s}_{k+2}, \cdots, \mathbf{s}_M$ must reside in noise dimensions $\boldsymbol{\phi}_{N-k}, \boldsymbol{\phi}_{N-k-1}, \cdots, \boldsymbol{\phi}_1$, orthogonal to codewords $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k$ which reside in dimensions $\boldsymbol{\phi}_N, \boldsymbol{\phi}_{N-1}, \cdots, \boldsymbol{\phi}_{N-k+1}$. Similarly, if the conditions of Theorem 11 are satisfied, then the bound specified in equation (61) is met with equality and all the eigenvalues of $\mathbf{SS}^\top + \mathbf{W}$ are identical.

In each of these three cases, the codeword and dimension partitioning implied by the fixed point exactly parallels the partitioning implied by equation (61). Thus, equation (61), Theorem 9, Theorem 10 and Theorem 11 can be applied at each step to an independent smaller problem until no codewords or dimensions remain.

Since at each recursive step the bound of equation (61) is met with equality, stable equilibrium ensembles for warfare-assisted greedy interference avoidance are only those ensembles which minimize TSC and maximize sum capacity. •

So in summary, recursive expurgation of largest eigenvalue signal dimensions and codewords in the equilibrium set is identical to the codeword/dimension expurgation and recursive application of equation (61) used to establish the lower bound. Thus, the equilibrium set meets the bound at every step. Therefore, the equilibrium point for greedy interference avoidance, augmented by warfare techniques, is a codeword ensemble which maximizes sum capacity and minimizes TSC. Thus, a fixed point ensemble which is stable under class warfare in general has a codeword power distribution as depicted in FIGURE 2.

5.4 A Level Playing Field Attained, Eventually

So far, we have provided algorithms which strictly reduce TSC at fixed points not satisfying a set of stopping conditions, and we have shown that only optimum codeword ensembles can satisfy the stopping conditions. However, showing reduction can always be forced does not guarantee that an optimal value is achieved since there could be an infinite number of fixed point values which taken in some monotonic sequence might approach any number of values. Therefore, we now show that the number of possible greedy interference algorithm equilibrium TSC values is finite and that the warfare procedure must therefore eventually attain some unique TSC specified by the stopping conditions.

This possibly subtle assertion bears repetition. Although the number of possible algorithm fixed point ensembles is potentially infinite, we can show that the number of TSC fixed point *values* is finite. Since escape methods strictly reduce TSC at suboptimal equilibria, only a finite number of TSC values need be traversed and convergence to the optimal is assured.

For a set of codewords $\{\mathbf{a}_i\}$ which spans n_1 dimensions and shares the same eigenvalue λ_a at equilibrium, by Theorem 2 the set must span some n_1-dimensional partition of the eigenspace of $\mathbf{W}$. Let this partition be described by $\{\boldsymbol{\psi}_i\}$, $i = 1, 2, \cdots, n_1$ where the $\boldsymbol{\psi}_i$ are eigenvectors of $\mathbf{W}$ (i.e., $\mathbf{W}\boldsymbol{\psi}_i = \sigma_i \boldsymbol{\psi}_i$). If there are m_1 codewords in the set we must have

$$\lambda_a = \frac{1}{n_1}\left[\sum_{i=1}^{m_1} |\mathbf{a}_i|^2 + \sum_{i=1}^{n_1} |\mathbf{W}\boldsymbol{\psi}_i|^2\right] \tag{80}$$

The maximum number of possible choices (without regard for feasibility) for assemblages of n_1 noise covariance eigenvectors and m_1 codewords is $\binom{N}{n_1}\binom{M}{m_1}$. Therefore, with the set $(\{n_i\}, \{m_i\})$ denoting all possible (but again, potentially infeasible) partitions of M codewords into N dimensions, there are at most a (possibly large) but finite number of fixed point TSC values

$$\prod_{(\{n_i\},\{m_i\})} \binom{N}{n_i}\binom{M}{m_i} \tag{81}$$

This result along with those for warfare techniques and stopping rules admits the following summarizing theorem:

Theorem 13 *Greedy interference avoidance coupled with class warfare allows escape from any fixed point which does not meet the stopping conditions for optimality. Since there are a finite number of fixed points, the composite algorithm is therefore guaranteed to attain the stopping points specified in section 5.3 and therefore achieve the minimum possible TSC and maximum sum capacity.*

5.5 Practical Considerations: a quantitative sketch

In the previous development we have assumed that the warfare procedure is started from a fixed point ensemble. In a theoretical sense, this is perfectly acceptable since we have shown that greedy interference avoidance does converge to such fixed points asymptotically.

