Since $\mathcal{Q} \subset L^1(\partial\mathcal{D})$, this proves the right-hand side of (29). Now, we consider a special function $q_\epsilon \in \mathcal{Q}$

$$q_\epsilon(e^{i\tau}) = \begin{cases} -\epsilon/4 & \text{for} \quad -\pi \leq \tau \leq -\pi + 1/\epsilon \\ \epsilon/4 & \text{for} \quad -1/\epsilon \leq \tau \leq 1/\epsilon \\ -\epsilon/4 & \text{for} \quad \pi - 1/\epsilon \leq \tau \leq \pi \\ 0 & \text{otherwise} \end{cases}$$

By construction $\int_{-\pi}^{\pi} q_\epsilon(e^{i\tau})\mathrm{d}\tau = 0$. For this function, the following lower bound for $|I_r|$ is obtained

$$|I_r(q_\epsilon)| \geq \frac{1}{8\pi}\log\left(\frac{(1+r)^2}{1-2r\cos(1/\epsilon)+r^2}\right) - \frac{1}{8\pi}\log\left(\frac{(1+r)^2}{1-2r\cos(\pi-1/\epsilon)+r^2}\right).$$

This $|I_r(q_\epsilon)|$ is monotonic increasing as ϵ increases with

$$\lim_{\epsilon\to\infty}|I_r(q_\epsilon)| = \frac{1}{4\pi}\log\left(\frac{1+r}{1-r}\right).$$

Since q_ϵ is one special element of $\mathcal{Q}$ this proves the left-hand inequality of (29). □

References

[1] N. Wiener and E. Hopf, "Über eine Klasse singulärer Integralgleichungen," *Sitzungsberichte der Königlich Preußische Akademie der Wissenschaften zu Berlin*, pp. 696–706, Dez. 3, 1931.
[2] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. Cambridge, MA: The MIT Press and Wiley, 1949.
[3] A. N. Kolmogorov, "Interpolation und extrapolation von stationären zufälligen folgen," *Bull. Acad. Sci. (URSS), Ser. Math.*, vol. 5, pp. 3–14, 1941.
[4] L. A. Zadeh and J. R. Ragazzini, "Optimum filters for the detection of signals in noise," in *Proc. I.R.E.*, Oct. 1952, vol. 40, pp. 1223–1231.
[5] S. Darlington, "{Linear Least-Square Smoothing and Prediction, with Applications}," *Bell Sys. Techn. J.*, vol. 37, pp. 1221–1294, Sept. 1958.
[6] H. W. Bode and C. E. Shannon, "A simplified derivation of linear least square smoothing and prediction theory," *Proc. IRE*, vol. 38, pp. 417–425, Apr. 1950.
[7] L. Franks, *Signal Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1969.
[8] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996, etc..
[9] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.
[10] S. Elliott, *Signal Processing for Active Control*. : Elsvier, 2000.
[11] A. H. Sayed and T. Kailath, "A survey of spectral factorization methods," *Numer. Linear Algebra Appl.*, vol. 8, pp. 467–496, 2001.
[12] L. Y. Wang, "Lipschitz continuity of inner-outer factorization," *Syst. Contr. Lett.*, vol. 16, pp. 281–287, 1991.
[13] J. B. Garnett, *Bounded Analytic Functions*. New York: Academic, 1981.
[14] W. Rudin, *Real and Complex Analysis*. London, U.K.: McGraw-Hill, 1970.
[15] H. Boche and V. Pohl, "Continuity of the outer factorization and mapping properties with applications," in *Proc. Joint 44th IEEE Conf. Decision Contr. European Contr. Conf. (CDC-ECC)*, Seville, Spain, Dec. 2005, pp. 548–553.
[16] A. Zygmund, *Trigonometric Series*. Cambridge, U.K.: Cambridge University Press, 1968.
[17] R.-L. Baire, "Sur les fonctions des variables réeles," *Ann. of Math.*, vol. 3, pp. 1–123, 1899.
[18] L. V. Kantorovich and G. P. Akilow, *Functional Analysis in Normed Spaces*. Oxford, U.K.: Pergamon, 1964, etc.
[19] K. M. Dyakonov, "Equivalent norms on Lipschitz-type spaces of holomorphic functions," *Acta Math.*, vol. 178, pp. 143–167, 1997.
[20] H. Boche and V. Pohl, "Characterization of holomorphic functions in terms of their moduli," *Complex Var. Theory Appl.*, vol. 50, no. 13, pp. 1025–1029, Oct. 2005.
[21] ——, "Spectral factorization, whitening- and estimation filter—Stability, smoothness properties and FIR approximation behavior," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Adelaide, Australia, Sept. 2005, pp. 1701–1705.
[22] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.
[23] L. Cohen, "The history of noise," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 20–45, Nov. 2005.
[24] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.*, vol. 23, no. 3, pp. 282–332, 1944, vol. 24, no. 1, pp. 46–156, 1945.
[25] H. Boche and V. Pohl, "BIBO norm in spectral factorization and wiener filtering," in *Proc. 9th Canadian Workshop Inf. Theory (CWIT)*, Montréal, QC, Canada, Jun. 2005, pp. 347–350.
[26] D. Jackson, "On approximations by trigonometric sums and polynomials," *Trans. Amer. Math. Soc.*, vol. 13, pp. 491–515, 1912.
[27] V. Kucera, "Factorization of rational spectral matrices: A survey of methods," in *Proc. IEE Int. Conf. Contr.*, Edinburgh, Scotland, UK, Mar. 1991, vol. 2, pp. 1047–1048.
[28] P. L. Butzer and R. J. Nessel, *Fourier analysis and approximation*. Basel, Germany: Birkhäuser-Verlag, 1971.

