Group-Metric Multiuser Decoding

Eric A. Fain, *Member, IEEE*, and Mahesh K. Varanasi, *Senior Member, IEEE*

***Abstract*—We propose the new group metric (GM) soft-decision decoder for convolutionally coded synchronous multiple-access channels. The GM decoder exploits the independently operating encoders of the multiuser channel by making decoding decisions for a subset of the users, but incorporating all the multiuser information in its metrics. For a single user, this decoder will have a reduced complexity that is exponential in the *sum* of encoder memory and the number of users. The soft-decision maximum-likelihood (ML) joint decoder is well known. This optimal decoder suffers from a high complexity requirement that is exponential in the *product* of encoder memory and the number of users. The size of the decoded subset is a design parameter which allows a tradeoff between complexity and performance.**

The performance of the GM decoder, once properly characterized, can be analyzed using standard techniques. In addition, a new analysis technique is presented which considers decomposable sequences for the fading channel. With this analysis, we have a new tool for bounding error probabilities for multiuser decoders. Applying this technique to the GM decoder, we can directly identify sequences that are decomposable some fraction of the time, and obtain a new upper bound. Further, this improved bound can be expressed in closed form. Numerical results show that the actual performance gap between the GM and ML decoders can be quite small.

***Index Terms*—Diversity methods, fading channels, multiuser channels, reduced-state decoding, transfer function bounds.**

I. Introduction

WE CONSIDER a correlated-waveform multiple-access (CWMA) channel where each user employs a single-user convolutional code. The maximum-likelihood (ML) joint decoder is easily specified and analyzed for this channel, but it has the disadvantage of requiring a computational complexity that is exponential in the product of the number of users and the encoder memory of the codes.

The group metric (GM) decoder is a new reduced-complexity alternative to the ML joint decoder and was first presented in [1] and [2]. This new scheme has a complexity that is exponential in the *sum* of encoder memory and the number of users. This complexity reduction is realized by assuming that only a subset of the users' codes are known. This approach is therefore particular to the multiuser channel. While less complex than the optimal decoder, it is naturally more complex than a single-user decoder, unless there is only one user.

Paper approved by G. Cherubini, the Editor for CDMA Systems of the IEEE Communications Society. Manuscript received October 26, 1999; revised March 27, 2000 and August 30, 2000. This work was supported in part by the National Science Foundation under Grant NCR-9406069 and Grant NCR-9725778. This work was presented in part at the IEEE Global Communications Conference, Phoenix, AZ, 1997, and at the IEEE International Symposium on Information Theory, Cambridge, MA, 1998.

The authors are with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: varanasi@schof.colorado.edu).

Publisher Item Identifier S 0090-6778(01)04876-0.

The GM decoder is developed in the context of a Rayleigh fading channel, although, in principle, it could be applied to any coded multiuser channel. The reason for this choice is the well-known fact that the soft-decision coding gain on fading channels is typically much larger than for Gaussian nonfading channels, and so more benefit is derived from using the GM strategy.

Standard error probability bounds for convolutional decoders can be applied to estimate the GM decoder's performance. To tighten this bound, a new bounding technique is developed. This bound extends the idea of decomposable sequences [3] to the fading channel. Previous work in this area did not achieve any tightening of the standard bound [4]. However, by introducing conditionally decomposable error sequences (conditioning a sequence's decomposability on the fading realization), we are able to improve upon the classic bounding technique. Further, for the GM decoder, this improved bound can be expressed in a easily computable form, in contrast with the technique of [3]. Conditionally decomposable sequences were suggested in [5] in the context of bit-interleaved trellis-coded modulation, but closed-form results were not obtained.

This paper is organized as follows. First, we summarize convolutional coding for the multiuser channel in Section II. We describe our general fading channel model in Section III. Section IV describes both the ML and GM decoders for our channel. Section V states the upper bound for the ML decoder, and Section VI develops a simple transfer function bound on the error rate of the GM decoder. Section VII develops the conditionally decomposable sequence bound. Lastly, Section IX presents simulation results for the decoders, and a summary is given in Section X.

II. Multiuser Convolutional Super-Encoder

In the multiuser channel, we consider the situation where each user employs a single-user (n, k, ν) convolutional code. For the purposes of joint decoding, it will be convenient to consider all the users to be participating in a single "super-code," as in [6]. To this end, we first define a single-user convolutional code.

We begin with some notation: an (n, k, ν) convolutional encoder takes k input bits (in $\{0, 1\}$) at a time, and outputs n encoded bits, typically sequentially, so that the information rate is k/n times the output encoded bit rate. ν denotes the memory length of the encoder. Let user m have a vector[1] of kI information bits $\mathbf{x}^{(m)}$ so that $\mathbf{x}^{(m)^\top} = [\mathbf{x}^{(m)^\top}(0) \cdots \mathbf{x}^{(m)^\top}(I-1)]$, where the k-length subvector $\mathbf{x}^{(m)}(i) = [x_0^{(m)}(i) \cdots x_{k-1}^{(m)}(i)]^\top$. The

[1] All vectors are taken to be column vectors, and are denoted with lowercase bold type ($\mathbf{a}$). Uppercase boldface ($\mathbf{A}$) will denote matrices, and caligraphic type ($\mathcal{A}$) will denote sets or spaces.

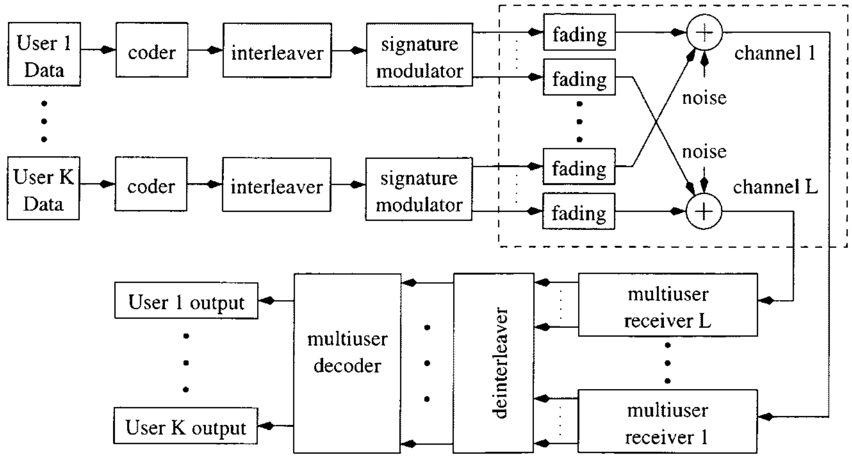

Fig. 1. A multiuser system block diagram with L diversity channels per user.

encoder will produce a vector of nI output bits $\mathbf{y}^{(m)}$ with $\mathbf{y}^{(m)\top} = [\mathbf{y}^{(m)\top}(0) \cdots \mathbf{y}^{(m)\top}(I-1)]$, and the n-length subvector $\mathbf{y}^{(m)}(i) = [y_0^{(m)}(i) \cdots y_{n-1}^{(m)}(i)]^\top$, according to the rule

$$y_l^{(m)}(i) = \sum_{\tau=0}^{k-1} \sum_{t=0}^{\nu} x_\tau^{(m)}(i-t) g_{l,\tau}^{(m)}(t), \qquad 0 \le l \le n-1 \quad (1)$$

where the $g_{l,\tau}^{(m)}(i)$'s are the binary generator sequences of the encoder with maximum length $\nu + 1$, and the sum and product operations are modulo-2.

We now assume that each of K users has an (n, k, ν) code. The multiuser super-code will, in its simplest interpretation, be a single (nK, kK, ν) convolutional code. This fact makes the joint decoding problem straightforward, albeit complex. The structure of the super-code is, of course, very special, and we describe it in terms of the single-user code for each user.

The hypothetical super-encoder takes an input vector $\mathbf{x} = [\mathbf{x}^\top(0), \ldots, \mathbf{x}^\top(I-1)]^\top$, composed of all K users' bits, where

$$\mathbf{x}(i) \triangleq \Big[x_0^{(1)}(i)\, x_0^{(2)}(i) \cdots x_0^{(K)}(i) \cdots x_{k-1}^{(1)}(i) \cdots x_{k-1}^{(K)}(i)\Big]^\top. \quad (2)$$

The outputs $\mathbf{y}$ are similarly $\mathbf{y} = [\mathbf{y}^\top(0), \ldots, \mathbf{y}^\top(I-1)]^\top$ with nK-dimensional subvector $\mathbf{y}(i)$

$$\mathbf{y}(i) \triangleq \Big[y_0^{(1)}(i)\, y_0^{(2)}(i) \cdots y_0^{(K)}(i) \cdots y_{n-1}^{(1)}(i) \cdots y_{n-1}^{(K)}(i)\Big]^\top \quad (3)$$

$$\triangleq \big[\mathbf{y}_0^\top(i) \cdots \mathbf{y}_{n-1}^\top(i)\big]^\top. \quad (4)$$

We let the $(nK \times kK)$ matrix

$$\mathbf{G}(j) = \begin{bmatrix} \mathbf{G}_{0,0}(j) & \cdots & \mathbf{G}_{0,k-1}(j) \\ \vdots & \ddots & \vdots \\ \mathbf{G}_{n-1,0}(j) & \cdots & \mathbf{G}_{n-1,k-1}(j) \end{bmatrix} \quad (5)$$

where each submatrix $\mathbf{G}_{l,\tau}(j)$ is a $K \times K$ *diagonal* matrix $\mathbf{G}_{l,\tau}(j) = \mathrm{diag}\{g_{l,\tau}^{(1)}(j) \cdots g_{l,\tau}^{(K)}(j)\}$. This allows us to write

$$\mathbf{y}(i) = \sum_{j=0}^{\nu} \mathbf{G}(j)\mathbf{x}(i-j). \quad (6)$$

Thus, the entire sequence is given by $\mathbf{y} = \mathbf{G}\mathbf{x}$, where $\mathbf{G}$ is block-Toeplitz with i, jth block $\mathbf{G}(i-j)$.

III. Discrete-Time Fading Diversity Channel

We assume a channel that is slowly fading, so that perfect channel estimates are available at the receiver, and the fading amplitudes are constant over an entire symbol interval. We also assume that ideal interleaving is used, so that the fading amplitudes are essentially independent from one symbol interval to the next.

We must first characterize the transmitted symbols. Here, we shall continue with the notion of a super-encoder, which takes multiuser input vector $\mathbf{x}$ and produces $\mathbf{y}$, with elements in $\{0, 1\}$. For binary phase-shift keying (BPSK) modulation, each $\mathbf{y}(i)$ will be transmitted over the space of n time-intervals, so we define $\mathbf{b}_l(i) \triangleq 2\mathbf{y}_l(i) - \mathbf{1}$ so that $\mathbf{y}(i) = 1/2[\mathbf{b}_0^\top(i) \cdots \mathbf{b}_{n-1}^\top(i)]^\top + (1/2)\mathbf{1}$, where $\mathbf{1}$ denotes a vector of ones of appropriate length. Each K-dimensional vector $\mathbf{b}_l(i)$, with elements in $\{\pm 1\}$, represents the coded symbols transmitted in each time-interval (in which the fading is considered constant) by all K users.

