Multiuser Decision-Feedback Receivers for the General Gaussian Multiple-Access Channel

Tommy Guess and Mahesh K. Varanasi †
Department of Electrical and Computer Engineering
University of Colorado, Boulder, CO 80309

Abstract

This paper considers multiple users accessing a common additive white Gaussian noise channel. The users are assigned signature waveforms, their signaling is synchronous, and hence the equivalent discrete-time model is a general Gaussian multiple-access channel (GMAC). A new class of receivers is proposed. The receivers within this class feature low-complexity detectors and allow the users to choose single-user codes independently of each other. The latter is accomplished by decoding the users sequentially, removing the contribution of the previously decoded user at each stage using feedback. These decision-feedback receivers (DFR) provide a continuum between single-output onion peeling and the Cholesky DFR. The latter uses the detector to remove inter-user interference, while the former does not. For a given set of users with fixed powers and signature waveforms, the receiver within this class that simultaneously maximizes the capacity of each user is found. This is the capacity-maximizing DFR.

1 Introduction

The capacity region for the conventional multiuser Gaussian channel is well known [1]. Although the boundary points of this capacity region can be achieved by random codes, the decoding complexity for such codes is exorbitant. Meanwhile, research into the design of joint codebooks for the users that allow them to approach the boundary of the capacity region is still in its fledgling stages. Another approach to this problem is to decode the users sequentially by means of *onion peeling* [1] [2] where the first user regards the other users as noise. The decoded symbols of the first user can be removed from the received codeword. If the symbols are decoded correctly, the result is that the first user is effectively removed from the system. This procedure is repeated for the second user and sequentially until the last user is all that is left to be decoded. This approach allays the primary difficulty of multiuser coding since it breaks the multiuser problem into several single-user problems, and single-user coding theory is well established. Such a strategy can achieve the vertices of the capacity region.

In this paper we apply the concept of sequentially decoding the users to the situation where they are allowed to use spreading waveforms. Spreading refers to the bandwidth

† This work was supported in part by NSF grant NCR-9406069.

expansion needed to accommodate the signature waveforms of the users in a code-division multiple-access system so as to provide some separation of the users. Whereas, if all of the users have the same signature waveform, then there is no spreading and error-control coding alone must be used to separate the users. When the users operate synchronously, the equivalent discrete-time model is the general Gaussian multiple-access channel (GMAC), a channel whose capacity region is given in [3]. This is a linear vector channel that is defined by a correlation matrix that contains the correlations of the users' signature waveforms with one another. When the users all have the same signature waveform (no spreading), this general GMAC reduces to the conventional GMAC [1] [3], and when the users have orthogonal signature waveforms (maximum spreading) this channel is simply a bank of single-user channels. In [4] it was shown that when the users are energy-weighted root-mean-squared (EW-RMS) bandlimited, the total capacity of the channel is in general maximized with signature waveforms that are neither orthogonal nor identical. This is an incentive for considering the general GMAC.

We want our receiver to have several properties. First, the detector should have a low complexity, preferably linear in the number of users, and it must be applicable to cases where the correlation matrix of the users is singular. Second, like onion peeling, it should decode the users sequentially (each user is decoded based on a single output of the vector channel) so that the users can choose single-user codes independently of each other. We propose a class of decision-feedback receivers (DFR) that satisfy the desired requirements. When the users' powers and signature waveforms are fixed, the receivers within this class cover the range between, and including, single-output onion peeling (SOOP) and the Cholesky-DFR (C-DFR) [5]. By SOOP we mean onion peeling that is applied to the general GMAC with each user being decoded based on a single output of the vector channel. For example, the first user is decoded based on the sampled outputs of the first matched filter as opposed to being decoded based on the sampled outputs of *all* the matched filters. The detector portion of the C-DFR uses the signature waveforms to separate the users by removing inter-user interference, but for correlated users the effective powers of the users can be severely reduced. SOOP, on the other hand, does not use the signature waveforms to reduce inter-user interference, and at high signal-to-noise ratios (SNR) this can significantly increase the effective noise powers seen by the users. General DFRs provide a continuum between these two extremes by varying the extent to which the receiver uses the detector to remove inter-user interference. The receiver that simultaneously maximizes the capacities of the users is found; it is the capacity-maximizing DFR (CM-DFR). In channels that emphasize spreading, the CM-DFR it looks like the C-DFR, but in regions where spreading is not important it behaves like the SOOP receiver.

The remainder of this paper is organized as follows: Section 2 introduces the general GMAC and shows its relation to synchronous CDMA when the users are bandlimited. Section 3 considers DFRs and develops the CM-DFR.