On the Limitation of Linear MMSE Detection

Mahesh K. Varanasi, *Senior Member, IEEE*, Clifford T. Mullis, and Ateet Kapur, *Member, IEEE*

***Abstract*—This correspondence highlights the performance limitation of linear minimum mean-squared error (mmse) detection in underdetermined vector Gaussian channels (as in overloaded code-division multiple-access (CDMA) systems) where the number of symbols (users) exceeds the signal space dimension (spread factor). It is shown that for such a simple receiver it is not possible to construct signal sets (or spreading codes) to even satisfy the basic requirement that every user's symbol error probability decays exponentially as noise power vanishes. This result holds for arbitrary received energies, modulation schemes, and any strictly underdetermined system with a finite signal space dimension and a finite number of users.**

***Index Terms*—Code-division multiple-access (CDMA), Gaussian multiple-access channels, multiuser detection, minimum mean square error (mmse) detection, signal design.**

I. Introduction

This correspondence analyzes the interference limitation of linear minimum mean square error (mmse) detection in underdetermined vector Gaussian channels such as those that arise, for example, in overloaded synchronous CDMA systems. A K-input, N-output Gaussian vector channel is described by the following discrete-time input–output relation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

Manuscript received February 2, 2005; revised May 13, 2006. This work was supported in part by the NSF under Grants CCR-0112977 and CCF-0431170. The material in this correspondence was presented in part at the International Symposium on Information Theory and Its Applications, Parma, Italy, Oct.2004.

M. K. Varanasi and C. T. Mullis are with the Department of Electrical Engineering, University of Colorado at Boulder, CO 80309-0425 USA (e-mail: varanasi@schof.colorado.edu; mullis@schof.colorado.edu).

A. Kapur was with the Department of Electrical Engineering, University of Colorado at Boulder, CO 80303 USA. He is now with Broadcom Corporation, Sunnyvale, CA 94086 USA.

Communicated by X. Wang, Associate Editor for Detection and Estimation.
Digital Object Identifier 10.1109/TIT.2006.880039

where $\mathbf{x}$ is the K-dimensional vector that represents the inputs, $\mathbf{y}$ is the N-dimensional vector of outputs, $\mathbf{H}$ the $N \times K$ is the channel transition matrix, and $\mathbf{n}$ is the N-dimensional Gaussian noise vector. This channel is said to be underdetermined if $K < N$.

The general model given by (1) describes a variety of communication systems, including multi-access channels, multi-antenna point-to-point links, bundled xDSL lines, etc. In particular, synchronous code-division multiple-access (CDMA) systems, wherein multiple users communicate simultaneously and synchronously with a common receiver by modulating their data onto user-specific spreading sequences, are an instance of vector Gaussian channels. In this case, K and N denote, respectively, the number of users and processing (or spreading) gain; x_k, the kth entry of $\mathbf{x}$, is the symbol transmitted by user k, and the channel matrix is the product of two matrices, i.e., $\mathbf{H} = \mathbf{SA}$, where $\mathbf{S} = [\mathbf{s}_1\mathbf{s}_2 \dots \mathbf{s}_K]$ and $\mathbf{A} = \mathrm{diag}\{A_1 A_2 \dots A_K\}$. These matrices denote, respectively, the matrix of spreading sequences and the matrix of (complex) received amplitudes. In other words, user k employs the spreading sequence $\mathbf{s}_k$, which we normalize to have unit norm, to carry its data and is received with amplitude A_k. We refer to the $\{\mathbf{s}_k\}_{k=1}^{K}$ simply as signals and to $\mathbf{S}$ as the signal matrix. Without loss of generality, we assume that $\mathbf{S}$ has full rank (since otherwise $\mathbf{y}$ would not be a minimum sufficient statistics for $\mathbf{x}$) and that the noise is white, i.e., $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_N)$.