We consider a synchronous multiple-access system with each user having L diversity channels. A system block-diagram is shown in Fig. 1. Each user digitally modulates a fixed signature signal in each channel in each symbol interval. This results in an equivalent discrete-time channel, for the ith symbol interval given as

$$\mathbf{q}_l(i) = \mathbf{F}\mathbf{C}_l(i)\mathbf{b}_l(i) + \mathbf{n}_l(i) \quad (7)$$

where $\mathbf{q}_l(i)$ has length $\rho(\le LK)$, $\mathbf{C}_l(i)$ is a $LK \times K$ matrix of complex fading amplitudes

$$\mathbf{C}_l(i) = \begin{bmatrix} \mathbf{c}_l^{(1)}(i) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{c}_l^{(K)}(i) \end{bmatrix} \quad (8)$$

with $\mathbf{c}_l^{(m)}(i)$ an L length vector of fading amplitudes for user m. The vector $\mathbf{n}_l(i)$ is complex Gaussian noise with covariance $N_0\mathbf{I}$. The $\rho \times LK$ matrix $\mathbf{F}$ is equal to $[\mathbf{F}_1 \cdots \mathbf{F}_K]$, where the L columns of $\mathbf{F}_m$ are the vector representations of the signals of user m in each of the diversity channels, with respect to a ρ-dimensional orthonormal basis for the signal space. Thus, $\mathbf{F}^*\mathbf{F} = \mathbf{R}$ is the rank ρ correlation matrix of the users' signature waveforms in each diversity channel, where $*$ denotes conjugate transpose. (For example, for a direct-sequence CDMA system with $L = 1$, each $\mathbf{F}_m$ could be a vector of the user's spreading sequence, with $\rho = K$ basis functions each corresponding to one chip. In this case, the (i, j)th element of $\mathbf{R}$ would be the correlation of the spreading sequences of users i and j. In a multipath channel, L would represent the number of multipaths, and each column of $\mathbf{F}_m$ would be the user's spreading sequence, shifted by the path delay.) A typical spread-spectrum signal would have $\rho = LK$ for a full-rank correlation matrix, but this restriction is not necessary for our model.

In the case of Rayleigh fading, the $\mathbf{c}_l^{(m)}(i)$'s are zero-mean complex normal vectors with covariance $\mathbf{\Sigma}_{mm}$. Under the assumption that each user fades independently, $\mathbf{C}_l(i)$ will be completely described by the block-diagonal channel envelope correlation matrix

$$\mathbf{\Sigma} \triangleq E(\mathbf{C}_l(i)\mathbf{C}_l^*(i)) = \begin{bmatrix} \mathbf{\Sigma}_{11} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{\Sigma}_{KK} \end{bmatrix}. \tag{9}$$

The mth user's average received energy is given by

$$E\left(\mathbf{c}_l^{(m)*}(i) * \mathbf{F}_m^*\mathbf{F}_m\mathbf{c}_l^{(m)}(i)\right) = \mathrm{tr}(\mathbf{R}_{mm}\mathbf{\Sigma}_{mm}) \tag{10}$$

where $\mathbf{R}_{mm}$ is the $(L \times L)$ mth diagonal block of $\mathbf{R}$.

IV. Multiuser Decoder Description

Our new soft-decision decoding scheme decodes the users individually, but uses decoding metrics that are a function of the entire group of users. We call this method GM decoding. For simplicity of description only, we consider the case where every user's code has the same parameters n, k, and ν, although this restriction is not necessary. To make the development clear, we first specify the ML joint decoder, then present the GM decoder. These parallel descriptions will emphasize the similarities and differences between the two decoders. Throughout this paper, we assume that the receiver has perfect knowledge of the signal matrix $\mathbf{F}$ as well as the fading processes.

A. ML Joint Decoder

To specify the ML decoding rule, we define the block-diagonal matrices[2] $\tilde{\mathbf{F}} = \mathrm{diag}\{\mathbf{F}, \ldots, \mathbf{F}\}$ and $\tilde{\mathbf{R}} = \mathrm{diag}\{\mathbf{R}, \ldots, \mathbf{R}\}$, and let $\mathbf{q} = [\mathbf{q}_0^\top(0), \ldots, \mathbf{q}_{n-1}^\top(0), \mathbf{q}_0^\top(1), \ldots, \mathbf{q}_{n-1}^\top(I-1)]^\top$, $\mathbf{b} = [\mathbf{b}_0^\top(0), \ldots, \mathbf{b}_{n-1}^\top(0), \mathbf{b}_0^\top(1), \ldots, \mathbf{b}_{n-1}^\top(I-1)]^\top$, and $\mathbf{C} = \mathrm{diag}\{\mathbf{C}_0(0), \ldots, \mathbf{C}_{n-1}(0), \mathbf{C}_0(1), \ldots, \mathbf{C}_{n-1}(I-1)\}$, where n is the parameter of the single-user codes as in Section II, and nKI is the coded message length.

[2]The tilde (˜) notation is used to indicate a block extension of the single-symbol parameters to message-length parameters.

The ML joint decoding rule for all the users, given that the CFI matrix $\mathbf{C}$ is known for the entire sequence, gives an estimate $\hat{\mathbf{b}}$ for $\mathbf{b}$ according to

$$\hat{\mathbf{b}} = \arg\min_{\beta \in \mathcal{A}} \left\| \tilde{\mathbf{F}}\mathbf{C}\beta - \mathbf{q} \right\|^2 \tag{11}$$

where $\mathcal{A}$ is the set of all possible coded K-user sequences.

This decoding rule can be implemented with the Viterbi algorithm [7], and it is an extension of the decoder in [6] to the synchronous fading diversity channel. For a single-user channel, in each n time-intervals, the Viterbi decoder considers every possible *state* of the encoder. If the stored memory of the encoder is ν bits, the Viterbi decoder will have $2^{k\nu}$ states. Since the multiuser joint decoder can be described as an (nK, kK, ν) convolutional decoder for the multiuser super-code, the Viterbi decoder here has $2^{k\nu K}$ states, 2^{kK} branches to and from every state, with a state transition every n bits.

The likelihood rule (11) is evaluated as a sum of *branch metrics*, which represent the transition for one encoder state to the next. Let $\hat{\mathbf{x}}$ be a candidate path through the decoding trellis, and let state A correspond to the subsection of $\hat{\mathbf{x}}$, $[\hat{\mathbf{x}}(i) \cdots \hat{\mathbf{x}}(i+\nu-1)]$, and state B correspond to $[\hat{\mathbf{x}}(i+1) \cdots \hat{\mathbf{x}}(i+\nu)]$. The branch metric between these two *adjacent* states is given by

$$d_{AB}(i+\nu) = \left\| \tilde{\mathbf{F}}^n\mathbf{C}(i+\nu)\mathbf{p}_{AB} - \mathbf{q}(i+\nu) \right\|^2 \tag{12}$$

where we define $\mathbf{C}(i+\nu) = \mathrm{diag}\{\mathbf{C}_0(i+\nu), \ldots, \mathbf{C}_{n-1}(i+\nu)\}$, $\mathbf{q}(i) = [\mathbf{q}_0^\top(i), \ldots, \mathbf{q}_{n-1}^\top(i)]^\top$, and $\tilde{\mathbf{F}}^n = \mathrm{diag}\{\mathbf{F}, \ldots, \mathbf{F}\}$. The nK-dimensional vector $\mathbf{p}_{AB}$ is formed from the K n-bit words generated by transition from state A to state B, that is, $(1/2)\mathbf{p}_{AB} + (1/2)\mathbf{1} = \hat{\mathbf{y}}(i+\nu) = \sum_{j=0}^{\nu} \mathbf{G}(j)\hat{\mathbf{x}}(i+\nu-j)$, where the operations on the right-hand side are modulo-2.

The computational burden of the decoding algorithm lies in the evaluation of the branch metrics. From (12), we see that each of these branch metrics requires the computation of n LK-dimensional quadratic forms. Thus, counting the number of LK-dimensional quadratic forms evaluated per coded bit, we find the joint decoder to have complexity of order $\mathcal{O}(2^{(\nu+1)kK})$.

B. GM Decoder

The GM decoder will reduce the complexity of the ML decoder by considering only a subset of the users' codes to be known. The other users' code symbols are (falsely) considered to be uncoded, independent, and equiprobable. These assumptions result in assuming a different multiuser super-code, for which the GM decoder would be the optimal decoder.

For clarity, we will consider the case where only one user is decoded for now and postpone the discussion of multiple-user decoding to Section VIII. The GM decoder in this case is a single-user decoding scheme that will use information from all users' matched-filter outputs for the decoding metrics, and will operate in a bank of K parallel units. In this case, the assumed super-encoder produces an output $\dot{\mathbf{y}}$ given an input $\dot{\mathbf{x}}$ as $\dot{\mathbf{y}} = \dot{\mathbf{G}}\dot{\mathbf{x}}$, where (assuming user 1 is the desired user)[3] $\dot{\mathbf{x}} = [\dot{\mathbf{x}}^\top(0) \cdots \dot{\mathbf{x}}^\top(I-1)]^\top$ with

[3]We use the dot (˙) notation to distinguish the matrices of the assumed super-encoder.

$\dot{\mathbf{x}}(i) = [\mathbf{x}_1^\top(i)\ y_0^{(2)}(i) \cdots y_0^{(K)}(i)\ y_1^{(2)}(i) \cdots y_{n-1}^{(K)}(i)]^\top$. Let $\mathbf{g}_i^{(m)}(l) = [g_{i,0}^{(m)}(l) \cdots g_{i,k-1}^{(m)}(l)]^\top$. The $nK \times (k + n(K-1))$ matrix $\dot{\mathbf{G}}(l)$ is defined as

$$\dot{\mathbf{G}}(l) = \begin{bmatrix} \mathbf{g}_0^{(1)^\top}(l) & 0 & \cdots & \cdots & \cdots & 0 \\ \mathbf{0} & \delta(l)\mathbf{I}_{K-1} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{g}_1^{(1)^\top}(l) & 0 & \cdots & \cdots & \cdots & 0 \\ \mathbf{0} & \mathbf{0} & \delta(l)\mathbf{I}_{K-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & & & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} & \delta(l)\mathbf{I}_{K-1} \\ \mathbf{g}_{n-1}^{(1)^\top}(l) & 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} \tag{13}$$

where $\mathbf{I}_K$ denotes an $K \times K$ identity matrix, and $\delta(l) = 1$, for $l = 0$, and $\delta(l) = 0$ otherwise. The large matrix $\dot{\mathbf{G}}$ is then block-Toeplitz with i, jth block $\dot{\mathbf{G}}(i-j)$. From the structure in the $\dot{\mathbf{G}}(l)$'s, we see that the for users $2, \ldots, K$, the coded outputs are passed as inputs to the assumed super-encoder. So by construction, we have that $\dot{\mathbf{y}} = \mathbf{y}$; the assumed coder's outputs are identical to the true super-coder's outputs. In the decoding stage, we will assume that these coded symbols from users $2, \ldots, K$ are unknown, equiprobable random quantities.

We additionally define the $k + n(K-1)$ length vectors $\dot{\mathbf{x}}^m(i) \triangleq [\mathbf{0}\ \ \mathbf{x}^{(m)^\top}(i)\ \ \mathbf{0}]^\top$, and $\dot{\mathbf{x}}^{\overline{m}}(i) \triangleq \dot{\mathbf{x}}(i) - \dot{\mathbf{x}}^m(i)$, with corresponding $\dot{\mathbf{x}}^m = [\dot{\mathbf{x}}^{m^\top}(0) \cdots \dot{\mathbf{x}}^{m^\top}(I-1)]^\top$, and $\dot{\mathbf{x}}^{\overline{m}} = [\dot{\mathbf{x}}^{\overline{m}^\top}(0) \cdots \dot{\mathbf{x}}^{\overline{m}^\top}(I-1)]^\top$ so that

$$\mathbf{y} = \dot{\mathbf{G}}\dot{\mathbf{x}} = \dot{\mathbf{G}}\dot{\mathbf{x}}^m + \dot{\mathbf{x}}^{\overline{m}}. \tag{14}$$

We can now state the group-metric decoding rule: under the assumption that only the mth user's code is known, the conditionally ML estimate of $\mathbf{b}$ is

$$\hat{\mathbf{b}} = \arg\min_{\beta \in \mathcal{A}_m} \left\| \tilde{\mathbf{F}}\mathbf{C}\beta - \mathbf{q} \right\|^2 \tag{15}$$

$$= \arg\min_{\beta^m \in \mathcal{X}_m} \left[\min_{\beta^{\overline{m}} \in \mathcal{X}_{\overline{m}}} \left\| \tilde{\mathbf{F}}\mathbf{C}\left(\beta^m + \beta^{\overline{m}}\right) - \mathbf{q} \right\|^2 \right] \tag{16}$$

where we replace $\mathcal{A}$ in (11) with $\mathcal{A}_m$, the set of all possible transmitted sequences assuming only the mth user is coded while the rest are uncoded, and further decompose this set by letting $\mathcal{X}_m$ be the set of all vectors that are coded outputs for user m and are zero in all other positions (i.e., for $m = 1$, if $\hat{\mathbf{y}} = \dot{\mathbf{G}}[\hat{\mathbf{x}}^{(1)^\top}(0)\ \mathbf{0}^\top\ \hat{\mathbf{x}}^{(1)^\top}(1)\ \mathbf{0}^\top\ \cdots]^\top$, then $2\hat{\mathbf{y}} - 1 \in \mathcal{X}_1$) , and let $\mathcal{X}_{\overline{m}}$ be the set of all vectors in $\{\pm 1\}$ with zeros in the positions for user m. Of course, we retain only those elements of $\hat{\mathbf{b}}$ corresponding to user m.

