On the Limitation of Linear MMSE Detection

Ateet Kapur†, Clifford T. Mullis‡, and Mahesh K. Varanasi‡

† Communication Theory Group
Swiss Federal Institute of Technology (ETH) Zurich
CH-8092 Zurich, Switzerland
e-mail: ateet@ee.ethz.ch

‡ Dept. of Electrical Engineering
University of Colorado at Boulder
Boulder, CO 80309-0425, USA
e-mail: {varanasi, mullis}@schof.colorado.edu

Abstract

This paper highlights the severe performance limitation of overloaded CDMA systems under linear minimum mean-squared error (MMSE) detection. We prove that with such a simple receiver structure it is not possible to construct signal sets to even satisfy the basic requirement that every user’s uncoded bit error rate (BER) decays exponentially as noise vanishes. This result holds for arbitrary received energies, modulation, and any load factor strictly greater than one. Moreover, it is proved for finite-size CDMA systems.

1 INTRODUCTION

This paper is concerned with analyzing the effect of the little-known interference limitation of linear MMSE detection on the performance of overloaded CDMA systems.

We consider the basic synchronous $K$-user CDMA channel as given by

$$r = \sum_{k=1}^{K} A_k b_k s_k + n$$  \hspace{1cm} (1)

where $A_k$, $b_k$, and $s_k$ represent the (complex) received amplitude, data, and unit-energy spreading sequence of user $k$, respectively, and $n$ white Gaussian noise. We refer to $s_k$ simply as signal, and to $S = [s_1, \ldots, s_K]$ as the signal matrix (without loss of generality, we assume that $S$ is full rank). A CDMA system with processing gain $N$ is said to be overloaded (low-rank, or oversaturated) if the load factor $\beta = K/N$ is strictly greater than one, which implies that the signals are linearly dependent and that the $N \times K$ signal matrix has full row-rank equal to $N$. Such systems are of practical importance when bandwidth is at a premium, such as in bandwidth-efficient multiuser communication (see, e.g., [1]).

The receiver employs linear MMSE detection, which is known to have a simple user-separating structure [2, 3]. Specifically, it consists, say for user $k$, of the filter

$$\hat{f}_k = A_k H^{-1} s_k,$$

where $H = E[rr^H]$ ($H$ stands for hermitian transpose), that minimizes the mean-squared error $E[|f^H r - b_k|^2]$ over all linear transformations $f \in \mathcal{F}^N$. The filter output is the linear MMSE estimate of the symbol $b_k$, and it is fed to the decision rule

$$\hat{b}_k \in \arg \min_{\alpha \in \mathcal{A}} |f_k^H (r - \alpha s_k)|^2.$$  \hspace{1cm} (2)

An alternative rule that specifies the biased linear MMSE detector is

$$\hat{b}_k \in \arg \min_{\alpha \in \mathcal{A}} |f_k^H r - \alpha|^2.$$  \hspace{1cm} (3)

Its asymptotic (low noise) BER, however, is provably no better than rule (2). For simplicity, we assume that the data are drawn from the common $M$-ary and unit average-energy alphabet $\mathcal{A} = \{\alpha_1, \ldots, \alpha_M\}$, but our results extend to the more general scenario where users employ distinct alphabets.

Linear MMSE detection has an attractive structure which, however, is too limited to efficiently cope with high interference such as arises in overloaded systems: a user whose signal is linearly independent of the interfering signals effectively experiences a single-user channel in the low-noise limit, albeit a degraded one; by contrast, a user with a linearly dependent signal experiences irreducible multiple access interference (MAI) even asymptotically [4]. In under or fully loaded CDMA systems, it is possible to allocate linear independent signals to all users and thus avoid irreducible MAI. Such an allocation is not possible, however, in overloaded CDMA systems.