A CDMA system is said to be *overloaded* (or *low-rank* or *oversaturated*) if the load $\beta = K/N$ is strictly greater than one, i.e., if the underlying vector Gaussian channel is underdetermined. Since the signals are vectors in an N-dimensional linear signal space, overload implies that they are linearly dependent. Such systems are of practical importance when bandwidth is at a premium, such as in bandwidth-efficient multiuser communication (see, e.g., [1]).

Linear minimum mean-squared error (mmse) detection is a popular approach to recover the transmitted data in vector Gaussian channels. It offers a suboptimal but simple structure that decouples the joint detection problem into K single-input problems, and it is arguably the "best" among the class of linear detectors. We show here, however, that this detector cannot cope with high interference as that arises in underdetermined channels. For concreteness, we use the terminology of overloaded CDMA systems in the rest of the correspondence.

Linear mmse detection in a CDMA system consists of a bank of K filters followed by K single-user decision rules [2], [3]. Specifically, one branch of the detector, say for user k, consists of the filter $\mathbf{f}_k = A_k \mathbf{K}^{-1}\mathbf{s}_k$, where $\mathbf{K} = \mathrm{E}[\mathbf{y}\mathbf{y}^{\mathrm{H}}]$ ($^{\mathrm{H}}$ stands for hermitian transpose), which minimizes the mean-squared error $\mathbf{E}[\|\mathbf{f}^{\mathrm{H}}\mathbf{y} - x_k\|_2^2]$ over all $\mathbf{f} \in \mathbb{C}^N$. The filter output, which is a linear mmse estimate of x_k, is used in the *unbiased* mmse detector specified via the decision rule

$$\hat{x}_k \in \arg\min_{\alpha \in \mathcal{A}} \left| \mathbf{f}_k^{\mathrm{H}}(\mathbf{y} - \alpha A_k \mathbf{s}_k) \right|^2. \tag{2}$$

The more common *biased* linear mmse detector makes the decision $\tilde{x}_k \in \arg\min_{\alpha \in \mathcal{A}} |\mathbf{f}_k^{\mathrm{H}}\mathbf{y} - \alpha|^2$; its asymptotic (low noise) error rate, however, is provably no better than rule (2), and hence the main result of this correspondence that is proved for the unbiased mmse detector also applies to the usual (biased) mmse detector.

For simplicity, we assume that the data symbols are drawn from a common M-ary and unit average-energy alphabet $\mathcal{A} = \{\alpha_1, \dots, \alpha_M\}$, but our results extend to the more general scenario where users employ distinct alphabets.

The performance measure of central interest in this correspondence is the *symmetric energy* introduced in [9]. It is a single-letter characterization of the overall error probability (that not all users are detected correctly) of a multiuser detector in the high SNR regime. Suppose that $P_k = \Pr(\mathcal{E}_k) = \Pr(\{\hat{x}_k \neq x_k\})$ is the probability that user k is received in error (with $\mathcal{E}_k$ denoting user k's error event $\{\hat{x}_k \neq x_k\}$). Define this user's *effective energy* as $e_k(\sigma)$ implicitly via the equality $P_k = Q(\sqrt{e_k(\sigma)}/\sigma)$ where $Q(x) = (2\pi)^{-1/2}\int_x^{+\infty} e^{-t^2/2}\, dt$ is the Gaussian complementary distribution function. In other words, $e_k(\sigma)$ is the energy required by the (optimum) matched-filter detector in a single-user channel to achieve the same error rate as the multiuser detector in the multiuser channel. While $e_k(\sigma)$ is simply an invertible function of the error probability, and hence only as tractable and insightful, its limit $e_k = \lim_{\sigma \to 0} e_k(\sigma)$ is more amenable to analysis and insight and has the simple interpretation as being the equivalent energy in a single user channel and is called the asymptotic effective energy (AEE) [9]. When normalized by $|A_k|^2$, it becomes the asymptotic efficiency [10]. The AEE is of course a function of the detector, the signals, and the energies, but it is always bounded above by the received energy and below by zero. It is identically equal to zero if and only if the corresponding user's error probability "floors" which in the AWGN channel means that it doesn't decay exponentially as noise vanishes.

Similarly, the *symmetric energy*, denoted by e, characterizes the rate of exponential decay of the *joint* error rate (JER) which is the probability that not all users are detected correctly of a multiuser detector in the low-noise limit. In particular, we define the symmetric energy as $e = \lim_{\sigma \to 0} e(\sigma)$ where $\Pr(\cup_{k=1}^{K}\mathcal{E}_k) = Q(\sqrt{e(\sigma)}/\sigma)$. The symmetric energy is easily shown to be equal to the smallest AEE, i.e., $e = \min_k e_k$ [9], and it is identically equal to zero if and only if the JER floors, i.e., *the error rate of at least one user floors*.

In this correspondence, we show that the symmetric energy of the linear mmse detector (biased or unbiased) is identically equal to zero for underdetermined vector Gaussian channels for any choice of signal matrix $\mathbf{S}$ and any set of user energies. It is important to distinguish this more elusive result from the one that simply asserts that a user whose signal is linearly dependent has zero near-far resistance [10]. We elaborate on this point next.