With this scheme, a Viterbi decoder for the mth user can be implemented, as in a single-user channel, with $2^{k\nu}$ states, and 2^k branches to and from each state. The branch metrics, however, will be unusual. Let $\hat{\mathbf{x}}$ be a candidate path through the decoding trellis, and, since the state information depends only on user m, let state A correspond to $[\hat{\mathbf{x}}^{(m)}(i) \cdots \hat{\mathbf{x}}^{(m)}(i+\nu-1)]$, and state B correspond to $[\hat{\mathbf{x}}^{(m)}(i+1) \cdots \hat{\mathbf{x}}^{(m)}(i+\nu)]$. The branch metric between these two *adjacent* states is given by

$$d_{AB}(i+\nu) = \min_{\mathbf{p}_{AB}^{\overline{m}}} \left\| \tilde{\mathbf{F}}^n \mathbf{C}(i+\nu)\left(\mathbf{p}_{AB}^m + \mathbf{p}_{AB}^{\overline{m}}\right) - \mathbf{q}(i+\nu) \right\|^2. \tag{17}$$

The vector $\mathbf{p}_{AB}^m$ makes up the mth user's encoded n-bit word generated from the transition to state A from state B, that is

$$\begin{aligned} \frac{1}{2}\mathbf{p}_{AB}^m + \frac{1}{2}\mathbf{1} &= \left[0\ \cdots\ y_0^{(m)}(i+\nu)\ 0\ \cdots\ 0\ y_{n-1}^{(m)}(i+\nu)\ \cdots\ 0\right]^\top \\ &= \sum_{j=0}^{\nu} \dot{\mathbf{G}}(j)\dot{\mathbf{x}}^m(i+\nu-j). \end{aligned} \tag{18}$$

So $\mathbf{p}_{AB}^m(i) = 0$ for $i \neq \{m, 2m, \ldots, nm\}$. The vector $\mathbf{p}_{AB}^{\overline{m}}$ represents the unconstrained interfering users, and we have $\mathbf{p}_{AB}^{\overline{m}}(i) \in \{\pm 1\}$ for $i \neq \{m, 2m, \ldots, nm\}$, and $\mathbf{p}_{AB}^{\overline{m}}(i) = 0$ for $i \in \{m, 2m, \ldots, nm\}$. Due to the block diagonal structure of the matrix $\tilde{\mathbf{F}}^n$, each branch metric requires a minimization over only $n2^{K-1}$ (as opposed to $2^{n(K-1)}$) possibilities. In order to compare with the ML decoder, we count the number of LK-dimensional quadratic forms evaluated per coded bit. We find a bank of K GM-decoders has a complexity of order $\mathcal{O}(K2^{(\nu+1)k+K-1})$ as follows. There are $2^{k\nu}$ states, 2^k branches per state, $n2^{K-1}$ computations per branch, for each n coded bits, and there is one decoder for each of K users. Note that this complexity is generally much less than that for the ML-decoder.

Finally, we note that while the GM decoder is most clearly presented with the assumption that only a single-user is coded, the strategy can easily be extended to consider a larger subset of the users to be coded.

V. ML Decoder Performance Analysis

The ML decoder is best analyzed by considering the effective super-coder. Fig. 2 shows a two-user example of such a super-coder.

Once the effective super-coder is identified, single-user convolutional code performance techniques can be readily applied.

Since the super-encoder's inputs are binary, it can also be viewed as a finite-state machine, where the past ν blocks of inputs can be thought of as the "stored memory" of the encoder. The state of the encoder is the value of this stored memory. We define the *zero state* to be simply the encoder state where all stored bits are zeros.

The *transfer function* [8] (alternatively, *generating function* [7]) of a convolutional encoder is a rational function of several variables which gives the number of possible coded output sequences which depart the zero state exactly once and have the same input and output Hamming weights. In other words, the series expansion for the transfer function can be viewed as a polynomial with terms of the form $\alpha X^d Y^i$, where α represents the number of codewords in the code with weight d, whose information sequence has weight i.

In principle, it is possible to obtain the closed-form transfer function of the super-code analytically, and compute the standard transfer function upper bound [8], [9]. We describe the bounding procedure as follows. First, we define the error sequence $\mathbf{e} = (\mathbf{b}_2 - \mathbf{b}_1)/2$, where $\mathbf{b}_1$ and $\mathbf{b}_2$ are any two distinct possible transmitted sequences. (It will also be convenient to write $\mathbf{e}^\top = [\mathbf{e}^\top(0), \ldots, \mathbf{e}^\top(n(I-1))]^\top$, where $\mathbf{e}(i)$ is a length K vector.)

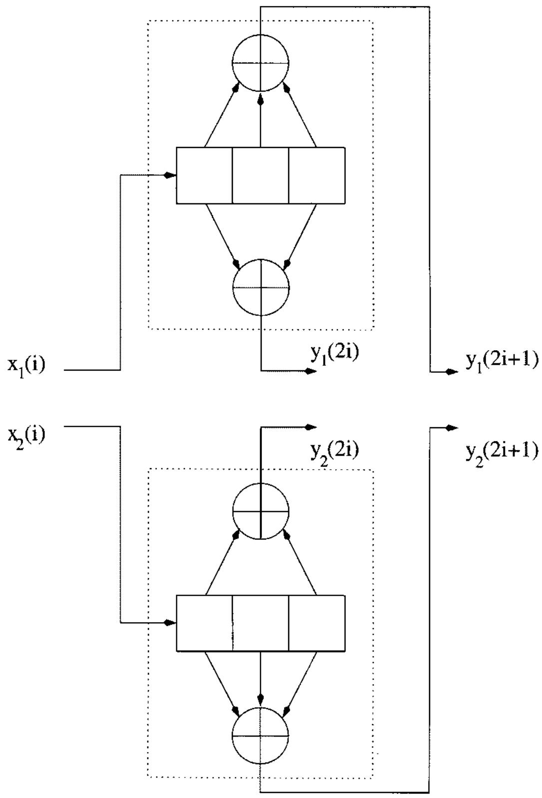

Fig. 2. A two-user example of an equivalent (4, 2, 2) super-coder. The inputs x_1 and x_2 correspond to users 1 and 2, respectively. These also represent the input bits to the super-encoder as in Section II. Dotted lines delineate the two (2, 1, 2) single-user encoders, which are represented as shift registers connected to modulo-2 adders.

We define $\mathcal{P}(\mathbf{e})$ as the *pairwise event* that an erroneous sequence $\mathbf{b}+2\mathbf{e}$ will be chosen over the correct sequence $\mathbf{b}$. The pairwise event will occur if

$$\mathbf{e}^\top \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e} < \mathrm{Re}\left\{\mathbf{n}^* \tilde{\mathbf{F}} \mathbf{C} \mathbf{e}\right\}. \tag{19}$$

In order to characterize the *pairwise probability* (probability of a pairwise event), we introduce a few definitions. Let $\mathbf{\Omega} = \mathbf{\Sigma}^{1/2}\mathbf{R}\mathbf{\Sigma}^{1/2}$ and $\tilde{\mathbf{\Omega}} = \mathrm{diag}\{\mathbf{\Omega}, \ldots, \mathbf{\Omega}\}$. We define the matrix $\mathbf{A}(\mathbf{e})$ as $\mathbf{A}(\mathbf{e}) \triangleq \mathbf{E}^*\tilde{\mathbf{\Omega}}\mathbf{E}$, where $\mathbf{E}$ is diagonal given by $\mathrm{diag}(\mathbf{e}) \otimes \mathbf{I}_L$, and $\otimes$ denotes Kronecker product, so that $\mathbf{e}^*\mathbf{C}^*\tilde{\mathbf{R}}\mathbf{C}\mathbf{e} = \mathbf{a}^*\mathbf{A}(\mathbf{e})\mathbf{a}$, where $\mathbf{a}$ is a zero-mean, circularly symmetric, complex Gaussian vector of whitened fading parameters with $E(\mathbf{a}\mathbf{a}^*) = \mathbf{I}$, and we have assumed $\mathbf{\Sigma}$ to be block diagonal. We also let $\mathbf{A}_i(\mathbf{e})$ be the ith diagonal block of $\mathbf{A}(\mathbf{e})$, and alternately write $\mathbf{A}_i(\mathbf{e}) \triangleq \mathbf{A}(\mathbf{e}(i)) \triangleq \mathbf{E}^*(i)\mathbf{\Omega}\mathbf{E}(i)$, where $\mathbf{E}(i)$ is $(KL \times KL)$.

We can now write the probability of (19), conditioned on the fading, as

$$\Pr(\mathcal{P}(\mathbf{e})|\mathbf{C}) = P_2(\mathbf{e}|\mathbf{C}) = Q\left(\sqrt{\frac{2}{N_0}\mathbf{a}^*\mathbf{A}(\mathbf{e})\mathbf{a}}\right). \tag{20}$$

The quantity $\mathbf{a}^*\mathbf{A}(\mathbf{e})\mathbf{a}$ has characteristic function [10]

$$\psi(s) = \prod_{i=1}^{P}(1-s\lambda_i)^{-m_i} = \sum_{i=1}^{P}\sum_{j=1}^{m_i}\beta_{i,j}(1-s\lambda_i)^{-j} \tag{21}$$

where $\mathbf{A}(\mathbf{e})$ has nonzero eigenvalues $\{\lambda_i\}_{i=1}^P$, each with multiplicity m_i. Taking the expectation of (20) as in [8] yields the pairwise probability of $\mathbf{e}$, $\Pr(\mathcal{P}(\mathbf{e}))$ [denoted $P_2(\mathbf{e})$]

$$P_2(\mathbf{e}) = \sum_{i=1}^{P}\sum_{j=1}^{m_i}\beta_{i,j}\left[\frac{1}{2}(1-\mu_i)\right]^j \cdot \sum_{l=0}^{j-1}\binom{j-1+l}{l}\left[\frac{1}{2}(1+\mu_i)\right]^l \tag{22}$$

where $\mu_i = \sqrt{\lambda_i/(N_0+\lambda_i)}$.

Note that this result does not depend on the actual order of the $\mathbf{e}(i)$'s in the error sequence, due to the perfect interleaving assumption. Also since $\mathbf{A}(\mathbf{e})$ has the same eigenvalues as $\mathbf{A}(|\mathbf{e}|)$, where $|\cdot|$ denotes element-wise absolute value, so that $P_2(\mathbf{e}) = P_2(|\mathbf{e}|)$. For simplicity, therefore, we may assume that the "all -1s" sequence is transmitted, so that every error sequence will have elements only in $\{0, 1\}$. These error sequences are exactly those binary sequences enumerated by the code generating function. By the modulo-2 linearity of the convolutional code, these sequences are valid for any transmitted sequence [9]. Thus, we can consider error sequences independently of the transmitted sequence.