However, in a practical setting, the greedy procedure would be stopped at some point and this would be close to, but perhaps not exactly a fixed point. Thus, for completeness, it is useful to consider the case where the codeword ensemble is "almost" at a fixed point and how warfare would be applied for such ensembles.

So, consider that greedy interference avoidance can be applied until codewords are within some small ν of being eigenvectors of the signal plus noise covariance matrix (see Appendix A). More formally, for suitably large ℓ and any codeword $\mathbf{s}_i(\ell)$ we have

$$|\mathbf{R}(\ell)\mathbf{s}_i(\ell) - \lambda_i(\ell)\mathbf{s}_i(\ell)|^2 \leq \nu^2 \tag{82}$$

Now suppose single user warfare is applied. We would have exactly equation (17)

$$\Delta(\omega,\chi) = -2\mathrm{Trace}[(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{Q}(\omega,\chi)] - \mathrm{Trace}[\mathbf{Q}(\omega,\chi)\mathbf{Q}(\omega,\chi)] \tag{83}$$

but then instead of equation (18) we would have

$$\begin{aligned} \tfrac{1}{2}\Delta(\omega,\chi) &= \delta(\sin^2\omega - \beta^2\sin^2\chi) - [\sin^2\omega - \beta^2\sin^2\chi]^2 - \tfrac{1}{4}[\sin(2\omega) + \beta^2\sin(2\chi)]^2 \\ +\mathbf{O}(\nu) & \end{aligned} \tag{84}$$

where the term in $\mathbf{O}(\nu)$ comes from the term $-2\mathrm{Trace}[(\mathbf{S}\mathbf{S}^\top + \mathbf{W})\mathbf{Q}(\omega,\chi)]$ and the imperfect eigenvectors which comprise the matrix $\mathbf{Q}$. This basic form applies to the other types of warfare as well.

A "practical" warfare method would therefore apply greedy interference avoidance until an approach to a fixed point was detected via an initial ν criterion, and then evaluate the potential decrease in TSC after warfare. If the potential decrease were much larger in magnitude than the $\mathbf{O}(\nu)$ term, then warfare would be applied. If not, then greedy interference would continue to decrease $|\nu|$ until such time as warfare would be guaranteed effective. The algorithm would then stop when TSC was within some tolerance of optimal.

Thus, we close with the following more formal statement of warfare-assisted greedy interference avoidance:

1. Apply greedy interference avoidance until convergence to within some tolerance ν of a fixed point codeword ensemble. If the fixed point is suboptimal, choose ν so that $|\mathbf{O}(\nu)|$ is much less than the potential benefit from warfare.

2. If warfare can reduce TSC, apply it and then go to 1).

3. If warfare cannot reduce TSC, the ensemble has attained nearly minimum TSC, and greedy interference avoidance may be further applied to reduce TSC until satisfied.

6 Summary and Conclusion

We have shown that in synchronous CDMA systems, greedy adaptation of codewords to achieve maximum SINR (interference avoidance), augmented by a method dubbed *class warfare* guarantees convergence of codeword ensembles to optimal sets which minimize total square correlation (TSC) and maximize sum capacity. In passing, the equivalence between sum capacity maximization and TSC minimization was shown. Application of the warfare procedure constitutes the first complete analytic proof that interference avoidance algorithms produce optimal codeword ensembles.