A user with a signal that is linearly independent of the interfering signals effectively experiences a single-user channel in the low-noise limit because the linear mmse detector asymptotically becomes the decorrelator for that user. By contrast, a user whose signal is linearly dependent experiences irreducible multiple access interference (MAI) even asymptotically, as shown in [4]. Consequently, its near-far resistance [10] can easily be shown to be equal to zero, which implies that there exist user powers for which the error rate floors in the low-noise limit.[1]

In under- or fully loaded CDMA systems, linear independent (in fact, orthogonal) signals can be allocated to all users to avoid irreducible MAI. Since such an allocation is not possible in the overloaded case, good signal alternatives are sought. In this respect, the so-called generalized Welch Bound Equality (WBE) signals are especially attractive because they minimize the total squared correlation [12] and have a wealth of optimality properties as a consequence (see, e.g., the introduction of [13]). Nevertheless, it was shown by the authors that their JER performance breaks down under linear mmse detection [13] in that the symmetric energy is identically equal to zero. In this correspondence, we prove a stronger result that, no matter how the signals are designed or what the received powers are, the symmetric energy of the linear mmse detector is identically equal to zero in an overloaded CDMA system.

We describe some related results in the literature here. The limitations of the linear mmse detector in terms of the signal-to-interference ratio (SIR) criterion are known: there exist combinations of SIRs that are not simultaneously achievable by all users, no matter what signal

[1]There may be a large set of user powers, and in fact there often is, for which the asymptotic efficiency, or equivalently, the AEE of a user with a linearly dependent signal is *not* equal to zero and hence this user with a linearly dependent signal can enjoy an exponential rate of decay of error probability as noise power vanishes.

set or power distribution is chosen [5]. By contrast, any combination of SIR's is achievable when the receiver employs decision feedback mmse detection, and in fact the power and signal allocation that minimize the total power are specified in [6]. A large system analysis (i.e., as K and $N \to \infty$ with β fixed) with random signals yields that the spectral efficiency of linear detection in overloaded systems is severely limited compared to nonlinear detection [7], [8].

II. Main Result

Do there exist signal sets that have nonzero symmetric energy under linear mmse detection? If so, can we characterize signal sets that maximize symmetric energy? Unfortunately—and this is the main result of this correspondence—even the answer to the first question is negative. Specifically, consider the following rather general result in which we let $e(\mathbf{S})$ denote the symmetric energy corresponding to the signal set $\mathbf{S}$ and $\mathbf{E} = \mathrm{diag}\{E_1, \ldots, E_K\}$ be the diagonal matrix of received energies, i.e., $E_k = |A_k|^2$.

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

$$\forall \mathbf{S} \in \mathbb{C}^{N\times K} \text{ with unit-norm columns}, \quad e(\mathbf{S}) = 0. \tag{3}$$

III. Proof

The low-noise limit of the linear mmse detector, and hence its AEE, depend on the signal space geometry as characterized in [4]. In particular, the linear mmse detector 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 pseudodecorrelator (which only partially cancels the MAI), otherwise. Therefore, user k has an AEE under linear mmse detection that is given in the linear independent and dependent cases as in [10] and [4, Prop. 2], respectively, by (4) shown at the bottom of the page. In (4), $\mathbf{x} = (\mathbf{S}_k\mathbf{A}_k)^{+}\mathbf{s}_k$ where the superscript $+$ stands for the pseudoinverse, $x_n = |x_n|e^{j\arg x_n}$ is its nth component, the matrices $\mathbf{S}_k$, $\mathbf{A}_k$ and $\mathbf{E}_k$ denote, respectively, the matrices of signals, amplitudes and energies of the interfering users (i.e., users $1, \ldots, k-1, k+1, \ldots, K$). In addition, we assume without loss of generality that user k is received with a phase equal to zero, and we denote for any real number a, $\max\{0, a\}$ as $[a]_+$, and $\alpha_{\mathcal{A}} = \frac{\max_{\alpha\in\mathcal{A}}|\alpha|}{\min_{i\neq j}|\alpha_j - \alpha_i|}$ (for square M-QAM constellations $(M = 2^{2m})$, $\alpha_{\mathcal{A}} = \frac{\sqrt{M}-1}{\sqrt{2}}$). Note that $\mathbf{S}_k$ has full row-rank in the linear dependent case, and hence $\mathbf{S}_k\mathbf{E}_k\mathbf{S}_k^{\mathrm{H}}$ is positive definite and its inverse is well defined. In the expression for AEE in the linearly independent case, $\mathbf{P}^{\perp}_{\mathrm{span}(\mathbf{S}_k)}$ denotes an orthogonal projection matrix that projects onto the orthogonal complement of $\mathrm{span}(\mathbf{S}_k)$.