The upper bound on error probability can now be stated in terms of the pairwise probabilities as in [7] and [9]. Let $\mathcal{E}$ denote the set of all multiuser error sequences enumerated by the transfer function, and define $\mathcal{E}_d = \{\mathbf{e} \in \mathcal{E}\colon w(\mathbf{e}) = d\}$, where $w(\mathbf{e})$ denotes the Hamming weight of $\mathbf{e}$. In the limit as $I \to \infty$, the probability that one of the mth user's information bits is in error (P_e^m) is bounded by

$$P_e^m < \frac{1}{k}\sum_{d=d_{\mathrm{free}}}^{\infty}\sum_{\mathbf{e}\in\mathcal{E}_d} N_m(\mathbf{e})P_2(\mathbf{e}) \tag{23}$$

where $N_m(\mathbf{e})$ is the number of information bit errors affecting the mth user in $\mathbf{e}$. An approximate upper bound is obtained by truncating the above summation to a large value of d to include sufficiently many error sequences.

VI. GM Decoder Performance Analysis

The performance analysis for the GM decoder is developed in two stages. First, in Section VI-A, we consider the standard upper bound, which applies the well-known techniques for a general convolutional code, as in Section V. Additionally, in Section VI-B, we present the "simple sequence" bound using the concepts of decomposable sequences to lay the groundwork for the second stage of the analysis. This second stage (Section VII) extends the idea of decomposable sequences to the fading channel.

A. Standard Upper Bound

Like the ML decoder, the GM decoder can also be analyzed by considering the assumed super-encoder. Of course, there will be far more possible error sequences for the GM decoder than for the ML decoder since the interfering users' symbols are assumed to be unconstrained. Fig. 3 depicts the assumed super-encoder for the encoder shown in Fig. 2.

Consider the GM decoder for the mth user, and let $\mathcal{E}$ be the set of error sequences enumerated by the transfer function of

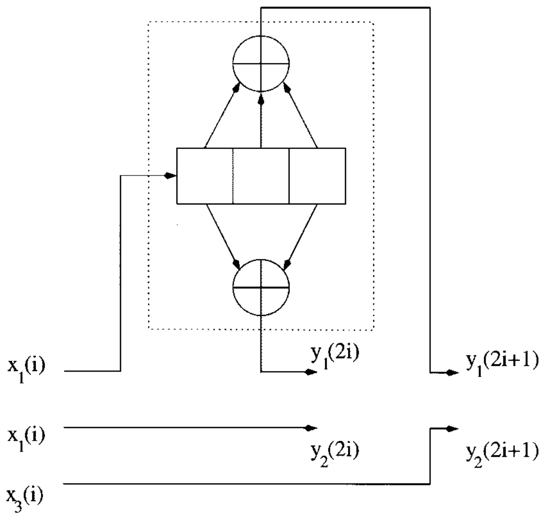

Fig. 3. A two-user example of the (4, 3, 2) assumed super-coder for the group-metric decoder. The inputs $x_2(i)$ and $x_3(i)$ are actually the encoded bits of user 2.

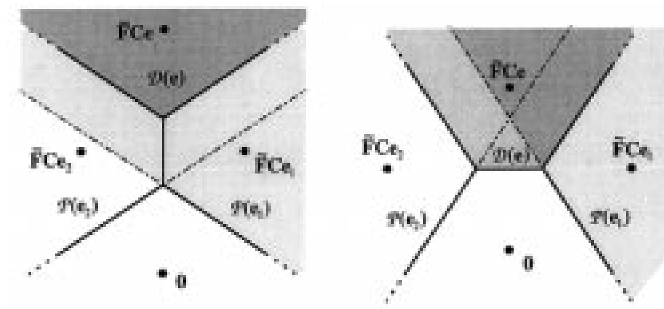

Fig. 4. A geometrical interpretation of decomposable sequences. Here, $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$. The left-hand diagram shows the decision regions for the decomposable case, the right-hand side shows the nondecomposable case.

the assumed super-encoder. We also define $\mathcal{S}^m$ to be the set of all single-user error sequences for user m (obtained from the transfer function of the mth user's single-user encoder). Given a single-user error sequence $\mathbf{s}$, we define the set of *associated* error sequences as

$$\mathcal{E}(\mathbf{s}) = \{\mathbf{e} \in \mathcal{E}\colon m\text{th subsequence of } \mathbf{e} = \mathbf{s}\} \tag{24}$$

i.e., the set of all multiuser error sequences which contain the single-user error sequence $\mathbf{s}$ such that for each $\mathbf{e}(i) = [e_1(i) \cdots e_K(i)]^\top$, $e_m(i) = s(i)$. Note that the set $\mathcal{E}$ can be partitioned so that $\mathcal{E} = \bigcup_{\mathbf{s} \in \{\mathcal{S}^m, \mathbf{0}\}} \mathcal{E}(\mathbf{s})$. In this way, the upper bound can be written as

$$P_e^m < \frac{1}{k} \sum_{\mathbf{s} \in \mathcal{S}^m} \sum_{\mathbf{e} \in \mathcal{E}(\mathbf{s})} N_m(\mathbf{e}) P_2(\mathbf{e}) \tag{25}$$

$$< \frac{1}{k} \sum_{d=d_{\text{free}}}^{\infty} \sum_{\mathbf{s} \in \mathcal{S}_d^m} \sum_{\mathbf{e} \in \mathcal{E}(\mathbf{s})} N_m(\mathbf{e}) P_2(\mathbf{e}) \tag{26}$$

where $\mathcal{S}_d^m = \{\mathbf{s} \in \mathcal{S}^m\colon w(\mathbf{s}) = d\}$, since $N_m(\mathbf{e}) = 0$ for $\mathbf{e} \in \mathcal{E}(\mathbf{0})$.

B. Decomposable Sequence Bound

We now wish to tighten the bound given in (26) by reducing the size of the set $\mathcal{E}$. To accomplish this, we will work on each $\mathcal{E}(\mathbf{s})$ separately by reducing the number of associated sequences for each single-user error sequence. We first require the following definitions: we define the matrix $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$ as a generalization of $\mathbf{A}(\mathbf{e})$

$$\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2) \triangleq \frac{1}{2}\left(\mathbf{E}_1^* \tilde{\mathbf{\Omega}} \mathbf{E}_2 + \mathbf{E}_2^* \tilde{\mathbf{\Omega}} \mathbf{E}_1\right) \tag{27}$$

where $\mathbf{E}_1$ is diagonal given by $\text{diag}(\mathbf{e}_1) \otimes \mathbf{I}_L$, and $\mathbf{E}_2 = \text{diag}(\mathbf{e}_2) \otimes \mathbf{I}_L$, so that

$$\frac{1}{2}\left[\mathbf{e}_1^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_2 + \mathbf{e}_2^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_1\right] = \mathbf{a}^* \mathbf{A}(\mathbf{e}_1, \mathbf{e}_2) \mathbf{a} \tag{28}$$

where $\mathbf{a}$ is a zero-mean, circularly symmetric, complex Gaussian vector of whitened fading parameters with $E(\mathbf{a}\mathbf{a}^*) = \mathbf{I}$. We also let $\mathbf{A}_i(\mathbf{e}_1, \mathbf{e}_2)$ be the ith diagonal block of $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$, so that

$$\begin{aligned}\mathbf{A}_i(\mathbf{e}_1, \mathbf{e}_2) &\triangleq \frac{1}{2}(\mathbf{E}_1^*(i) \mathbf{\Omega} \mathbf{E}_2(i) + \mathbf{E}_2^*(i) \mathbf{\Omega} \mathbf{E}_1(i)) \\ &\triangleq \mathbf{A}(\mathbf{e}_1(i), \mathbf{e}_2(i)).\end{aligned} \tag{29}$$

When $\mathbf{e}_1 = \mathbf{e}_2$, we have $\mathbf{A}(\mathbf{e}_1) = \mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$.

Decomposable sequences were first defined for the uncoded Gaussian multiuser channel in [3], and similarly defined for the uncoded multiuser nonselective fading channel in [4]. We extend the definition to any fading channel model with diversity.

Definition 1 (Decomposable Sequence): An error sequence $\mathbf{e} \in \mathcal{E}$ will be *decomposable* into $\mathbf{e}_1$ and $\mathbf{e}_2$ with $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$ and $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{E}$ if there is no cancellation between $\mathbf{e}_1$ and $\mathbf{e}_2$, $(w(\mathbf{e}) = w(\mathbf{e}_1) + w(\mathbf{e}_2))$ and if $\mathbf{a}^* \mathbf{A}(\mathbf{e}_1, \mathbf{e}_2) \mathbf{a} \geq 0$.

Such a decomposition will be denoted as $\mathbf{e} \stackrel{\text{dec}}{=} \mathbf{e}_1 + \mathbf{e}_2$.

For our derivation, we introduce the following. Let $\mathcal{D}(\mathbf{e})$ be the *decoding event* that the decoder returns $\mathbf{b} + 2\mathbf{e}$, where $\mathbf{b}$ is the transmitted sequence; i.e., the decoder makes the error $\mathbf{e}$. Clearly, the decoding event is a subset of the pairwise event for $\mathbf{e}$, i.e., $(\mathcal{D}(\mathbf{e}) \subset \mathcal{P}(\mathbf{e}))$. But the key result for decomposable sequences, due to [3], but formulated for our framework, is as follows.

Proposition 1: If $\mathbf{e} \stackrel{\text{dec}}{=} \mathbf{e}_1 + \mathbf{e}_2$, then

$$\mathcal{D}(\mathbf{e}) \subset \mathcal{P}(\mathbf{e}_1) \quad \text{and} \quad \mathcal{D}(\mathbf{e}) \subset \mathcal{P}(\mathbf{e}_2). \tag{30}$$

Proof: If $\mathcal{D}(\mathbf{e})$ occurs, then from (11)

$$\mathbf{e} = \arg \min_{\mathbf{x} \in \{\mathcal{E}, \mathbf{0}\}} \left\| 2\tilde{\mathbf{F}}\mathbf{C}\mathbf{x} - \mathbf{n} \right\|^2 \tag{31}$$

$$\mathbf{e} = \arg \min_{\mathbf{x} \in \{\mathcal{E}, \mathbf{0}\}} \mathbf{x}^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{x} - \text{Re}(\mathbf{n}^* \tilde{\mathbf{F}} \mathbf{C} \mathbf{x}). \tag{32}$$

Therefore,

$$\begin{aligned}\mathbf{e}^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e} - \text{Re}(\mathbf{n}^* \tilde{\mathbf{F}} \mathbf{C} \mathbf{e}) < \mathbf{e}_2^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_2 \\ - \text{Re}(\mathbf{n}^* \tilde{\mathbf{F}} \mathbf{C} \mathbf{e}_2)\end{aligned} \tag{33}$$

$$\mathbf{e}_1^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_1 + 2\text{Re}(\mathbf{e}_1^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_2) < \text{Re}(\mathbf{n}^* \tilde{\mathbf{F}} \mathbf{C} \mathbf{e}_1). \tag{34}$$

Since decomposability requires that $\text{Re}(\mathbf{e}_1^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_2) \geq 0$, we have

$$\mathbf{e}_1^* \mathbf{C}^* \tilde{\mathbf{R}} \mathbf{C} \mathbf{e}_1 < \text{Re}(\mathbf{n}^* \tilde{\mathbf{F}} \mathbf{C} \mathbf{e}_1) \tag{35}$$

and so $\mathcal{P}(\mathbf{e}_1)$ occurs. Thus, $\mathcal{D}(\mathbf{e}) \subset \mathcal{P}(\mathbf{e}_1)$ and similarly, $\mathcal{D}(\mathbf{e}) \subset \mathcal{P}(\mathbf{e}_2)$. ■

Fig. 4 illustrates (30) with $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$ by showing the decision regions for four error vectors, with $\mathbf{0}$ corresponding to a correct decision. The left-hand figure shows the decomposable case, where the decoding event for $\mathbf{e}$ ($\mathcal{D}(\mathbf{e})$) is contained within the half-plane of $\mathcal{P}(\mathbf{e}_1)$ (dark shading is used to indicate the overlap). Note that the angle between $\tilde{\mathbf{F}}\mathbf{C}\mathbf{e}_1$ and $\tilde{\mathbf{F}}\mathbf{C}\mathbf{e}_2$ is less than 90° (corresponding to $\mathrm{Re}(\mathbf{e}_1^*\mathbf{C}^*\tilde{\mathbf{R}}\mathbf{C}\mathbf{e}_2) \geq 0$). The right-hand figure depicts the nondecomposable case. Here, we see that neither of the pairwise events [the half-planes $\mathcal{P}(\mathbf{e}_1)$ and $\mathcal{P}(\mathbf{e}_2)$] contain the decoding event for $\mathbf{e}$, and that the angle between $\tilde{\mathbf{F}}\mathbf{C}\mathbf{e}_1$ and $\tilde{\mathbf{F}}\mathbf{C}\mathbf{e}_2$ is greater that 90°.