2 The System

Consider a PAM, synchronous, Gaussian CDMA channel where users are assigned signature waveforms that are time-limited to $[0, T]$. The signal that is received at some central location is modeled as the sum of M transmitting users and white Gaussian noise; it is given by

$$r(t) = \sum_{k=1}^{N} \sum_{i=1}^{M} b_i(k) u_i(t - kT) + n(t), \tag{1}$$

where $\{b_i(k)\}_k$ is the sequence of information symbols sent by the i^{th} user, $u_i(t)$ is the signature waveform of the i^{th} user, and $n(t)$ is an additive white Gaussian noise process with a power-spectral density of $N_0/2$. If each user's input waveform is power constrained to be w_i, then without loss of generality it can be assumed that $\frac{1}{NT}\sum_{k=1}^{N} b_i^2(k) = w_i$ for each i, and that the energy of the i^{th} user's waveform is $1 = \int_0^T u_i^2(t)dt$. Define the elements of the correlation matrix, $\mathbf{R}$, associated with a given set of users as

$$\mathbf{R}_{mn} = \int_0^T u_m(t) u_n(t) dt, \tag{2}$$

so that the each diagonal element of $\mathbf{R}$ is unity, and let $\mathbf{W}$ be a diagonal matrix with $\{w_1, w_2, \ldots, w_M\}$ as its diagonal. The received signal is input into a bank of filters that are matched to the M signature waveforms of the users, and the outputs of these filters are sampled at integer multiples of T to yield a set of sufficient statistics for the users' symbols. The channel is memoryless so it can be considered on a per-symbol basis. Letting $\mathbf{r}(k)$ denote an M-dimensional vector containing these sufficient statistics at time k, we have the following general GMAC

$$\mathbf{r}(k) = \mathbf{R}\mathbf{b}(k) + \mathbf{n}(k), \tag{3}$$

where $\mathbf{b}(k)$ is the vector of information symbols of the users and $\{\mathbf{n}(k)\}$ is a sequence of independent, identically-distributed(i.i.d.), zero-mean, Gaussian random vectors, each with covariance $\frac{N_0}{2}\mathbf{R}$.

3 Decision-Feedback Receivers for the GMAC

Consider a receiver structure where $\mathbf{r}(k)$ is input into a linear feedforward filter, the matrix $\mathbf{F}$. The first user is decoded based on the output $\text{row}_1(\mathbf{F})\mathbf{r}(k)$ (where $\text{row}_1(\mathbf{F})$ denotes the first row vector of $\mathbf{F}$), and these symbol decisions are subtracted from the other users after being fed back through a linear feedback filter, the matrix $\mathbf{B}$. Similarly, the second user is then decoded and the symbol decisions of this user are subtracted from users 3 to M after being fed back through $\mathbf{B}$. This process is repeated for users 3 up to M. Since we are considering capacity as the performance measure, we assume that all users are decoded correctly. Then, if $\mathbf{y}_i(k)$ denotes the effective channel of the i^{th} user, we have that

$$\mathbf{y}(k) = (\mathbf{FR} - \mathbf{B})\mathbf{b}(k) + \tilde{\mathbf{n}}(k), \tag{4}$$

where $\{\tilde{\mathbf{n}}(k)\}$ is a sequence of i.i.d., zero-mean, Gaussian random vectors with covariance $\frac{N_0}{2}\mathbf{FRF}^{\mathrm{T}}$, where the superscript T means transpose. Of course $\mathbf{B}$ is strictly lower triangular since only the decoded symbols of previous users are fed back. So for the i^{th} user we have

$$\mathbf{y}_i(k) = \mathbf{K}_{ii}\mathbf{b}_i(k) + \sum_{j\neq i}\mathbf{K}_{ij}\mathbf{b}_j(k) - \sum_{j=1}^{i-1}\mathbf{B}_{ij}\mathbf{b}_j(k) + \tilde{\mathbf{n}}_i(k), \tag{5}$$

where $\mathbf{K} = \mathbf{FR}$. When the input symbol distributions of the users are Gaussian, then it is readily seen that from the perspective of the i^{th} user this is a single-user, discrete-time Gaussian channel where the effective noise power is the sum of the system noise, $\tilde{\mathbf{n}}_i(k)$, and the "noise" of the interfering users, $\sum_{j\neq i}\mathbf{K}_{ij}\mathbf{b}_j(k) - \sum_{j=1}^{i-1}\mathbf{B}_{ij}\mathbf{b}_j(k)$. Thus the capacity of the effective single-user channel in nats per second is

$$C_i = \frac{1}{2T}\log\left(1 + \frac{\mathbf{K}_{ii}^2 w_i}{\sum_{j\neq i}(\mathbf{K}_{ij} - \mathbf{B}_{ij})^2 w_j + \frac{N_0}{2T}(\mathbf{FRF}^{\mathrm{T}})_{ii}}\right). \tag{6}$$

The ratio within the large parentheses of (6) shall be referred to as the signal-to- interference ratio of user i (SIR_i) because it includes interference from both the noise source and from the other users.