To prove our claim that $e = \min_k e_k = 0$ for any signal matrix $\mathbf{S}$, we make a series of simplifications. First notice that we only need to consider the modulation size M for which $\alpha_{\mathcal{A}}$ is the smallest. For complex-valued (real-valued) modulations, this corresponds to QPSK (BPSK), for which $\alpha_{\mathcal{A}} = 1/\sqrt{2} (\alpha_{\mathcal{A}} = 1/2)$. Indeed, it is clear from the subtractive term in the numerator of (4) that any modulation with a smaller minimum distance yields a smaller $\alpha_{\mathcal{A}}$ and hence worse symmetric energy.

Second, if $\mathbf{S}$ is an $N \times K$ matrix with unit-norm columns, we can convert the problem to one that involves a reduced signal matrix with simpler signal space geometry. Considering that the rank of $\mathbf{S}$ is N, there exists a set of N mutually linearly independent columns which, without loss of generality, we take to be the first N columns of $\mathbf{S}$ (else, one could permute the columns, renumber users and produce such an $\mathbf{S}$). Consider any column of $\mathbf{S}$ which is not one of the first N columns, say the $(N+1)$st column. Evidently, this column can be written as a linear combination of the first N columns. Suppose that $J \subseteq \{1, \ldots, N\}$ is the index set corresponding to the nonzero coefficients of that linear combination. Construct an expurgated matrix $\mathbf{S}_{\exp}$ which is obtained by starting with $\mathbf{S}$ and deleting all columns from $\mathbf{S}$ except those indexed by J and $N+1$. Evidently, $\mathbf{S}_{\exp}$ can be thought of as a signal matrix in a fictitious expurgated channel with $|J|+1$ users and signal space dimension $|J|$ (the cardinality of J) with the property that all users in this channel are linearly dependent. Moreover, if *any* one user is removed from this system, the resulting signal space dimension remains unchanged (this need not be true of an overloaded system in general). One could write the input–output model this user expurgated system as

$$\mathbf{y}_{\exp} = \mathbf{S}_{\exp}\mathbf{A}_{\exp}\mathbf{x}_{\exp} + \mathbf{n} \tag{5}$$

where $\mathbf{A}_{\exp}$ and $\mathbf{x}_{\exp}$ are obtained by retaining the diagonal elements in $\mathbf{A}$ and the elements in $\mathbf{x}$ that correspond to the expurgated channel. Note that $\mathbf{S}_{\exp}$ has dimension $N \times (|J|+1)$ and $N \geq |J|$. Two cases need to be considered. If $N = |J|$ then we work with the model in (5). If $N > |J|$, reduce the expurgated channel output without loss of sufficiency to have dimension $|J|$. This can be accomplished by using the singular value decomposition $\mathbf{S}_{\exp} = \mathbf{U}\Lambda\mathbf{V}^*$, where $\mathbf{U}$ is $N \times |J|$ and $\mathbf{V}$ is $|J|+1 \times |J|$ with $\mathbf{U}^*\mathbf{U} = \mathbf{I}$, $\mathbf{V}^*\mathbf{V} = \mathbf{I}$, and the diagonal matrix Λ of positive singular values is $|J| \times |J|$. Pre-multiply $\mathbf{y}_{\exp}$ in (5) by $\mathbf{U}^*$ to obtain the model

$$\tilde{\mathbf{y}}_{\exp} = \tilde{\mathbf{S}}_{\exp}\mathbf{A}_{\exp}\mathbf{x}_{\exp} + \tilde{\mathbf{n}} \tag{6}$$

where $\tilde{\mathbf{S}}_{\exp} = \mathbf{\Lambda}\mathbf{V}^*$ is $|J| \times (|J|+1)$ and has full row-rank and $\tilde{\mathbf{n}}$ is additive zero-mean, white, complex Gaussian random $|J|$-dimensional vector with covariance $\sigma^2\mathbf{I}$. In both cases therefore, we have produced a user expurgated channel (by removing some users from the original channel) and with one more user than the signal space dimension with the property that all users in this expurgated channel are linearly dependent. Note also that the expurgated channel has at least two users.

The point of constructing the expurgated channel would be evident from the following lemma whose proof is left to the reader.

Lemma 1: In a vector Gaussian channel (underdetermined or not) with additive white Gaussian noise, the per-user AEE performance in any user expurgated channel is lower bounded by the AEE achieved in the original channel from which the expurgated channel is obtained when linear mmse detectors (tailored to each channel) are used.