The result in (30) implies that in bounding the union of all the $\mathcal{D}(\mathbf{e})$'s, it is sufficient to include the pairwise events for only those sequences which are not decomposable. Hence, the bound obtained in this way is known as the indecomposable sequence bound. When only decompositions with zero inner product are allowed, this becomes the simple-sequence bound [11]. Clearly, the upper bound (26) does not fully exploit (30), since it includes some sequences which are decomposable. Of course, the larger problem is that the indecomposable sequence set depends on the fading realization; however, we now show that it is possible to characterize additional classes of decomposable sequences for the GM decoder, and remove them from the bound (26).

We first decompose sequences for which $\mathbf{a}^*\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)\mathbf{a} = 0$ for all $\mathbf{a}$, which must occur when $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2) = \mathbf{0}$ (since this matrix is Hermitian). A sufficient condition for this to occur is when $\mathbf{e}_1$ is all zeros over a block where $\mathbf{e}_2$ is nonzero, and vice-versa; that is, for each i, either $\mathbf{e}_1(i) = \mathbf{0}$ or $\mathbf{e}_2(i) = \mathbf{0}$. Therefore, any sequence can be decomposed into one which is nonzero in only the time intervals where the mth user's error bit is nonzero, and the remainder which does not affect the mth user's bits. (For example, suppose $K = 2$, $m = 1$, and $\mathbf{e}^\top = [1\ 1|1\ 0|0\ 1|0\ 0]$. For the decomposition $\mathbf{e}_1^\top = [1\ 1|1\ 0|0\ 0|0\ 0]$ and $\mathbf{e}_2^\top = [0\ 0|0\ 0|0\ 1|0\ 0]$, the inner product is always zero. $\mathbf{e}_1$ will always be a valid error sequence in $\mathcal{E}$ for the GM decoder.) We wish to exclude sequences which are decomposable in this way from our union bound (26). Let $\hat{\mathcal{E}}(\mathbf{s})$ be the accordingly expurgated subset of the associated sequences $\mathcal{E}(\mathbf{s})$, so that every element of $\hat{\mathcal{E}}(\mathbf{s})$ is nonzero in the time intervals in which $\mathbf{s}$ contains a 1, and zero in the time intervals in which $\mathbf{s}$ contains a 0. Now the upper bound becomes

$$P_e^m < \frac{1}{k} \sum_{d=d_{\mathrm{free}}}^{\infty} \sum_{\mathbf{s}\in\mathcal{S}_d^m} N_m(\mathbf{s}) \sum_{\mathbf{e}\in\hat{\mathcal{E}}(\mathbf{s})} P_2(\mathbf{e}) \tag{36}$$

where $\mathcal{S}_d^m = \{\mathbf{s} \in \mathcal{S}^m\colon w(\mathbf{s}) = d\}$ as in (26).

For the single-user ML decoder, a technique for evaluating the upper bound in closed form is well known. This requires the code transfer function, a rational function $T(X, Y)$ which enumerates the set of error sequences. The infinite series expansion for $T(X, Y)$ resembles the series for error probability given in (23) and (26), where each term of the series $T(X, Y)$ has the form $\alpha X^d Y^i$, representing α sequences of weight d affecting i information bits. The error probability is then bounded as [9]

$$P_e < \frac{c}{k} \frac{\partial}{\partial Y} T(X, Y)\bigg|_{X=p,\, Y=1} \tag{37}$$

where pairwise probability is bounded as $P_2(\mathbf{e}) < cp^{w(\mathbf{e})}$.

We seek a transfer function bound for the GM decoder. Since the bound in (36) depends only on the sequences in $\mathcal{S}^m$, we need only find an upper bound on pairwise probability $P_2(\mathbf{e})$ which is exponential in the weight of the single-user sequence.

The conditional pairwise probability of $\mathbf{e}$, given $\mathbf{C}$, is upper bounded by

$$P_2(\mathbf{e}|\mathbf{C}) < \frac{1}{2} \exp\left(-\frac{1}{N_0}\mathbf{a}^*\mathbf{A}(\mathbf{e})\mathbf{a}\right) \triangleq P_2^{ub}(\mathbf{e}|\mathbf{C}). \tag{38}$$

The average of (38) is simply the characteristic function (21) evaluated at $s = -1/N_0$, which yields

$$P_2(\mathbf{e}) \leq \frac{1}{2} \det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{A}(\mathbf{e})\right)^{-1} \triangleq P_2^{ub}(\mathbf{e}). \tag{39}$$

Because $\mathbf{A}(\mathbf{e})$ is block diagonal, we can write $P_2^{ub}(\mathbf{e})$ as

$$P_2^{ub}(\mathbf{e}) = \frac{1}{2} \prod_{i=1}^{nI} \det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{A}_i(\mathbf{e})\right)^{-1}. \tag{40}$$

Thus, by defining

$$P_2^{ub}(\mathbf{e}(i)) \triangleq \frac{1}{2} \det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{A}_i(\mathbf{e})\right)^{-1} \tag{41}$$

and combining this result with (36), we have

$$P_e^m < \frac{1}{k} \sum_{d=d_{\mathrm{free}}}^{\infty} \sum_{\mathbf{s}\in\mathcal{S}_d^m} N_m(\mathbf{s}) \sum_{\mathbf{e}\in\hat{\mathcal{E}}(\mathbf{s})} \frac{1}{2} \prod_{i=1}^{nI} 2P_2^{ub}(\mathbf{e}(i)). \tag{42}$$

Since the set $\hat{\mathcal{E}}(\mathbf{s})$ is the set of all sequences that are nonzero only in the time intervals where $\mathbf{s}$ is nonzero and take on all possible combinations of $\{0, 1\}$ in the positions not affecting user m, (36) can be written solely in terms of $\mathbf{s}$. Define $\mathcal{U}^m = \{\mathbf{u}_i\}$ to be the set of all $U \triangleq 2^{K-1}$ binary K-length vectors which have a one in the mth position, so that for every $\mathbf{e} \in \hat{\mathcal{E}}(\mathbf{s})$, $\mathbf{e}(i) \in \{\mathcal{U}^m, \mathbf{0}\}\ \forall i$. Enumerating all combinations of these vectors gives

$$P_e^m < \frac{1}{k} \sum_{d=d_{\mathrm{free}}}^{\infty} \sum_{\mathbf{s}\in\mathcal{S}_d^m} N_m(\mathbf{s}) \sum_{\mathbf{x}\colon \mathbf{1}^\top\mathbf{x}=d} \binom{d}{x_1\ x_2 \cdots x_U} \cdot \frac{1}{2}\left(\prod_{i=1}^{U} [2P_2^{ub}(\mathbf{u}_i)]^{x_i}\right) \tag{43}$$

$$= \frac{1}{k} \sum_{d=d_{\mathrm{free}}}^{\infty} \sum_{\mathbf{s}\in\mathcal{S}_d^m} N_m(\mathbf{s}) \frac{1}{2}\left[\sum_{i=1}^{U} 2P_2^{ub}(\mathbf{u}_i)\right]^d \tag{44}$$

where we have used multinomial coefficients, $\mathbf{1}$ is a vector of ones, and $\mathbf{x}$ is a U-length vector of nonnegative integers. We have chosen the form of (42) in (43) to allow the collection of multinomial terms.

It is now clear that given the transfer function $T(X, Y)$ for the mth user's code, each term in the expansion of (37), $\alpha i p^d$ corresponds to α sequences in $\mathcal{S}_d^m$, with $N_m(\mathbf{s}) = i$. Thus, (44) is given in closed form by

$$P_e^m < \frac{1}{2k} \frac{\partial}{\partial Y} T(X, Y)\bigg|_{X=p,\, Y=1} \tag{45}$$

with $p = \sum_{i=1}^{2^{K-1}} 2P_2^{ub}(\mathbf{u}_i)$.

Example: For the single-user encoder shown in Fig. 3, we have: [9] $T(X, Y) = (X^5Y)/(1-2XY) = X^5Y + 2X^6Y^2 + 4X^7Y^3 + 8X^8Y^4 \cdots$. Hence the error probability of the GM decoder is simply bounded as $P_e^m < (1/2)(p^5/(1-2p)^2)$.

The tightening of the transfer-function bound by removing orthogonally decomposable sequences can also be applied to the ML bound, although it may be more difficult and less fruitful.

For example, for the coder specified in Fig. 2 with $d_{\text{free}} = 5$, one of the eight shortest orthogonally decomposable sequences is $[10|10|00|10|01|01|10|01|00|10|10|01|10|10|00|01|01|01]^\top$, with total weight 15. It is not known, however, whether this tightening might be more beneficial for more complex codes.

VII. Conditionally Decomposable Sequence Bound

The upper bound for the GM decoder here improves upon the bound of the preceeding section by considering sequences which are decomposable some fraction of the time; that is, sequences for which $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$ is not identically zero, and therefore depends on the channel fading parameters. These sequences may be conditionally decomposable, given the fading realization. In order to develop this bound, we need some specialized results which follow.

A. Probability of Decomposability

We now consider those sequence decompositions for which $\mathbf{a}^*\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)\mathbf{a} > 0$. We must therefore characterize the behavior of this random quadratic form in the same way as we examined the form $\mathbf{a}^*\mathbf{A}(\mathbf{e})\mathbf{a}$. To do so, we consider the eigenvalues of the constant matrix $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$ as in [4]. If $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$ were positive semi-definite, then $\mathbf{a}^*\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)\mathbf{a}$ would always be nonnegative, and the sequence $\mathbf{e}$ would be decomposable for any channel realization. However, we shall see shortly that this will only occur when $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2) = \mathbf{0}$ (and this is the point where the analysis in [4] concluded).

By the block diagonal nature of $\tilde{\mathbf{\Omega}}$, we can consider the diagonal blocks of $\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)$ separately, since $\mathbf{a}^*\mathbf{A}(\mathbf{e}_1, \mathbf{e}_2)\mathbf{a} = \sum_i \mathbf{a}^*(i)\mathbf{A}(\mathbf{e}_1(i), \mathbf{e}_2(i))\mathbf{a}(i)$. This leads to the following result.

Lemma 1: If $\mathbf{e}(i) = \mathbf{e}_1(i) + \mathbf{e}_2(i)$, with no cancellation between $\mathbf{e}_1$ and $\mathbf{e}_2$, then the random variable $\theta = \mathbf{a}^*(i)\mathbf{A}(\mathbf{e}_1(i), \mathbf{e}_2(i))\mathbf{a}(i)$ has a symmetric probability density function.

(A proof is deferred to the Appendix.) As a consequence of Lemma 1, we have that $\Pr(\theta > 0) = \Pr(\theta < 0) = 0.5$.

The important application of this lemma is of course in computing the probability of a sequence being decomposable. Consider a two-user example: for some error sequence $\mathbf{e}$, let $\mathbf{e}(i) = [1\ 1]^\top$. Then with probability 1/2, the sequence $\mathbf{e}$ can be decomposed into $\mathbf{e}_1$ and $\mathbf{e}_2$ so that $\mathbf{e}_1(i) = [1\ 0]^\top$, $\mathbf{e}_2(i) = [0\ 1]^\top$, and $\mathbf{e}_2(j) = [0\ 0]^\top$ and $\mathbf{e}_1(j) = \mathbf{e}(j)$ for $i \neq j$. Note that this is not an orthogonal decomposition used to obtain (44).