3.1 The Orthogonal Decision-Feedback Receiver

The orthogonal DFR (O-DFR) is applicable whenever the users' signature waveforms are orthogonal. The correlation matrix, $\mathbf{R}$, is the identity matrix and the system reduces to M single-user channels. Letting the feedfoward filter be the identity matrix and the feedback filter be the zero matrix(i. e., there is no feedback) allows all users to achieve their single-user capacities, that is, the capacities they would achieve if no interfering users were present. Of course this ideal performance comes at the cost of bandwidth. Time-division multiple access (TDMA) is an example of an orthogonal signaling scheme. It is not, however, the most efficient from a bandwidth point of view. That is, there are orthogonal signature waveforms that require less EW-RMS bandwidth than TDMA [4]. The set of orthogonal signature waveforms that minimize the bandwidth over all possible sets of orthogonal signature waveforms will be referred to as optimum-orthogonal multiple access (OOMA). OOMA is not a subset of TDMA since the non-zero portions of the OOMA signature waveforms, in general, overlap each other. This point was overlooked in [5], where TDMA should be replaced by OOMA.

3.2 Single-Output Onion Peeling

Single-output onion peeling (SOOP) is simply the application of onion peeling to the general GMAC with all of the users considering only the sampled outputs of their respective matched filters. It is a scheme that makes no effort to remove the interference of the other users with the feedforward filter. So $\mathbf{F}$ is once again chosen to be the identity matrix, and the feedback filter, $\mathbf{B}$, is the strictly lower-triangular part of the correlation matrix, $\mathbf{R}$. The capacity of the i^{th} user is

$$C_i = \frac{1}{2T}\log\left(1 + \frac{w_i}{\sum_{j=i+1}^{M}\mathbf{R}_{ij}^2 w_j + \frac{N_0}{2T}}\right). \tag{7}$$

If $\mathbf{R} = \mathbf{I}$ then this is the same as the orthogonal case, but here $\mathbf{R}$ can even be such that every element is unity. The latter situation is ideal from the point of view of bandwidth conservation, but at high SNRs the interfering users significantly limit the capacity since their interference term dominates the denominator of the SIR_i.

3.3 The Cholesky Decision-Feedback Receiver

The Cholesky DFR (C-DFR) was introduced in [5] (there it was called the Cholesky sequential decoder, but the term sequential decoder has been abandoned to avoid confusion with sequential decoding which is a method of decoding on single-user channels). It is a valid decoder whenever the correlation matrix is positive definite, so it inherently requires all the signature waveforms be linearly independent. When this is true, the correlation matrix, $\mathbf{R}$, is positive definite, and a Cholesky decomposition can be used to write $\mathbf{R} = \mathbf{L}^{\mathrm{T}}\mathbf{L}$ where $\mathbf{L}$ is a lower triangular matrix. The feedforward filter is given by $\mathbf{F} = (\mathbf{L}^{\mathrm{T}})^{-1}$ and the feedback filter by $\mathbf{B} = \mathbf{L} - \mathrm{diag}(\mathbf{L})$, where the $\mathrm{diag}(\mathbf{A})$ operation yields a diagonal matrix whose diagonal elements are the same as those of $\mathbf{A}$. Note that the noise term, $\tilde{\mathbf{n}}$, is white with each element having a power of $N_0/2$, and that $\mathbf{K} = \mathbf{L}$; both are true because of the decorrelating effect of $\mathbf{F}$. The resulting expression for (5) is $\mathbf{y}_i(k) = \mathbf{L}_{ii}\mathbf{b}_i(k) + \tilde{\mathbf{n}}_i(k)$, and this is independent of interference from other users. The capacity of the i^{th} user is given as

$$C_i = \frac{1}{2T}\log\left(1 + \frac{\mathbf{L}_{ii}^2 w_i}{N_0/(2T)}\right). \tag{8}$$

The cost of removing inter-user interference is that the $\mathbf{L}_{ii}^2$ term is less than unity unless the i^{th} user is orthogonal to users $i+1$ to M; this is a property of the Cholesky decomposition. Thus there is a potential tradeoff between the ability of the feedfoward filter to remove the effects of the interfering users, and the simultaneous reduction in the effective energy of the user of interest.

Finally, it should be pointed out that the C-DFR is unique among DFRs (excluding the relatively uninteresting case of orthogonal signaling) in the sense that it is applicable even when the users' symbol distributions are not Gaussian. This is true because the decorrelating action of the feedforward filter means that once the interference of the previous users has been removed by feedback, there is no residual interference (from other users).