The application of the above lemma to our problem is as follows. The lemma implies that the symmetric energy obtained in the original

$$\text{if } \mathbf{s}_k \notin \mathrm{span}(\mathbf{S}_k), \quad e_k = E_k\mathbf{s}_k^{\mathrm{H}}\mathbf{P}^{\perp}_{\mathrm{span}(\mathbf{s}_k)}\mathbf{s}_k$$
$$\text{if } \mathbf{s}_k \in \mathrm{span}(\mathbf{S}_k),$$
$$e_k = E_k\left(\left[\frac{\mathbf{s}_k^{\mathrm{H}}\left(\mathbf{S}_k\mathbf{E}_k\mathbf{S}_k^{\mathrm{H}}\right)^{-1}\mathbf{s}_k - \alpha_{\mathcal{A}}\sqrt{\frac{2}{E_k}}\sum_{n=1}^{K-1}|x_n|(|\cos\arg x_n| + |\sin\arg x_n|)}{\sqrt{\mathbf{s}_k^{\mathrm{H}}\left(\mathbf{S}_k\mathbf{E}_k\mathbf{S}_k^{\mathrm{H}}\right)^{-2}\mathbf{s}_k}}\right]_+\right)^2. \tag{4}$$

channel is less than or equal to the symmetric energy resulting in the expurgated system. Therefore, if we can prove that the symmetric energy of the linear mmse detector for the particular expurgated system we constructed above is zero then the symmetric energy obtained the original (general) overloaded system using a linear mmse detector cannot be positive either, and we would be done. This is in fact what we will do.

Consequently, from now on we will consider an overloaded system with one more user than signal space dimension with the additional properties that 1) all users are linearly dependent and 2) if one user were removed from this system, then the remaining users would all be linearly independent. In order to avoid cumbersome notation, we will take the signal space dimension to be N (instead of $|J|$) and hence the number of users to be $N+1$ and the signal matrix to be $\mathbf{S}$ (instead of $\tilde{\mathbf{S}}_{\exp}$) which is now $N \times (N+1)$ and the amplitude, energy matrices and symbol vector in the expurgated system to be $\mathbf{A}$, $\mathbf{E}$ and $\mathbf{x}$, respectively. The abuse of notation is slight since the strategy is to prove the result of Theorem 1 for particular overloaded systems which have the parameters and signal space properties of the above mentioned expurgated channel and then to use the fact that such a particular overloaded system can always be constructed by the process of expurgation from any overloaded system and then to use the lemma above to complete the proof of Theorem 1 for general overloaded systems.

The last simplification is as follows. Since the system under consideration has only linearly dependent users, the AEE for each user is given by (4) (which holds for arbitrary modulation) and can be upper bounded by using the fact that $|\cos\theta| + |\sin\theta| \geq 1$. Specifically, if $\mathbf{s}_k \in \operatorname{span}(\mathbf{S}_k)$, then we have (7) shown at the bottom of the page, where $\|\cdot\|_1$ denotes the ℓ_1-norm, i.e., $\|\mathbf{z}\|_1 = \sum_i |z_i|$. Here, we have recognized that the submatrix $\mathbf{S}_k$ (and hence $\mathbf{S}_k\mathbf{A}_k$) is invertible. From this upper bound, we have that the symmetric energy of $\mathbf{S}$ is upper bounded by $\min_k \mathrm{U}_k$. Therefore, to prove that $e(\mathbf{S}) = 0$, we need to find at least one user for which the upper bound on its AEE in (7) is nonpositive, or equivalently, for which the numerator of the ratio in the square brackets its upper bound is nonpositive.

The two terms of that numerator of the upper bound (7) can both be expressed in terms of some N-dimensional vectors $\{\mathbf{y}_k\}$. Let $\mathbf{T} = \mathbf{SA} = [\mathbf{t}_1, \ldots, \mathbf{t}_K]$ where $K = N+1$. Let $\mathbf{T}_k$ be the $N \times N$ matrix constructed by removing the kth column of $\mathbf{T}$. By assumption, the remaining N columns are linearly independent and therefore $\mathbf{T}_k$ is nonsingular. Set $\mathbf{y}_k = \mathbf{T}_k^{-1}\mathbf{t}_k$. This vector consists of the coordinates of $\mathbf{t}_k$ along the interfering vectors in $\mathbf{T}_k$. The two numerator terms in (7) become

$$\begin{aligned}\mathbf{s}_k^{\mathrm{H}}\left(\mathbf{S}_k\mathbf{E}_k\mathbf{S}_k^{\mathrm{H}}\right)^{-1}\mathbf{s}_k &= \mathbf{s}_k^{\mathrm{H}}(\mathbf{S}_k\mathbf{A}_k)^{+\mathrm{H}}(\mathbf{S}_k\mathbf{A}_k)^{+}\mathbf{s}_k \\ &= E_k^{-1}\|\mathbf{y}_k\|_2^2\end{aligned}$$

and

$$\begin{aligned}\|(\mathbf{S}_k\mathbf{A}_k)^{-1}\mathbf{s}_k\|_1 &= E_k^{-1/2}\|\mathbf{y}_k\|_1.\end{aligned}$$

Thus, the numerator is proportional to $\|\mathbf{y}_k\|_2^2 - \|\mathbf{y}_k\|_1$, and to prove the theorem we need only show that $\min_k(\|\mathbf{y}_k\|_2^2 - \|\mathbf{y}_k\|_1) \leq 0$.