The effects of the resulting interference limitation have been analyzed in different contexts. With respect to the signal-to-interference ratio (SIR) criterion, there exists user-specified SIR requirements that are not admissible, in that there are no power and signal allocations that meet these requirements [5]. By contrast, when the receiver employs decision feedback MMSE detection, every set of SIR requirements is admissible. In fact, the power and signal allocation that minimize the
total power are specified in [6]. Finally, analysis based on the asymptotics of large systems (i.e., as \( K \) and \( N \to \infty \) with \( \beta \) fixed) yields that the spectral efficiency of linear detection (not just linear MMSE detection) with random signals is severely limited in overloaded systems compared to nonlinear detection [7, 8].

In this paper, we prove that no matter how the signals are designed and powers allocated, the uncoded BER of at least one user floors when the receiver in an overloaded CDMA system employs linear MMSE detection.

2 OUR MAIN RESULT

Uncoded BER has received much attention as a performance measure of multiuser detection. It is not tractable in general, but is characterized asymptotically by the asymptotic effective energy (AEE), which is defined for any user (and denoted by \( e_k \) for user \( k \)) as the energy required by the (optimum) matched-filter detector in a single-user channel to achieve, in the low-noise limit, the same BER as the energy required by the (optimum) matched-filter detector in the multiuser channel [9]. An AEE that is identically equal to zero means that the corresponding user’s BER floors. Similarly, the joint error rate (JER) is characterized asymptotically by the symmetric energy, denoted by \( e \) (or \( e(S) \) for the signal matrix \( S \)), which is equal to the smallest AEE [9]. When \( e \) is identically equal to zero, the BER of at least one user, and hence the JER too, floors.

In a recent paper, we have proved that the so-called generalized Welch-bound equality (WBE) signals, which are known to have a host of optimality properties in the context of overloaded CDMA systems, always yield a JER that floors (i.e., \( e = 0 \)) when the receiver employs linear MMSE detection [10]. In fact, the BER of every user floors when in addition the received energies are equal. So the natural question is as to whether there exist signal sets that yield non-zero symmetric energy? If so, how can we characterize signal sets that yield maximum symmetric energy? Unfortunately – and this is the main result of this paper – even the answer to the former question is “no”. Specifically, consider the following rather general theorem, where \( E = \text{diag}\{E_1, \ldots, E_K\} \) is the diagonal matrix of received energies, i.e., \( E_k = |A_k|^2 \).

**Theorem 1** Given an arbitrary system triplet \((K, N, E)\) where \( N < K \), any \( N \times K \) signal matrix has symmetric energy equal to zero under linear MMSE detection and for any modulation size \( M \), i.e.,

\[
\forall S \in \mathbb{C}^{N \times K} \text{ with unit-norm columns, } \quad e(S) = 0.
\]

3 PROOF

The asymptotic (in the low-noise limit) form of linear MMSE detection and hence its AEE depend critically on the signal space geometry. We have fully characterized this dependence in [4], and we succinctly summarize it here. The linear MMSE detector of any user is asymptotically equivalent to the decorrelator (defined as a detector that projects out the MAI) if and only if the desired signal is linearly independent of the interfering signals, and to the so-called pseudo-decorrelator otherwise. Consequently, if \( A = \text{diag}\{A_1, \ldots, A_K\} \) denotes the diagonal matrix of received amplitudes, the AEE of any user, say user \( k \), under linear MMSE detection is given in the linear independent and dependent case as in [11] and [4, Prop. 2], respectively, by:

\[
\begin{align*}
&\text{if } s_k \notin \text{span}(S_k), \quad e_k = E_k s_k^H P^\perp_{\text{span}(S_k)} s_k \quad (3) \\
&\text{if } s_k \in \text{span}(S_k), \quad \text{see (4) at top of next page}
\end{align*}
\]