Any sequence $\mathbf{e}$ can be written uniquely as $\mathbf{e} = \mathbf{e}^m + \mathbf{e}^{\overline{m}}$ where $\mathbf{e}^m(i)$ is zero in all but the mth position and each $\mathbf{e}^{\overline{m}}(i)$ is zero in the mth position. We define

$$v_i(\mathbf{e}) \triangleq \mathbf{a}^*(i)\mathbf{A}(\mathbf{e}^m(i), \mathbf{e}^{\overline{m}}(i))\mathbf{a}(i) \tag{46}$$

for any $\mathbf{e}$, as in Lemma 1. We now make a the following few remarks about the implications of what has been said up to this point.

- The $v_i(\mathbf{e})$'s are uniquely defined for a given $\mathbf{e}$ and fading realization.
- If either $\mathbf{e}^m(i) = \mathbf{0}$ or $\mathbf{e}^{\overline{m}}(i) = \mathbf{0}$, then $v_i(\mathbf{e}) = 0$, otherwise $v_i(\mathbf{e}) \neq 0$ with probability 1.
- For the GM decoder for user m, $\mathbf{e}^m \in \mathcal{E}(\mathbf{s})$ and $\mathbf{e}^{\overline{m}} \in \mathcal{E}(\mathbf{0})$ for some single-user sequence $\mathbf{s}$.

In the analysis that follows, we shall be concerned with a sequence's decomposability only as a function of the $v_i(\mathbf{e})$'s, defined above. This does not consider all possible decomposable sequences.

Next, we will find the probability of a decomposition, and then compute the conditional pairwise probability of the appropriate sequences. Using only the v_is to indicate decomposability, we wish to compute

$$P_{2,S}(\mathbf{e}) \triangleq E[P_2(\mathbf{e}|\mathbf{C})|v_i(\mathbf{e}) < 0,\ i \in S] \tag{47}$$

where S is a set of time indices. The upper bound then comes from the decomposability as

$$P_e^m < \frac{1}{k} \sum_{d=d_{\text{free}}}^{\infty} \sum_{\mathbf{s}\in\mathcal{S}_d^m} N_m(\mathbf{s}) \sum_{\mathbf{e}\in\hat{\mathcal{E}}(\mathbf{s})} \sum_{S} \cdot P_{2,S}(\mathbf{e}) \Pr[v_i(\mathbf{e}) < 0,\ i \in S]. \tag{48}$$

Since $P_2(\mathbf{e}|\mathbf{C}) \leq P_2^{ub}(\mathbf{e}|\mathbf{C})$

$$P_{2,S}(\mathbf{e}) \leq E[P_2^{ub}(\mathbf{e}|\mathbf{C})|v_i(\mathbf{e}) < 0,\ i \in S]. \tag{49}$$

Furthermore, since $P_2^{ub}(\mathbf{e}|\mathbf{C}) = (1/2)\prod_i 2P_2^{ub}(\mathbf{e}(i)|\mathbf{C}(i))$, and the $\mathbf{C}(i)$s are independent, we have

$$P_{2,S}(\mathbf{e}) \leq \frac{1}{2} \prod_{i\in S} 2P_{2-}^{ub}(\mathbf{e}(i)) \prod_{i\notin S} 2E[P_2^{ub}(\mathbf{e}(i)|\mathbf{C}(i))] \tag{50}$$

where

$$P_{2-}^{ub}(\mathbf{e}(i)) \triangleq E[P_2^{ub}(\mathbf{e}(i)|\mathbf{C}(i))|v_i(\mathbf{e}) < 0]. \tag{51}$$

In general, this does not seem to have a closed-form expression, but it can be computed using Monte Carlo averaging.

B. Error Probability Bound

We are now ready to develop the new upper bound. We introduce another notational convention: given a vector $\mathbf{e}$ and a set of time indices S, we let the vector ${}^S\mathbf{e}$ be obtained by zeroing-out any elements $\mathbf{e}(i)$ with index i not in S. Note further that if $S \cup \overline{S} = \{0, \ldots, n(I-1)\}$ (where $\overline{S}$ is the complement of S), then $\mathbf{e} = {}^S\mathbf{e} + {}^{\overline{S}}\mathbf{e}$. For example, if $S = \{0, 2\}$, then ${}^S\mathbf{e} = [\mathbf{e}(0)^\top, \mathbf{0}^\top, \mathbf{e}(2)^\top, \mathbf{0}^\top, \ldots]^\top$.

Consider a sequence $\mathbf{e} = \mathbf{e}^m + \mathbf{e}^{\overline{m}}$ (where $\mathbf{e}^m$ affects only user m). Suppose that the fading is such that there is a S with $v_i(\mathbf{e}) < 0$ for $i \in S$ and $v_i(\mathbf{e}) \geq 0$ for $i \in \overline{S}$. In this case, let $\mathbf{e}_1 = \mathbf{e}^m + {}^S\mathbf{e}^{\overline{m}}$ and let $\mathbf{e}_2 = {}^{\overline{S}}\mathbf{e}^{\overline{m}}$. We have that

$$\mathbf{a}^*\mathbf{A}\left(\mathbf{e}^m + {}^S\mathbf{e}^{\overline{m}},\ {}^{\overline{S}}\mathbf{e}^{\overline{m}}\right)\mathbf{a} = \sum_{i\in\overline{S}} \mathbf{a}^*(i)\mathbf{A}_i\left(\mathbf{e}^m, \mathbf{e}^{\overline{m}}\right)\mathbf{a}(i) = \sum_{i\in\overline{S}} v_i(\mathbf{e}) \geq 0. \tag{52}$$

Hence, from (30)

$$\mathcal{D}(\mathbf{e}) \subseteq \mathcal{P}\left(\mathbf{e}^m + {}^S\mathbf{e}^{\overline{m}}\right) \tag{53}$$

so $\mathbf{e}$ is decomposable into $\mathbf{e} \overset{\text{dec}}{=} \mathbf{e}_1 + \mathbf{e}_2$, and $\mathbf{e}$ need not be counted in the error bound (when conditioned on the fading). We call such a decomposition a *GM decomposition* (when ${}^{\overline{S}}\mathbf{e}^{\overline{m}} \neq \mathbf{0}$). In fact, we can similarly show that for any other sequence $\mathbf{e}^m + {}^T\mathbf{e}^{\overline{m}}$ with $S \subset T$ we have

$$\mathcal{D}\left(\mathbf{e}^m + {}^T\mathbf{e}^{\overline{m}}\right) \subseteq \mathcal{P}\left(\mathbf{e}^m + {}^S\mathbf{e}^{\overline{m}}\right), \qquad \text{for } S \subset T. \tag{54}$$

Note that the conditions on the $v_i(\mathbf{e})$s above are sufficient for decomposability, but not necessary.

We now apply this result to tighten the simple sequence bound of (45). For each $\mathbf{e} = \mathbf{e}^m + \mathbf{e}^{\overline{m}} \in \hat{\mathcal{E}}(\mathbf{s})$ [recall (36)], let $S_{\mathbf{e}}$ be the set of indices for which $\mathbf{e}^{\overline{m}}(i) \neq \mathbf{0}$. (So in fact, ${}^{S_\mathbf{e}}\mathbf{e}^{\overline{m}} = \mathbf{e}^{\overline{m}}$.) This leads to the key result, which is that the sequence $\mathbf{e}$ in $\hat{\mathcal{E}}(\mathbf{s})$ must have a GM decomposition with $\mathbf{e}_1$ in $\hat{\mathcal{E}}(\mathbf{s})$ unless $v_i(\mathbf{e}) < 0$ for all $i \in S_\mathbf{e}$. This will occur with probability $2^{-|S_\mathbf{e}|}$, since from Lemma 1, $v_i(\mathbf{e}) < 0$ with probability 0.5, and the v_is are independent.

The only sequences that we include in the union bound are those for which we are not assured of a GM decomposition, so we now require the joint probability of the pairwise event for $\mathbf{e} = \mathbf{e}^m +^S \mathbf{e}^{\overline{m}}$ and the event that $v_i(\mathbf{e}) < 0$ for $i \in S$. We upper bound this probability using (51) as follows: let $\mathbf{u}_1$ be the K-length vector which is one in the mth position and zero elsewhere, then

$$\begin{aligned} &P_{2,S}(\mathbf{e}) \Pr(v_i(\mathbf{e}) < 0 \, \forall i \in S) \\ &\quad < \left[\frac{1}{2} \prod_{i \notin S} [2P_2^{ub}(\mathbf{e}(i))] \prod_{i \in S} [2P_{2-}^{ub}(\mathbf{e}(i))] \right] 2^{-|S|} \quad (55) \\ &\quad < \frac{1}{2} [2P_2^{ub}(\mathbf{u}_1)]^{w(\mathbf{e}^m)-|S|} \prod_{i \in S} [P_{2-}^{ub}(\mathbf{e}(i))]. \quad (56) \end{aligned}$$

The exponent $w(\mathbf{e}^m) - |S|$ arises as the number of occurrences of $\mathbf{u}_1$ in $\mathbf{e}$, since $\mathbf{e}^m(i)$ is either $\mathbf{u}_1$ or $\mathbf{0}$.

We enumerate the associated sequences by considering the alphabet of vectors for $\mathbf{e}(i)$, as in (43). The union bound on the probability of the associated sequences in $\hat{\mathcal{E}}(\mathbf{s})$, allowing for decomposability some fraction of the time, is given by summing over (56)

$$\Pr \left[\bigcup_{\mathbf{e} \in \hat{\mathcal{E}}(\mathbf{s})} \mathcal{D}(\mathbf{e}) \right] < \frac{1}{2} \sum_{\mathbf{e} \in \hat{\mathcal{E}}(\mathbf{s})} (2P_2^{ub}(\mathbf{u}_1))^{w(\mathbf{s})-|S_\mathbf{e}|} \cdot \prod_{i \in S_\mathbf{e}} \left[P_{2-}^{ub}(\mathbf{e}(i)) \right]. \quad (57)$$

However, (57) can be written in a more concise form by explicitly considering the structure of the associated sequences. Recall that the set $\hat{\mathcal{E}}(\mathbf{s})$ consists of all $2^{d(K-1)} = U^d$ combinations of U vectors in $d = w(\mathbf{s})$ time-intervals. We write these combinations explicitly using multinomial coefficients, and by indexing the alphabet of vectors for $\mathbf{e}(i)$, as in (43) so that the right-hand side of (57) is

$$\frac{1}{2} \sum_{\mathbf{x}:\, \mathbf{1}^\top \mathbf{x} = d} \binom{d}{x_1 \; x_2 \cdots x_U} [2P_2^{ub}(\mathbf{u}_1)]^{x_1} \prod_{i=2}^{U} (P_{2-}^{ub}(\mathbf{u}_i))^{x_i} \quad (58)$$

where the sum is over all U-length vectors $\mathbf{x}$ of nonnegative integers with $\sum x_j = d$. By virtue of the form of the probability terms, we can collect this multinomial expansion as

$$\frac{1}{2} \left[2P_2^{ub}(\mathbf{u}_1) + \sum_{i=2}^{2^{K-1}} P_{2-}^{ub}(\mathbf{u}_i) \right]^d. \quad (59)$$

The conditionally decomposable error bound for the GM decoder can now be stated using (48) as

$$P_e^m < \frac{1}{2k} \sum_{d=d_{\text{free}}}^{\infty} \sum_{\mathbf{s} \in \mathcal{S}_d^m} N_m(\mathbf{s}) \left[2P_2^{ub}(\mathbf{u}_1) + \sum_{i=2}^{2^{K-1}} P_{2-}^{ub}(\mathbf{u}_i) \right]^d. \quad (60)$$

The essence of this idea is that we count all of the sequences some of the time, we count some of the sequences all of the time, but we do not count all of the sequences all of the time.

It should be clear that the decomposable sequence bound can also be written in closed form, with the appropriate substitution for p in (45), so the bound (60) can be simply restated as

$$P_e^m < \frac{1}{2k} \frac{\partial}{\partial Y} T(X, Y) \bigg|_{X=\hat{p},\, Y=1} \quad (61)$$

with $\hat{p} = 2P_2^{ub}(\mathbf{u}_1) + \sum_{i=2}^{2^{K-1}} P_{2-}^{ub}(\mathbf{u}_i)$. Of course, we have not presented a closed-form expression for $\hat{p}$, however, its computation requires only 2^{K-1} numerical averages and is independent of the choice of codes. The expression (61) is closed form in the sense that we have avoided the infinite double summation of (60).