It must be noted, however, that in *all* numerical experiments with both greedy interference avoidance [2,4,5] and MMSE interference avoidance [1,3], when starting from randomly chosen initial codewords, essentially optimal sets were obtained within approximately three cycles of codeword adaptation and *never* required intervention with warfare-like methods. That is, direct application of interference avoidance principles was enough to ensure empirical convergence. Why this convergence to optimal codeword ensembles is so robust for simple interference avoidance methods remains an open question. We note that implicit in each warfare method described is a number of (gradient-like) directions for escape from local minima. Perhaps careful characterization of the dimensionality of potential escape trajectories relative those trajectories which tend to trap an ensemble in a local minimum might prove useful. Specifically, if approach trajectories are of zero measure, then small stochastic perturbations would be effective in avoiding local minima as was used in recent work [10] by Pablo Anigstein at UCB.

A The Greedy Interference Avoidance Algorithm

An informal statement of a greedy interference avoidance algorithm would be:

1. A minimum eigenvalue eigenvector, $\boldsymbol{\phi}$ of $\mathbf{R}_i = \mathbf{S}\mathbf{S}^\top + \mathbf{W} - \mathbf{s}_i\mathbf{s}_i^\top$ is determined.

2. $\mathbf{s}_i$ is replaced by $\boldsymbol{\phi}$ **iff**
$$\boldsymbol{\phi}^\top \mathbf{R}_i \boldsymbol{\phi} < \mathbf{s}_i^\top \mathbf{R}_i \mathbf{s}_i \tag{85}$$
which strictly reduces TSC. Otherwise, $\mathbf{s}_i$ is unchanged.

3. The procedure is repeated sequentially for every user i.

The algorithm is informal since no stopping criterion is supplied, even though each iteration cannot increase TSC and TSC bounded from below imply theoretical convergence to some TSC value. Practically, the algorithm is run until a cycle of iterations (all codewords) does not decrease TSC by some threshold amount.

However, TSC convergence does not necessarily imply codeword ensemble convergence. And since warfare methods presume the attainability of ensembles where every codeword is an eigenvector of the noise plus signal covariance matrix, provable convergence to such sets is a necessity. This is provided in the following theorem for a slightly modified version of greedy interference avoidance where at each iteration, the codeword chosen for replacement is that whose replacement will minimize TSC. We call this variant *greedy+* interference avoidance.

Theorem 14 *With iterative application of greedy+ interference avoidance, codewords must converge to ensembles where each codeword is an eigenvector of the signal plus noise covariance matrix* $\mathbf{R} = \mathbf{R}_i + \mathbf{s}_i\mathbf{s}_i^\top$.

Proof: Theorem 14 First we define an *iteration* ℓ as a greedy interference avoidance step where a codeword $\mathbf{s}_{i_\ell}(\ell)$ is replaced. We assume that codeword $\mathbf{s}_{i_\ell}(\ell)$ is the codeword which when replaced will produce the greatest reduction in TSC. We then have

$$\delta_{i_\ell}(\ell) = \mathbf{s}_{i_\ell}^\top(\ell)\mathbf{R}_{i_\ell}(\ell)\mathbf{s}_{i_\ell}(\ell) - \mathbf{x}_{i_\ell}(\ell)^\top\mathbf{R}_{i_\ell}(\ell)\mathbf{x}_{i_\ell}(\ell) \geq 0 \tag{86}$$

as the difference between the TSC value at iteration ℓ and $\ell+1$ where $\mathbf{x}_{i_\ell}(\ell)$ is the minimum eigenvalue eigenvector of $\mathbf{R}_{i_\ell}(\ell)$ [5]. We note that

$$\delta_{i_\ell}(\ell) = \max_i \delta_i(\ell) \tag{87}$$

and is therefore the maximum possible $\delta_i(\ell)$ at iteration ℓ.