In (4), \( S_k, A_k \) and \( E_k \) denote, respectively, the matrices of signals, complex amplitudes, and energies of the interfering users (i.e., users \( 1, \ldots, k-1, k+1, \ldots, K \)), and \( P^\perp_{\text{span}(S_k)} \) denotes the projection orthogonal to the subspace spanned by \( S_k \). We have assumed without loss of generality that user \( k \) is received with a phase equal to zero. We have also introduced the function \( [a]_+ = \max\{0, a\} \), the vector \( x = (S_k A_k)^+ s_k \) (plus stands for the pseudo-inverse) where \( x_n = |x_n|e^{j\arg x_n} \) denotes its \( n \)-th component, and the modulation parameter \( \alpha_A = \frac{\max_{\min(s_j)} |a_j - a_i|}{\min(s_j)} \). For a square \( M\)-QAM constellation (i.e., when \( M = 2^m \)), \( \alpha_A = \frac{\sqrt{M - 1}}{\sqrt{2}} \).

Note that in the linear dependent case, \( S_k \) has full row-rank, so that \( S_k E_k S_k^H \) is positive definite and its inverse is well defined.

In the linear independent case, we recognize in (3) the AEE of the decorrelator, i.e., the squared norm of the desired signal after its projection orthogonal to the interference subspace and scaled by the received energy. The AEE of linear MMSE detection as given in (4) for the linear dependent case is more involved. The reason, as we explain in [4], is that the pseudo-decorrelator (to which the linear MMSE detector converges) cancels out the interference selectively. Specifically, it cancels out only the interfering signals that the desired signal is linearly independent of. Note that by contrast, the decorrelator cancels out the desired signal along with the interference in this case.

To prove our claim that \( \min_k e_k = 0 \) for any signal matrix \( S \), we make a series of simplifications. First notice that we only need to consider the modulation size \( M \) for which \( \alpha_A \) is the smallest. For complex-valued (real-valued) modulations, this corresponds to QPSK.
\[ e_k = E_k \left[ \frac{s_k^H (S_k E_k S_k^H)^{-1} s_k - \alpha_A \sqrt{\frac{2}{E_k}} \sum_{n=1}^{K-1} |x_n| (|\cos \arg x_n| + |\sin \arg x_n|)}{\sqrt{s_k^H (S_k E_k S_k^H)^{-2} s_k}} \right]^2 + \right. \\
\left. \left. (4) \right] \]

(BPSK), for which \( \alpha_A = 1/\sqrt{2} \) (\( \alpha_A = 1/2 \)). Indeed, it is clear from the subtractive term in the numerator of (4) that any modulation with a larger \( \alpha_A \) cannot have a better maximum symmetric energy. Intuitively, we only need to consider the modulation that has the largest minimum distance between any two symbols (for a given average energy).

Second, if \( S \) is an arbitrary \( N \times K \) matrix (\( N < K \)) with unit-norm columns, we can reduce the problem to one that involves a simpler signal matrix and simpler signal space geometry. Indeed, consider the largest partition of the form \( S = [S_1, \ldots, S_L] \) [largest in the sense that \( L \) is maximum] where the subspaces spanned by the submatrices \( \{S_i\}_{i=1}^L \) are linearly independent, i.e., \( \text{span}(S_i) \cap \text{span}(S_j) = \{0\}, \forall i \neq j \). Because \( N < K \), there is at least one submatrix that is comprised of at least two signals. Suppose it is \( S_1 \), and denote by \( N_1 \) and \( K_1 \) its rank and number of columns, respectively. Now, consider the following straightforward lemma.

**Lemma 1** In a multiuser system with \( K \) active users and linear MMSE detection, the AEE of any user cannot decrease if one user leaves the system.

It follows from this lemma that the symmetric energy of the signal matrix \( S \), i.e., when all \( K \) users are active, is less than or equal to the symmetric energy for the reduced system specified by \( S_1 \), i.e., when only users whose signals are in \( S_1 \) are active. We will show in fact that the symmetric energy of this reduced \( K_1 \)-user CDMA system is identically equal to zero. But rather than working with the signal matrix \( S_1 \) which need not have full row-rank, we further reduce it to a full row-rank matrix \( \bar{S}_1 \) that has dimensions \( N_1 \times K_1 \) and that is obtained from the singular value decomposition of \( S_1 \). Finally, we appeal to Lemma 1 once more, and conclude that we need to consider only a submatrix of \( \bar{S}_1 \) that is comprised of \( N_1 + 1 \) signals that are linearly independent. In other words, we don’t need to account for all \( K_1 \) users in the reduced CDMA system, only for sufficiently many so that the system be indeed overloaded.