Since the matrix $\mathbf{T}$ is of dimension $N \times N+1$ and has rank N, its null space has dimension one. There is a nonzero vector $\mathbf{a}$ for which $\mathbf{Ta} = 0$. All the components of $\mathbf{a}$ are nonzero for the same reason that the matrices $\mathbf{T}_k$ are all nonsingular. From the equation

$$\mathbf{0} = \mathbf{Ta} = \mathbf{T}_k\mathbf{a}_k + a_k\mathbf{t}_k$$

where $\mathbf{a}_k$ is the result of removing the kth component of $\mathbf{a}$, we can get insight about the vectors $\mathbf{y}_k$, namely

$$\mathbf{y}_k = -\frac{1}{a_k}\mathbf{a}_k.$$

Consider now the quadratic

$$f(\gamma) = \|\mathbf{a}\|_2^2\gamma(\gamma - \gamma_0).$$

where $\gamma_0 = \|\mathbf{a}\|_1/\|\mathbf{a}\|_2^2$. This quadratic has zeros at 0 and γ_0. For $0 < \gamma < \gamma_0$ it is negative. But by evaluating this function at the points $\gamma = 1/|a_k|$, we will recover the terms in which we are most interested

$$\begin{aligned}f\left(\frac{1}{|a_k|}\right) &= \|\mathbf{a}\|_2^2\frac{1}{|a_k|}\left(\frac{1}{|a_k|} - \frac{\|\mathbf{a}\|_1}{\|\mathbf{a}\|_2^2}\right) \\ &= \frac{\|\mathbf{a}\|_2^2}{|a_k|^2} - \frac{\|\mathbf{a}\|_1}{|a_k|} \\ &= \frac{\|\mathbf{a}_k\|_2^2}{|a_k|^2} - \frac{\|\mathbf{a}_k\|_1}{|a_k|} \\ &= \|\mathbf{y}_k\|_2^2 - \|\mathbf{y}_k\|_1. \end{aligned} \tag{8}$$

The third line in this chain is the result of eliminating the common terms at component k. If the components of $\mathbf{a}$ are identical in absolute value, $|a_k| = \alpha$, then $\gamma_0 = 1/|\alpha|$ and $f(\frac{1}{|a_k|}) = f(\frac{1}{|\alpha|}) = 0$. The upper bound is then zero for all users. This is the case when, for instance, the received energies are equal and the users are allocated WBE signals (in this case, the WBE signals form a simplex signal set).[2]

On the other hand, when the components of $\mathbf{a}$ are not identical then

$$|a_\nu| = \min\{|a_k|\} < |a_\mu| = \max\{|a_k|\}.$$

In this case the term of interest for user μ becomes negative

$$\begin{aligned}\|\mathbf{y}_\mu\|_2^2 - \|\mathbf{y}_\mu\|_1 &= f\left(\frac{1}{|a_\mu|}\right) \\ &= \frac{\sum_{n=1}^{N+1}|a_n|(|a_n| - |a_\mu|)}{|a_\mu|^2} < 0.\end{aligned}$$

This proves the claim that the numerator of the expression in the square brackets of the upper bound (7) is negative for at least one user, and hence this user's AEE is zero.

IV. Conclusion

This correspondence highlights a severe limitation of linear mmse detection in the context of overloaded CDMA systems by showing that

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

$$e_k \leq \mathrm{U}_k = E_k\left(\left[\frac{\mathbf{s}_k^{\mathrm{H}}\left(\mathbf{S}_k\mathbf{E}_k\mathbf{S}_k^{\mathrm{H}}\right)^{-1}\mathbf{s}_k - E_k^{-1/2}\|(\mathbf{S}_k\mathbf{A}_k)^{-1}\mathbf{s}_k\|_1}{\sqrt{\mathbf{s}_k^{\mathrm{H}}\left(\mathbf{S}_k\mathbf{E}_k\mathbf{S}_k^{\mathrm{H}}\right)^{-2}\mathbf{s}_k}}\right]_+\right)^2 \tag{7}$$

there do not exist spreading sequences that satisfy even the basic requirement that the error rate of every user decreases exponentially as the noise power vanishes. This result holds for arbitrary overload factor, modulation scheme, and received energies. One consequence is that power control even without a constraint on the total power cannot mitigate this limitation. This strongly suggests that reliable and spectrally efficient multiple access requires a high-performance receiver structure based on nonlinear detection. Such a structure is in fact better suited to uplink transmission and it need not necessarily entail a prohibitive increase in receiver complexity. As a case in point, systematic constructions of bandwidth-efficient signal sets have been proposed under mmse decision feedback detection that provably achieve arbitrary user-specified rate-tuples (with or without power control) [1], [6] or AEE's [11].