Since greedy interference avoidance converges in TSC we have

$$\lim_{\ell\to\infty} \delta_{i_\ell}(\ell) = 0 \tag{88}$$

Now consider that the difference in TSC values before and after the *potential* replacement of *any* codeword $\mathbf{s}_i$ at iteration ℓ can be written as

$$\delta_{i_\ell}(\ell) \geq \delta_i(\ell) = \mathbf{s}_i^\top(\ell)\mathbf{R}_i(\ell)\mathbf{s}_i(\ell) - \mathbf{x}_i(\ell)^\top\mathbf{R}_i(\ell)\mathbf{x}_i(\ell) \geq 0 \tag{89}$$

We define the eigenvalues of $\mathbf{R}_i(\ell)$ as $\{\lambda_{ij}(\ell)\}$, $j = 1, 2, \cdots, N$ and assume that they are ordered from largest to smallest. If we further define the corresponding eigevectors as $\boldsymbol{\phi}_{ij}(\ell)$, $j = 1, 2, \cdots, N$ we can rewrite $\mathbf{s}_i(\ell)$ as

$$\mathbf{s}_i(\ell) = \sum_{j=1}^{N} \alpha_{ij}(\ell)\boldsymbol{\phi}_{ij}(\ell) \tag{90}$$

where we assume

$$\sum_{j=1}^{N} \alpha_{ij}^2(\ell) = |\mathbf{x}_i(\ell)|^2 = p_i \tag{91}$$

This leads to

$$\delta_i(\ell) = \sum_{j=1}^{N} \alpha_{ij}^2(\ell)(\lambda_{ij}(\ell) - \lambda_{iN}(\ell)) \tag{92}$$

Since all terms in the sum are non-negative we must have

$$\delta_i(\ell) \geq \alpha_{ij}^2(\ell)(\lambda_{ij}(\ell) - \lambda_{iN}(\ell)) \tag{93}$$

for $j = 1, 2, \cdots, N$. Now suppose via equation (93) we define $\epsilon_{ij}(\ell) \leq \delta_i(\ell)$ as

$$\epsilon_{ij}(\ell) = \alpha_{ij}^2(\ell)(\lambda_{ij}(\ell) - \lambda_{iN}(\ell)) \tag{94}$$

Dividing by nonzero $\alpha_{ij}(\ell)$ results in

$$\frac{\epsilon_{ij}(\ell)}{\alpha_{ij}(\ell)} = \alpha_{ij}(\ell)\lambda_{ij}(\ell) - \alpha_{ij}(\ell)\lambda_{iN}(\ell) \tag{95}$$

To see how closely each $\mathbf{s}_i(\ell)$ approximates an eigenvector of $\mathbf{R}(\ell) = \mathbf{R}_i(\ell) + \mathbf{s}_i(\ell)\mathbf{s}_i^\top(\ell)$ the signal plus noise covariance matrix at iteration ℓ, we form the product

$$\mathbf{R}(\ell)\mathbf{s}_i(\ell) \;=\; \sum_{j\in\mathbf{J}_i(\ell)} \lambda_{ij}(\ell)\alpha_{ij}(\ell)\boldsymbol{\phi}_{ij}(\ell) + p_i \sum_{j\in\mathbf{J}_i(\ell)} \alpha_{ij}(\ell)\boldsymbol{\phi}_{ij}(\ell) \tag{96}$$

where $\mathbf{J}_i(\ell)$ is the set of all j such that $\alpha_{ij}(\ell) \neq 0$. Using equation (95) in equation (96) yields

$$\mathbf{R}(\ell)\mathbf{s}_i(\ell) = \sum_{j\in\mathbf{J}_i(\ell)} \left(\frac{\epsilon_{ij}(\ell)}{\alpha_{ij}(\ell)} + \lambda_{iN}(\ell)\alpha_{ij}(\ell)\right)\boldsymbol{\phi}_{ij}(\ell) + p_i \sum_{j\in\mathbf{J}_i(\ell)} \alpha_{ij}(\ell)\boldsymbol{\phi}_{ij}(\ell) \tag{97}$$

Regrouping we have

$$\mathbf{R}(\ell)\mathbf{s}_i(\ell) = (\lambda_{iN}(\ell) + p_i)\mathbf{s}_i(\ell) + \sum_{j\in\mathbf{J}_i(\ell)} \frac{\epsilon_{ij}(\ell)}{\alpha_{ij}(\ell)}\boldsymbol{\phi}_{ij}(\ell) \tag{98}$$