Table I illustrates the decomposability bound of (59) for all four sequences associated with the single-user sequence $\mathbf{s} = [1 \; 1]$ for user 1 by considering separately the channel realizations for a two-user system. These sequences all share the same v_is. Each row in the table is equally likely, and lists the set inclusions and minimum bound required. The first line of decompositions is made by noting that $\mathbf{e}_b = \mathbf{e}_a + [0 \; 1 | 0 \; 0]^\top$ so that $\mathbf{e}_b$ is decomposed into $\mathbf{e}_1 = \mathbf{e}_a$ and $\mathbf{e}_2 = [0 \; 1 | 0 \; 0]^\top$. The quadratic form $\mathbf{a}^* \mathbf{A}(\mathbf{e}_a, [0 \; 1 | 0 \; 0]^\top) \mathbf{a}$ is positive because of the condition on $v_1 = \mathbf{a}^*(1) \mathbf{A}([1 \; 0], [0 \; 1]) \mathbf{a}(1)$. The sequence $[0 \; 1 | 0 \; 0]^\top$ is valid since user 2's bits are unconstrained. Therefore, we have $\mathcal{D}(\mathbf{e}_b) \subset \mathcal{P}(\mathbf{e}_a)$, and $\mathbf{e}_b$ is not counted in the bound for this fading realization. The other decompositions follow similarly. The probabilities are all conditional on the v_is.

VIII. Multiple-User Subsets

While the GM decoder is most clearly presented with the assumption that only a single-user is coded, the strategy can easily be extended to consider a larger subset of the users to be coded. In this case, we define G to be the set of user indices of the assumed-coded users. Equation (15) can then be restated as follows.

Under the assumption that only the codes of the users in set G are known, the conditionally ML joint decoding rule for the users in G is

$$\begin{aligned} \hat{\mathbf{b}} &= \arg \min_{\mathbf{p} \in \mathcal{A}_G} \left\| \tilde{\mathbf{F}} \mathbf{C} \mathbf{p} - \mathbf{q} \right\|^2 \quad (62) \\ &= \arg \min_{\mathbf{p}^G \in \mathcal{X}_G} \left[\min_{\mathbf{p}^{\overline{G}} \in \mathcal{X}_{\overline{G}}} \left\| \tilde{\mathbf{F}} \mathbf{C} \left(\mathbf{p}^G + \mathbf{p}^{\overline{G}} \right) - \mathbf{q} \right\|^2 \right] \quad (63) \end{aligned}$$

where we replace $\mathcal{A}_m$ in (15) with $\mathcal{A}_G$, the set of all possible transmitted sequences assuming only the codes for users in G are known, and further decompose this set by letting $\mathcal{X}_G$ be the set of all vectors that are coded outputs for the users in G and are zero in all other positions, and let $\mathcal{X}_{\overline{G}}$ be the set of all vectors in $\{\pm 1\}$ which are zero in the positions for G.

If we let $|G|$ be the number of users in G, and let $|\overline{G}| = K - |G|$, then the GM decoder can be implemented with $2^{k\nu|G|}$ states, $2^k|G|$ branches from each state, with each branch minimizing over $n2^{|\overline{G}|}$ possibilities, resulting in a complexity of order $\mathcal{O}(2^{k(\nu+1)|G| + |\overline{G}|})$.

The performance analysis for this decoder is also a simple extension of the previous section. We proceed as in the single-user

TABLE I

TWO-USER EXAMPLE OF DECOMPOSABILITY AS A FUNCTION OF THE PARAMETERS v_1 AND v_2 IN (46). THE FOUR POSSIBLE SEQUENCES ARE $\mathbf{e}_a = [1\ 0|1\ 0]^\top$, $\mathbf{e}_b = [1\ 1|1\ 0]^\top$, $\mathbf{e}_c = [1\ 0|1\ 1]^\top$, AND $\mathbf{e}_d = [1\ 1|1\ 1]^\top$. $v_1 = \mathbf{a}^*(1)\mathbf{A}([1\ 0],\ [0\ 1])\mathbf{a}(1)$, $v_2 = \mathbf{a}^*(2)\mathbf{A}([1\ 0],\ [0\ 1])\mathbf{a}(2)$.

v_1	v_2	Decompositions	Event Inclusion	Error Bound for User 1
$+$	$+$	$\mathbf{e}_b = \mathbf{e}_a + [01\|00]^\top$ $\mathbf{e}_c = \mathbf{e}_a + [00\|01]^\top$ $\mathbf{e}_d = \mathbf{e}_a + [01\|01]^\top$	$\mathcal{P}(\mathbf{e}_a) \supseteq \{\mathcal{D}(\mathbf{e}_a), \mathcal{D}(\mathbf{e}_b), \mathcal{D}(\mathbf{e}_c), \mathcal{D}(\mathbf{e}_d)\}$	$P_2^{ub}(\mathbf{e}_a)$
$+$	$-$	$\mathbf{e}_b = \mathbf{e}_a + [01\|00]^\top$ $\mathbf{e}_d = \mathbf{e}_c + [01\|00]^\top$	$\mathcal{P}(\mathbf{e}_a) \supseteq \{\mathcal{D}(\mathbf{e}_a), \mathcal{D}(\mathbf{e}_b)\}$ $\mathcal{P}(\mathbf{e}_c) \supseteq \{\mathcal{D}(\mathbf{e}_c), \mathcal{D}(\mathbf{e}_d)\}$	$P_2^{ub}(\mathbf{e}_a) + P_2^{ub}(\mathbf{e}_c \| v_2 < 0)$
$-$	$+$	$\mathbf{e}_c = \mathbf{e}_a + [00\|01]^\top$ $\mathbf{e}_d = \mathbf{e}_b + [00\|01]^\top$	$\mathcal{P}(\mathbf{e}_a) \supseteq \{\mathcal{D}(\mathbf{e}_a), \mathcal{D}(\mathbf{e}_c)\}$ $\mathcal{P}(\mathbf{e}_b) \supseteq \{\mathcal{D}(\mathbf{e}_b), \mathcal{D}(\mathbf{e}_d)\}$	$P_2^{ub}(\mathbf{e}_a) + P_2^{ub}(\mathbf{e}_b \| v_1 < 0)$
$-$	$-$	-	-	$P_2^{ub}(\mathbf{e}_a) + P_2^{ub}(\mathbf{e}_b \| v_1 < 0)$ $+ P_2^{ub}(\mathbf{e}_c \| v_2 < 0)$ $+ P_2^{ub}(\mathbf{e}_d \| v_1, v_2 < 0)$

case, except that we must consider the joint codewords for the users in group G. We then make decompositions of the form $\mathbf{e} = \mathbf{e}^G + \mathbf{e}^{\overline{G}}$, where $\mathbf{e}^G$ affects only the users in group G and $\mathbf{e}^{\overline{G}}$ does not. For the upper bound, we have that the error rate for user m in the group G can be stated as follows: let $\mathcal{U}^G$ be the set of all K-vectors in $\{0,\ 1\}$ that are one in position m and zero for the positions in $\overline{G}$, let $\mathcal{U}^{\overline{G}}$ be the set of all K-vectors that are zero in the positions in G, and let $\mathcal{U}$ be the set of vectors that are nonzero in $\overline{G}$ and one in position m. Thus, every vector in $\mathbf{w} \in \mathcal{U}$ can be written as $\mathbf{w} = \mathbf{u} + \mathbf{v}$, with $\mathbf{u} \in \mathcal{U}^G$ and $\mathbf{v} \in \mathcal{U}^{\overline{G}}$. Then

$$P_e^m < \frac{1}{2k} \sum_{d=d_{\text{free}}}^{\infty} \sum_{\mathbf{s} \in \mathcal{S}_d^{G,\,m}} \sum_{\mathbf{e} \in \hat{\mathcal{E}}(\mathbf{s})} N_m(\mathbf{e}) \cdot \left[\sum_{\mathbf{u} \in \mathcal{U}^G} 2P_2^{ub}(\mathbf{u}) + \sum_{\mathbf{w} \in \mathcal{U}:\ \mathbf{w}=\mathbf{u}+\mathbf{v}} \cdot P_2^{ub}(\mathbf{w} | \mathbf{a}^* \mathbf{A}(\mathbf{u},\ \mathbf{v})\mathbf{a} < 0) \right]^d. \tag{64}$$

where $\mathcal{S}_d^{G,\,m}$ is the set of all shift-distinct multiuser error sequences for group G which affect user m and are nonzero in d symbol intervals, and $\hat{\mathcal{E}}(\mathbf{s})$ is the set of all K-user sequences of which $\mathbf{s}$ is a subsequence, and are zero for those time intervals where $\mathbf{s}(i) = \mathbf{0}$.

IX. NUMERICAL RESULTS

In order to assess the effectiveness of the bounds for both decoders, it is necessary to conduct computer simulations to estimate the true bit-error rate (BER). The simulations presented here assume the channel model is exactly as given in Section III, with perfect interleaving and perfect channel fading estimates at the receiver. The modulation is BPSK.

Fig. 5 shows simulation results for a two-user single-path ($K = 2$, $L = 1$) equal-power example with a signal cross correlation of 0.5; $\mathbf{R} = \left[\begin{smallmatrix} 1 & 0.5 \\ 0.5 & 1 \end{smallmatrix}\right]$. The BER is plotted versus

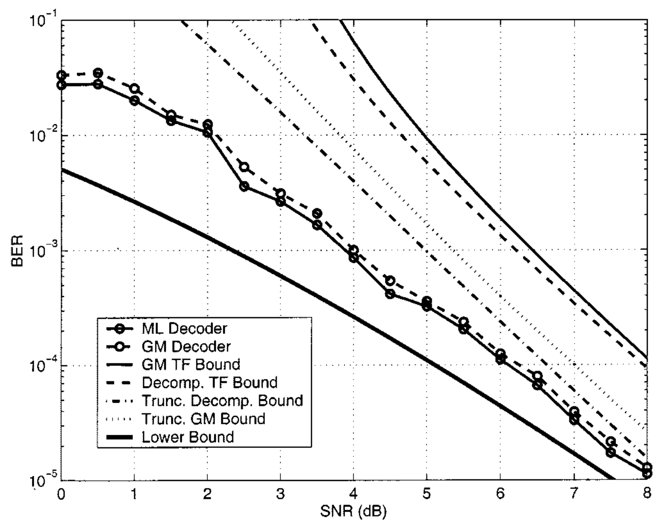

Fig. 5. Bounds and simulation results for ML and GM decoders for two users with medium (0.5) cross correlation.

signal-to-noise ratio E_c/N_0, where E_c is the energy per coded symbol. Both users employ the same rate 1/2 code with $d_{\text{free}} = 5$. (This is the encoder of Fig. 2.) We see that in this case the simulated performance of the GM decoder is quite close to that of the ML decoder.

The curve labeled "GM TF bound" is obtained from (45), while the "Decomp. TF Bound" is obtained from (61). The conditional pairwise probabilities in (61) are computed by Monte Carlo averaging. Truncated series bounds are also shown. The curve "Trunc. GM Bound" is obtained by computing (22) and substituting into (36), with outer summation evaluated from $d = 5 \cdots 10$. The "Trunc. Decomp. Bound" curve is obtained by numerically averaging (47) with conditioning on $v_i(\mathbf{e}) < 0$ and evaluating (48) for all associated sequences with $d \leq 10$. These truncated series bounds are much tighter than the closed form bounds, since P_2^{ub} is a loose upper bound on P_2. The lower bound is the minimum-distance bound: the probability of a single error sequence with weight d_{free}.