References

[1] M. K. Varanasi and T. Guess, "Bandwidth-efficient multiple-access (BEMA): A new strategy based on signal design for multiuser receivers under quality-of-service constraints," *IEEE Trans. Commun.*, vol. 49, no. 5, pp. 844–854, May 2001.

[2] Z. Xie, R. T. Short, and C. T. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 4, pp. 683–690, May 1990.

[3] U. Madhow and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.

[4] A. Kapur and M. K. Varanasi, "Multiuser detection for overloaded CDMA systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1728–1742, July 2003.

[5] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE receivers," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1968–1983, Sep. 1999.

[6] T. Guess, "Optimal sequences for CDMA with decision-feedback receivers," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 886–900, Apr. 2003.

[7] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[8] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[9] M. K. Varanasi, "Decision feedback multiuser detection: A systematic approach," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 219–240, Jan. 1999.

[10] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.

[11] A. Kapur and M. K. Varanasi, "Signal design for bandwidth efficient multiple access with guaranteed bit error rate," in *Proc. IEEE Int. Conf. Commun.*, Anchorage, AK, May 2003.

[12] J. L. Massey and T. Mittelholzer, "Welch's bound and sequence sets for code-division and multiple-access systems," in *Sequences II, Methods in Communication, Security, and Computer Science*. New York: Springer-Verlag, 1993.

[13] M. K. V. A. Kapur and C. Mullis, "On the limitation of generalized Welch-bound equality signals," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2220–2224, Jun. 2005.

Optimal Power/Rate Allocation and Code Selection for Iterative Joint Detection of Coded Random CDMA

Christian Schlegel, *Senior Member, IEEE*, Zhenning Shi, *Member, IEEE*, and Marat Burnashev

***Abstract*—Iterative interference cancellation of coded code-division multiple access (CDMA) using random spreading with linear cancellation is analyzed. If users are grouped into power classes and Shannon bound approaching codes are used, a geometric power distribution achieves the additive white Gaussian noise (AWGN) channel Shannon bound as the numbers of classes becomes large. The optimal distribution of the size of these classes is shown to be uniform. If users are grouped into different rate classes with equal powers among equal rate users, the Shannon bound for AWGN channels can be achieved with an arbitrary distribution of the classes sizes, provided that the size of the largest rate class obeys the mild condition that its ratio of size to processing gain is much smaller than the inverse of the signal-to-noise ratio (SNR). The case of equal powers and equal rates among all users is addressed as a "worst case" scenario. It is argued that simple repetition codes provide for a larger achievable capacity than stronger codes. It is shown that this capacity monotonically increases as the rate of the code decreases. A density evolution analysis is used to show that the achievable rates exceed those of a minimum-mean square error filter applied to the uncoded signals. This lower bound is tight for small ratios of bit energy to noise power, and otherwise the iterative cancellation receiver provides an appreciably larger capacity. Relating to recent result from the application of statistical mechanics it is shown that the repetition-coded system with iterative cancellation achieves the performance of an equivalent optimal joint detector for uncoded transmission.**

***Index Terms*—Iterative decoding, joint detection, optimal power, optimal rate, random code-division multiple access (CDMA).**

I. Introduction

Iterative joint decoding of code-division multiple-access (CDMA) systems using forward error control coding is based on the success of turbo coding [3], and has the potential of realizing a significant portion of the multiple-access channel capacity [1], [21], [16], [2], [24], [37]. Iterative decoding breaks the complex task of a multiuser decoder into two operations, viz. 1) *a posteriori* probability (APP) estimation of the coded symbols, or an approximation thereof, and 2), parallel soft decoding of single-user forward error control (FEC) codes.

The asymptotic analysis of large-scale CDMA systems has been studied for optimal and some suboptimal multiuser detectors [33], [32], and [28], where it is shown that the output SINRs, approach certain constant values if random spreading is used. The convergence behavior of an iterative decoder can be studied using density evolution (DE) analysis [22], [4]. In [7] and [6] different power levels and a linear programming solution to optimizing these power levels are presented.

Manuscript received May 3, 2004; revised April 20, 2006. This work was supported in part by NSF under Grant CCR 9732962 and iCORE Alberta, and by Grant 03-01-00098 from the Russian Fund for Fundamental Research. The material in this correspondence was presented in part at Zurich Seminar on Digital Communications, Zurich, Switzerland, February 2004.

C. Schlegel is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: schlegel@ece.ualberta.ca).

Z. Shi is with National ICT Australia, and affiliated with the Australian National University, Canberra, ACT, Australia (e-mail: zhenning.shi@nicta.com.au).

M. Burnashev is with the Institute for Problems of Information Transmission, Moscow 101447, Russia (e-mail: burn@ee.ucr.edu).

Communicated by X. Wang, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2006.880015