However, since $\lim_{\ell\to\infty} \delta_i(\ell) = 0$ and $\alpha_{ij}^2(\ell) < p_i$, then for any $\lim_{\ell\to\infty} \alpha_{ij}(\ell) \neq 0$ we must have by equation (94)

$$\lim_{\ell\to\infty} \lambda_{ij}(\ell) - \lambda_{iN}(\ell) = 0 \tag{99}$$

Therefore, we have

$$\lim_{\ell\to\infty} \frac{\epsilon_{ij}(\ell)}{\alpha_{ij}(\ell)} = \lim_{\ell\to\infty} \alpha_{ij}(\ell)\left(\lambda_{ij}(\ell) - \lambda_{iN}(\ell)\right) = 0 \tag{100}$$

for any $\alpha_{ij}(\ell)$ which does not approach zero.

Thus, we can choose ℓ such that the terms $\epsilon_{ij}(\ell)$ are arbitrarily small. This implies that for suitably large ℓ, all codewords $\mathbf{s}_i(\ell)$ are arbitrarily close to being eigenvectors of $\mathbf{R}(\ell)$, thus completing the proof.

We define such codeword convergence as *convergence in class.* $\bullet$

Of course, Theorem 14 is mute on the relative values of the $\lambda_{iN}(\ell) + p_i$ which define classes. We thus have the following corollary, stated without proof:

Corollary 2 *With greedy+ interference avoidance, codewords might not necessarily converge to a set of optimal classes, but could theoretically become trapped in local TSC minima corresponding to a suboptimal set of classes.*

We can now formalize the greedy interference avoidance algorithm by adding an additional step.

4. Stop if $\forall \mathbf{s}_i$ and some suitably small threshold $\nu > 0$ we have

$$|\sum_{j \in \mathbf{J}_i(\ell)} \frac{\epsilon_{ij}(\ell)}{\alpha_{ij}(\ell)} \boldsymbol{\phi}_{ij}(\ell)| < \nu$$

What constitutes a "suitably small" ν is discussed in Section 5.5 where some practical issues are explored.

B When Single User Warfare Fails

Suppose

$$\mathbf{W} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{101}$$

and

$$\mathbf{s}_1 = \mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{102}$$

so that

$$\mathbf{W} + \mathbf{S}\mathbf{S}^\top = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \tag{103}$$

with a TSC value of 5. Notice that in the context of unequal power signals, the conditions of (1) are violated by $\beta^2 = 0$ and $\delta = 1$ so that $\Delta \leq 0$. The optimal codeword set is

$$\mathbf{s}_1 = \begin{bmatrix} -\frac{1}{2} \\ \sqrt{\frac{3}{4}} \end{bmatrix} \tag{104}$$

$$\mathbf{s}_2 = \begin{bmatrix} \frac{1}{2} \\ \sqrt{\frac{3}{4}} \end{bmatrix} \tag{105}$$

where

$$\mathbf{W} + \mathbf{S}\mathbf{S}^\top = \begin{bmatrix} 3/2 & 0 \\ 0 & 3/2 \end{bmatrix} \tag{106}$$

is achieved with a TSC value of 9/2 and the users reduce their interference level from 1 to 1/2, the absolute minimum for this case [1, 5, 7, 8].

C Sum Capacity Derivation

We modify the approach in [6] to include the nonwhite, possibly correlated Gaussian channel described by equation (1). The mutual information between $\mathbf{y}$ and $\mathbf{b}$ is [24, 25]

$$I(\mathbf{y};\mathbf{b}) = h(\mathbf{y}) - h(\mathbf{y}|\mathbf{b}) = h(\mathbf{y}) - h(\mathbf{w}) \tag{107}$$

This quantity is upper bounded by assuming that $\mathbf{y}$ is a Gaussian random vector. Since $\mathbf{b}$ and $\mathbf{w}$ are assumed zero mean and independent and the components of b are also assumed independent, we have $\text{cov}(\mathbf{y}) = \mathbf{S}\mathbf{S}^\top + \mathbf{W}$ where $\mathbf{W}$ is the covariance of the noise vector. This leads directly to [24, 25]

$$C_{\mathbf{s}} = \frac{1}{2}\log\left[(2\pi e)^N \left|\mathbf{S}\mathbf{S}^\top + \mathbf{W}\right|\right] - \frac{1}{2}\log\left[(2\pi e)^N \left|\mathbf{W}\right|\right] \tag{108}$$

which reduces to

$$C_{\mathbf{s}} = \frac{1}{2}\log\left|\mathbf{S}\mathbf{S}^\top + \mathbf{W}\right| - \frac{1}{2}\log\left|\mathbf{W}\right| \tag{109}$$