We have thus reduced the problem to one that involves only a full row-rank signal matrix that has exactly one more column than number of rows, and that is comprised only of linear dependent and unit-norm signals. For simplicity, we denote this matrix by \( S \) and its dimensions by \( N \times (N + 1) \). It is clear that the AEE of any user in this reduced system is given by the expression in (4).

As a final simplification, we note that the expression for the AEE in the linear dependent case as given in (4) is in general unwieldy. Instead, we consider a simpler upper bound (valid for any \( M \)) which is derived by noting that \( |\cos \theta| + |\sin \theta| \geq 1 \). Specifically, if \( s_k \in \text{span}(S_k) \), then \( e_k \leq U_k \), where

\[ U_k = E_k \left[ \frac{s_k^H (S_k E_k S_k^H)^{-1} s_k - E_k^{-1/2} ||(S_k A_k)^{-1} s_k||_1}{\sqrt{s_k^H (S_k E_k S_k^H)^{-2} s_k}} \right]^2 + \]

and ||·|| denotes the \( f_1 \)-norm, i.e., \( ||z||_1 = \sum_i |z_i| \). Here we have recognized that the submatrix \( S_k \) (and hence \( S_k A_k \)) is invertible. From this upper bound, we have that the symmetric energy of \( S \) is upper bounded by \( \min_k U_k \). Therefore, to prove that \( e(S) = 0 \), we need to find at least one user for which the upper bound (5) is equal to zero, or equivalently for which the numerator of the upper bound is negative.

Consider the matrix \( T = SA = [t_1, \ldots, t_K] \) and the vectors \( y_k = T_k^T t_k \) for \( k = 1, \ldots, K \) (here, \( T_k = [t_n, n \neq k] \)). The \( N \)-dimensional vector \( y_k \) represents the coordinates of \( t_k \) along the interfering vectors in \( T_k \). It is then easy to express the two terms of the numerator of the upper bound (5) in terms of these coordinate vectors, namely:

\[ s_k^H (S_k E_k S_k^H)^{-1} s_k = E_k^{-1} ||y_k||_2^2, \]

\[ ||(S_k A_k)^{-1} s_k||_1 = E_k^{-1/2} ||y_k||_1. \]

Thus, the numerator is proportional to \( ||y_k||_2^2 - ||y_k||_1 \), and to prove that the minimum upper bound is equal to zero, we only need to show that \( \min_k (||y_k||_2^2 - ||y_k||_1) = 0 \).

Since the matrix \( T \) has rank equal to \( N \), its null space has dimension one, and hence there is a non-zero \((N + 1)\)-dimensional vector \( a \) (unique up to scaling) such that \( Ta = 0 \). Moreover, all components of \( a \) are non-zero. For now, we assume that the origin is in the convex hull of \( \{t_k\}_{k=1}^K \), so that these components are strictly positive, i.e., \( a_k > 0 \) for \( k = 1, \ldots, N + 1 \). We remove this restriction later. From the equality
\( T \mathbf{a} = T_k \mathbf{a}_k + a_k \mathbf{t}_k = \mathbf{0} \), where \( \mathbf{a}_k \) denotes the \( N \)-dimensional vector obtained from \( \mathbf{a} \) by removing its \( k \)-th component \( a_k \), it follows that the coefficient vectors can be expressed in terms of the null space vector \( \mathbf{a} \), namely \( \mathbf{y}_k = -(1/a_k) \mathbf{a}_k \) for all \( k \). Consider now the constant \( \gamma_0 = |a|_1/|a|_2^2 \) and the polynomial function

\[
f(\gamma) = |a|_2^2 \gamma (\gamma - \gamma_0).
\]