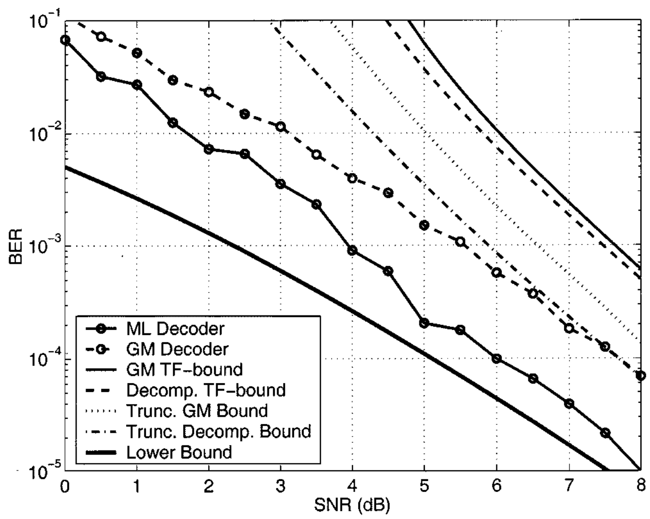

Fig. 6. Bounds and simulation results for ML and GM decoders for two users with singular correlation matrix.

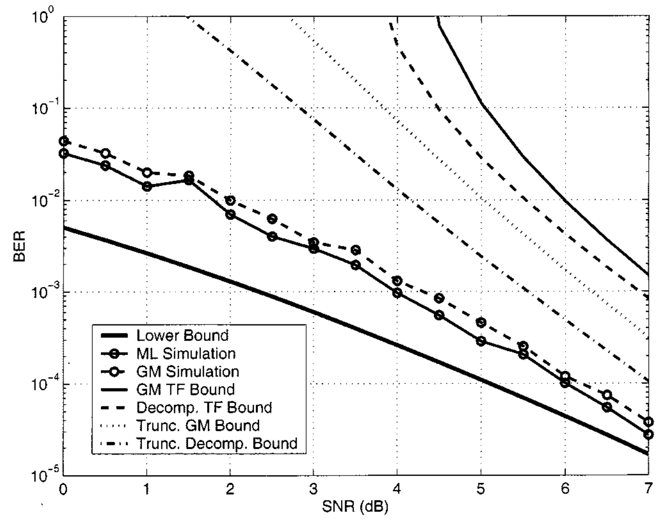

Fig. 8. Bounds and simulation results for a three-user example with 0.5 cross correlation. The GM decoder uses a group size of 1.

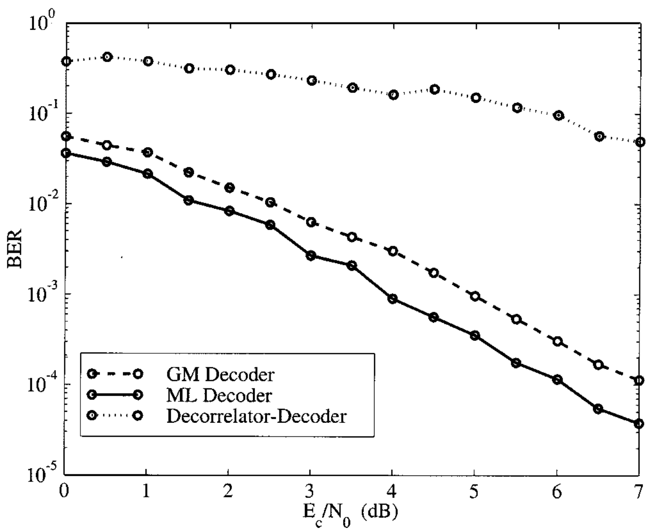

Fig. 7. Bounds and simulation results for a two-user example with 0.9 cross correlation. Also shown is the simulated performance of a decorrelator followed by a single-user decoder.

Fig. 6 depicts simulation results for the same example, but with a cross correlation of 1. In this case, $\mathbf{R}$ is singular so that we have a narrow-band channel. Here, the performance gap between the GM decoder and ML decoder is larger, but still within about 1.5 dB. The truncated bounds (labeled "Trunc. GM Bound" and "Trunc. Decomp. Bound") include single-user sequences with $d \leq 14$.

In order to see the advantage of using a multiuser decoder, even the reduced-complexity GM decoder, we compare the previous example with a single-user decoder in Fig. 7. Here, the users have been separated with a decorrelating transformation $\mathbf{R}^{-1}$, then those outputs are passed to a soft-input single-user decoder. In this case, the penalty due to the decorrelation is quite high.

A three-user example is shown in Fig. 8. In this case, we have assumed a single-path ($L = 1$) model with $\mathbf{R}$ being an equicorrelated matrix with off-diagonal elements equal to 0.5 and with equal-power users, and the same $d_{\text{free}} = 5$ code of the previous examples. Simulation results for the ML decoder, and for the GM decoder with group size of 1, are presented. The GM decoder closed-form bounds obtained from (45) and (61) are also shown, as in the previous figures, along with truncated series bounds. In this example, the ML decoder requires roughly 16 times the number of computations as a single GM decoder. However, the simulated performance is quite close, showing the advantage of the one-user GM decoding scheme.

X. Summary

The group-metric decoding strategy exploits a unique feature of the multiuser decoding problem for reducing complexity. This strategy performs well since the effective super-coder of the combined users must be noncooperative, and therefore cannot increase the free distance of the super-code above that of the single-user codes. The GM and ML decoders thus share the same minimum-distance performance, and should be expected to be asymptotically equivalent. Simulation results show that the performance loss relative to the ML decoder can be negligible.

The conditionally ML nature of the GM decoder allows for its performance to be completely analyzed using standard techniques for convolutional codes, once the assumed super-coder is characterized. In addition, a new analysis which removes conditionally decomposable sequences from the union bound can further tighten the performance bound for the GM decoder, providing a valuable design tool.

Appendix

Lemma 2: A matrix $\mathbf{X}$ of the form $\left[\begin{smallmatrix} \mathbf{0} & \mathbf{M}^* \\ \mathbf{M} & \mathbf{0} \end{smallmatrix}\right]$ has the property that if λ is an eigenvalue, $-\lambda$ is also an eigenvalue and $\lambda \in \{\pm\sqrt{\text{eig}(\mathbf{M}^*\mathbf{M})}\}$ for $\lambda \neq 0$.

Proof: Let $\lambda\mathbf{v} = \mathbf{X}\mathbf{v}$. Rewriting

$$\mathbf{X}\mathbf{v} = \begin{bmatrix} \mathbf{0} & \mathbf{M}^* \\ \mathbf{M} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{M}^*\mathbf{y} \\ \mathbf{M}\mathbf{x} \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x} \\ \lambda\mathbf{y} \end{bmatrix}. \quad (65)$$

Then we can also write

$$\begin{bmatrix} \mathbf{0} & \mathbf{M}^* \\ \mathbf{M} & \mathbf{0} \end{bmatrix} \begin{bmatrix} -\mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{M}^*\mathbf{y} \\ -\mathbf{M}\mathbf{x} \end{bmatrix} = -\lambda \begin{bmatrix} -\mathbf{x} \\ \mathbf{y} \end{bmatrix} \tag{66}$$

so $-\lambda$ is also an eigenvalue. Also, if $\lambda \neq 0$, $\mathbf{M}^*\mathbf{y} = \lambda\mathbf{x}$ and $\mathbf{M}\mathbf{x} = \lambda\mathbf{y}$, so $\mathbf{M}^*\mathbf{M}\mathbf{x} = \lambda^2\mathbf{x}$ and $\mathbf{M}\mathbf{M}^*\mathbf{y} = \lambda^2\mathbf{y}$. Thus λ^2 is an eigenvalue of $\mathbf{M}^*\mathbf{M}$. ■

Proof of Lemma 1:

Proof: Let the subsets of indices in $\{1, \ldots, K\}$ for which $\mathbf{e}_1(i)$ and $\mathbf{e}_2(i)$ are equal to zero be denoted as G and H, respectively. Hence θ can be written as

$$\theta = \frac{1}{2}[\mathbf{a}^*(i)_G \ \mathbf{a}^*(i)_H] \begin{bmatrix} \mathbf{0} & \mathbf{\Omega}_{HG}^* \\ \mathbf{\Omega}_{HG} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a}(i)_G \\ \mathbf{a}(i)_H \end{bmatrix} \tag{67}$$

where $\mathbf{a}(i)_G$ and $\mathbf{a}(i)_H$ are the sub-vectors of $\mathbf{a}(i)$ obtained by retaining the elements indexed by G and H, respectively. $\mathbf{\Omega}_{HG}$ is the sub-matrix of $\mathbf{\Omega}$ obtained by retaining the rows indexed by H and the columns indexed by G.

Using Lemma 2, we can form a spectral decomposition of $\hat{\mathbf{A}}(\mathbf{e}_1, \mathbf{e}_2)$

$$\hat{\mathbf{A}}(\mathbf{e}_1, \mathbf{e}_2) = \mathbf{U}^* \begin{bmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U} \tag{68}$$

where $\mathbf{U}$ is unitary and $\mathbf{D}$ is a diagonal matrix of eigenvalues. Since the vector $\mathbf{a}(i)$ is zero-mean, complex Gaussian with $E(\mathbf{a}\mathbf{a}^*) = \mathbf{I}$, so is $\mathbf{v} = \mathbf{U}[\mathbf{a}^*(i)_G \ \ \mathbf{a}^*(i)_H]^*$. Let $\mathbf{v}^* = [\mathbf{v}_1^* \ \mathbf{v}_2^* \ \mathbf{v}_3^*]$, then $E(\mathbf{v}_1\mathbf{v}_1^*) = E(\mathbf{v}_2\mathbf{v}_2^*) = \mathbf{I}$. Therefore

$$\theta = \mathbf{v}^* \begin{bmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{v} = [\mathbf{v}_1^* \ \mathbf{v}_2^*] \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} \tag{69}$$

$$= \mathbf{v}_1^*\mathbf{D}\mathbf{v}_1 - \mathbf{v}_2^*\mathbf{D}\mathbf{v}_2 \tag{70}$$

and since $\mathbf{v}_1$ and $\mathbf{v}_2$ are independent and identically distributed, so are $\mathbf{v}_1^*\mathbf{D}\mathbf{v}_1$ and $\mathbf{v}_2^*\mathbf{D}\mathbf{v}_2$. Thus, θ has the same probability density as $-\theta$. ■

REFERENCES

[1] E. A. Fain and M. K. Varanasi, "Group-metric decoding for frequency-selective Rayleigh fading multiple-access channels," in *Proc. IEEE GLOBECOM Communication Theory Mini-Conf.*, Nov. 1997, pp. 205–207.

[2] ——, "Error probability bound for reduced complexity multiuser decoding using orthogonal decomposability," in *Proc. IEEE Int. Symp. Information Theory*, Aug. 1998, p. 341.

[3] S. Verdú, "Minimum probability of error for Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[4] Z. Zvonar and D. Brady, "Multiuser detection in single-path fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 1729–1739, Feb./Mar./Apr. 1994.

[5] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.

[6] T. R. Giallorenzi and S. G. Wilson, "Multiuser ML sequence estimator for convolutionally coded asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, pp. 997–1008, Aug. 1996.

[7] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[8] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[9] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[10] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[11] S. Verdú, "Maximum likelihood sequence detection for intersymbol interference channels: A new bound on error probability," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 62–68, Jan. 1987.

Eric A. Fain received the B.S. degree in engineering from Harvey-Mudd University, Claremont, CA, and the Ph.D. degree in electrical engineering from the University of Colorado at Boulder in 1999.

He was a Research Assistant at the University of Colorado at Boulder from 1994 to 1999. His research activities were in the area of multiuser communications. In 1999, he joined Lockheed-Martin, San Jose, CA.

Mahesh K. Varanasi (S'87–M'89–SM–95) received the B.E. degree in electronics and communications engineering from Osmania University, Hyderabad, India, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1987 and 1989, respectively.

In 1989, he joined the faculty of the University of Colorado at Boulder in the Department of Electrical and Computer Engineering, where he is now an Associate Professor. His teaching and research interests include the areas of communication theory, information theory, and signal processing. His research has been in the areas of multiuser detection, space–time communications, equalization, signal design and power control for multiple access, fading channels and diversity systems, blind receivers and power and bandwidth-efficient multiuser communications.