We then define the eigenvalues of the noise covariance matrix $\mathbf{W}$ as $\{\sigma_i\}$, $i = 1, \cdots, N$. Likewise we define the eigenvalues of the matrix $\mathbf{S}\mathbf{S}^\top + \mathbf{W}$ as $\{\lambda_i\}$, $i = 1, \cdots, N$ and obtain the sum capacity in terms of eigenvalues

$$C_{\mathbf{s}} = \frac{1}{2}\sum_{i=1}^{N}\log\lambda_i - \frac{1}{2}\sum_{i=1}^{N}\log\sigma_i \tag{110}$$

Since the eigenvalues $\{\sigma_i\}$ are fixed, capacity maximization depends only on the choice of the $\{\lambda_i\}$.

D Proof of Theorem 5

Proof: Theorem 5

First form the function

$$Q(\mathbf{x}, \mathbf{r}) = \sum_{j=1}^{N} g(\mu x_j + (1-\mu) r_j) \tag{111}$$

where $\mu \in (0,1)$ and differentiate with respect to μ to obtain

$$\frac{dQ(\mathbf{x},\mathbf{r})}{d\mu} = \sum_{j=1}^{N} g'(\mu x_j + (1-\mu)r_j)(x_j - r_j) \tag{112}$$

We note that if $\frac{dQ(\mathbf{x},\mathbf{r})}{d\mu} \geq 0$ for all $\mu \in (0,1)$ then $G(\mathbf{x}) \geq G(\mathbf{r})$.

Now for notational convenience, define $\xi_j(\mu) = g'(\mu x_j + (1-\mu)r_j)$ and then define $\Delta\xi_j(\mu) = \xi_j(\mu) - \xi_{j-1}(\mu)$ with $\Delta\xi_1(\mu) = \xi_1(\mu)$. We then note that for any non-negative μ, the quantity $\mu x_j + (1-\mu)r_j$ is a non-increasing sequence since the x_j and r_j are non-increasing. Therefore, if $g()$ is convex, then $\xi_j(\mu)$ is non-increasing in j which in turn implies that $\Delta\xi_j(\mu) \leq 0$.

We then have

$$\sum_{j=1}^{N} \xi_j(\mu)x_j = \Delta\xi_1(\mu) + \xi_2(\mu)(1-x_1) + \cdots = \sum_{j=1}^{N} \Delta\xi_j(\mu)(1 - F_X(j-1)) \tag{113}$$

Likewise

$$\sum_{j=1}^{N} \xi_j(\mu)r_j = \sum_{j=1}^{N} \Delta\xi_j(\mu)(1 - F_R(j-1)) \tag{114}$$

so that

$$\sum_{j=1}^{N} \xi_j(\mu)[x_j - r_j] = \sum_{j=1}^{N} \Delta\xi_j(\mu)[F_R(j-1) - F_X(j-1)] \geq 0 \tag{115}$$

Thus, $G(\mathbf{x}) \geq G(\mathbf{r})$. For $g()$ concave, the reverse, $G(\mathbf{x}) \leq G(\mathbf{r})$ is true. •

E Acknowledgements

The author would like to thank Pablo Anigstein for careful early review of the manuscript and in particular catching a basic error in the exposition which had propagated through an earlier manuscript and would have caused reader consternation and great author embarassment. In addition, the author is indebted to the anonymous reviewers who carefully read the manuscript and offered constructive criticism and corrections where necessary. The paper is much better as a result of their efforts. In particular, Anonymous Reviewer B was an analytic bulldog who did not rest until certain points about the basic underlying algorithm (points originally taken for granted) were settled. Reviewer B went well beyond the call of duty on this paper and in fact, did more work than co-authors on some past papers! I cannot thank him or her enough.