Evaluating this polynomial for \( \gamma = 1/a_k \), we have that

\[
f(1/a_k) = \frac{1}{a_k^2} \sum_{n=1}^{N+1} a_n^2 - \frac{1}{a_k} \sum_{n=1}^{N+1} a_n
\]

\[
= |a_k|_2^2 - |a_k|_1
\]

\[
= |\mathbf{y}_k|_2^2 - |\mathbf{y}_k|_1,
\]

and also that

\[
\gamma_0 - 1/a_k = \frac{\sum_{n=1}^{N+1} a_n(a_k - a_n)}{a_k \sum_{n=1}^{N+1} a_n^2}.
\]

If the components \( \{a_k\}_{k=1}^{N+1} \) are all equal, it follows that \( \gamma_0 = 1/a_k \), that \( f(1/a_k) = 0 \), and hence that \( |\mathbf{y}_k|_2^2 - |\mathbf{y}_k|_1 = 0 \) for all \( k \). Consequently, the upper bound is identically equal to zero for all users. A sufficient condition for this case is that the received energies be equal and that the users be allocated WBE signals.\(^1\)

As an aside, we note that in this case (\( K = N + 1 \)), the WBE signals form a simplex signal set.

On the other hand, if the components \( \{a_k\}_{k=1}^{N+1} \) are not all equal, consider \( a_{\mu} = \min_k a_k \) and \( a_{\nu} = \max_k a_k \). It follows from (6) that \( 1/a_{\mu} < \gamma_0 \), and hence that \( f(1/a_{\mu}) = |\mathbf{y}_{\mu}|_2^2 - |\mathbf{y}_{\mu}|_1 < 0 \). This proves our claim that the numerator of the upper bound (5) is negative for at least one user, and hence that this user’s AEE is identically equal to zero. Furthermore, it can similarly be shown that \( |\mathbf{y}_{\nu}|_2^2 - |\mathbf{y}_{\nu}|_1 > 0 \), so that at least one user has an AEE that is non-zero.

In case the components of \( \mathbf{a} \) are not all positive, consider the diagonal matrix

\[
\mathbf{D} = \text{diag} \left( \text{sign}(a_1) e^{-j \arg(a_1)}, \ldots, \text{sign}(a_K) e^{-j \arg(a_K)} \right)
\]

The proof follows along the same arguments as in the previous case if we consider, instead of \( \mathbf{T} \) and \( \mathbf{a} \), the matrix \( \tilde{\mathbf{T}} = \mathbf{T} \mathbf{D} \) and the vector \( \tilde{\mathbf{a}} = \mathbf{D} \mathbf{a} \) which is in the null space of \( \mathbf{T} \) and which has positive components.

1Assuming that \( E_k = E, \forall k \), the signal matrix of a WBE signal set is characterized by \( SS^H = (KE)N \mathbf{I}_N \).

4 CONCLUSION

This paper has highlighted a severe limitation of linear MMSE detection in the context of overloaded CDMA systems. Namely, there do not exist spreading sequences that satisfy even the basic requirement that the BER of every user decreases exponentially as noise vanishes. This result holds for arbitrary overload, modulation, and received energies, and therefore power control cannot mitigate this limitation. Consequently, the poor JER performance of generalized WBE signals which we highlighted in [10] is not due to poor properties of these signals, but rather to the intrinsic limitations of linear MMSE detection. Furthermore, this limitation suggests that reliable and spectrally efficient multiple access requires a more robust receiver structure: in short, nonlinear detection. In fact, such a structure need not entail a prohibitive increase in receiver complexity, and it is better suited to uplink transmission. In particular, systematic constructions of bandwidth-efficient signal sets have been proposed for the case of decision feedback MMSE detection which is only marginally more complex than its linear counterpart we have studied in this paper. The proposed constructions achieve arbitrary user-specified rate-tuples [1] or AEEs [12].

References