References

[1] S. Ulukus and R. D. Yates. Iterative signature adaptation for capacity maximization of cdma systems. In *Allerton Conf. on Comm., Control and Computing*, September 1998.

[2] D. C. Popescu and C. Rose. Interference Avoidance and Dispersive Channels. A New Look at Multicarrier Modulation. In *Proc. 37th Allerton Conf. on Communication, Control, and Computing*, pages 505–514, Monticello, IL, September 1999.

[3] S. Ulukus and R. D. Yates. Iterative construction of optimum signature sequence sets in synchronous CDMA systems. *IEEE Trans. Info. Theory*, 2001. Accepted and available at http://www.winlab.rutgers.edu/∼ryates.

[4] C. Rose, S. Ulukus, and R. Yates. Interference avoidance for wireless systems. In *Vehicular Technology Conference*, pages 2.05–3, May 2000. Tokyo.

[5] C. Rose, S. Ulukus, and R. Yates. Interference Avoidance in Wireless Systems. *IEEE JSAC*. (submitted 4/2000, see www.winlab.rutgers.edu/∼crose/papers/avoid16.ps).

[6] M. Rupf and J.L. Massey. Optimum sequence multisets for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 40(4):1226–1266, July 1994.

[7] P. Viswanath, V. Anantharam, and D. Tse. Optimal Sequences, Power Control and Capacity of Spread Spectrum Systems with Multiuser Linear Receivers. *IEEE Transactions on Information Theory*, 45(6):1968–1983, September 1999.

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

[9] D. C. Popescu and C. Rose. Multiaccess Dispersive Channels: Maximizing Sum Capacity and Interference Avoidance. *IEEE Transactions on Information Theory*. submitted 12/2000.

[10] P. Anigstein and V. Anantharam. On the Convergence of the MMSE Algorithm for Interference Avoidance. In *Proc. 38th Allerton Conf. on Communication, Control, and Computing*, October 2000.

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

[12] P. Viswanath and V. Anantharam. Total Capacity of Vector Channels. College of Engineering UCB/ERL Memorandum 99/47, U.C. Berkeley, May 1999.

[13] Gilbert Strang. *Linear Algebra and Its Applications.* Academic Press, second edition, 1992.

[14] Special issue on software radio. *IEEE Personal Communications Magazine*, 6(4), August 1999. Editors: K-C. Chen and R. Prasad and H.V. Poor.

[15] I. Seskar and N. Mandayam. Software Defined Radio Architectures for Interference Cancellation in DS-CDMA Systems. *IEEE Pers. Comm. Mag.*, 6(4):26–34, August 1999.

[16] Ivan Seskar and Narayan B. Mandayam. A Software Radio Architecture for Linear Multiuser Detection. *IEEE JSAC*, 17(5):814–823, May 1999.

[17] T. Hentschel, M. Henker, and G. Fettweis. The Digital Front-End of Software Radio Terminals. *IEEE Personal Communications Magazine*, 6(4):40–46, August 1999.

[18] A.K. Salkintzis, H. Nie, and P.T. Mathiopoulos. ADC and DSP Challenges in the Development of Software Radio Base Stations. *IEEE Personal Communications Magazine*, 6(4):47–55, August 1999.

[19] S.P. Reichhart, B. Youmans, and R. Dygert. The Software Radio Development System. *IEEE Personal Communications Magazine*, 6(4):20–25, August 1999.

[20] H.L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I.* Wiley, New York, 1968.

[21] D. Gross and C.M. Harris. *Fundamentals of Queueing Theory.* Wiley, second edition, 1985.

[22] M. Shaked and J. G. Shanthikumar. *Stochastic Orders and Their Applications.* Acadmic Press, 1994.

[23] C. Rose and R. Yates. Minimizing the average cost of paging under delay constraints. *ACM Wireless Networks*, 1(2):211–219, 1995.

[24] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* Wiley-Interscience, 1991.

[25] R.G. Gallager. *Information Theory and Reliable Communication.* Wiley, 1968.