3.4 The Capacity-Maximizing Decision-Feedback Receiver

Motivated by the observations of the previous subsections, we wish to optimize the capacity of the i^{th} user over all possible choices for the feedforward and feedback filters when the order in which the users are decoded is given. This is done under the assumption that the users' powers are fixed, as are N_0 and T. Thus we seek to solve

$$\mathbf{F}_{\mathrm{opt}}, \mathbf{B}_{\mathrm{opt}} \in \arg\max_{\mathbf{F},\mathbf{B}}\{C_i\}, \tag{9}$$

where $\mathbf{F}$ and $\mathbf{B}$ are real $M \times M$ matrices with $\mathbf{B}$ being strictly lower triangular, and C_i given by (6). Because the logarithm is a monotonic increasing function and T is fixed, the problem can be expressed equivalently as

$$\mathbf{F}_{\text{opt}}, \mathbf{B}_{\text{opt}} \in \arg\max_{\mathbf{F},\mathbf{B}} \left\{ \frac{\mathbf{K}_{ii}^2 w_i}{\sum_{j\neq i}(\mathbf{K}_{ij} - \mathbf{B}_{ij})^2 w_j + \frac{N_0}{2T}(\mathbf{F}\mathbf{R}\mathbf{F}^{\mathrm{T}})_{ii}} \right\}. \tag{10}$$

This objective function depends only the i^{th} rows of $\mathbf{F}$ and $\mathbf{B}$. This means that each row of $\mathbf{F}$ and $\mathbf{B}$ can be chosen to maximize the capacity of the corresponding user, and hence the capacities of all the users can be simultaneously maximized. Let $\mathbf{f}^{\mathrm{T}}$ denote the i^{th} row of $\mathbf{F}$ and β^{T} the i^{th} row of $\mathbf{B}$ so that (10) becomes

$$\mathbf{f}_{\text{opt}}, \beta_{\text{opt}} \in \arg\max_{\mathbf{f},\beta} \left\{ \frac{(\mathbf{f}^{\mathrm{T}}\text{col}_i(\mathbf{R}))^2 w_i}{(\mathbf{f}^{\mathrm{T}}\mathbf{R} - \beta^{\mathrm{T}})\hat{\mathbf{W}}_i(\mathbf{R}\mathbf{f} - \beta) + \frac{N_0}{2T}\mathbf{f}^{\mathrm{T}}\mathbf{R}\mathbf{f}} \right\}, \tag{11}$$

where $\hat{\mathbf{W}}_i$ denotes $\mathbf{W}$ with the i^{th} diagonal element being set to zero, and $\text{col}_i(\mathbf{A})$ denotes the i^{th} column vector of $\mathbf{A}$. Since the numerator is independent of β, the best β can do to maximize the objective function is make the denominator as small as possible. It is readily seen that the nonzero elements of β_{opt} are the same as the first $i-1$ elements of $\mathbf{R}\mathbf{f}$ for any $\mathbf{f}$ that is used. That is, β_{opt} simply removes the interference from the previous users. Since, for any user i, $\mathbf{B}$ has removed the interference of users 1 to $i-1$, user i "looks" like a first user. Therefore, to facilitate notation, we will proceed by considering only the first user.

Let $\hat{\mathbf{W}}_{j:k}$ be $\mathbf{W}$ with the diagonal elements j to k set to zero. Thus we are left to find

$$\mathbf{f}_{\text{opt}} \in \arg\max_{\mathbf{f}} \left\{ \frac{\mathbf{f}^{\mathrm{T}}\mathbf{R}\hat{\mathbf{W}}_{2:M}\mathbf{R}\mathbf{f}}{\mathbf{f}^{\mathrm{T}}(\mathbf{R}\hat{\mathbf{W}}_1\mathbf{R} + \frac{N_0}{2T}\mathbf{R})\mathbf{f}} \right\}. \tag{12}$$

For the case where $\mathbf{R}$ is positive definite the term in the parentheses is positive definite and this maximization is solved via a change of variables to transform the problem into a Rayleigh-Ritz parameterization of the largest eigenvalue. First make a change of variables so that $\mathbf{k} = \mathbf{R}\mathbf{f}$, an invertible transformation by virtue of the invertibility of $\mathbf{R}$. The SIR of user 1 is now

$$\text{SIR}_1 = \frac{\mathbf{k}^{\mathrm{T}}\hat{\mathbf{W}}_{2:M}\mathbf{k}}{\mathbf{k}^{\mathrm{T}}(\hat{\mathbf{W}}_1 + \frac{N_0}{2T}\mathbf{R}^{-1})\mathbf{k}}. \tag{13}$$

Now let $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ be the Cholesky decomposition of

$$\mathbf{S} \triangleq \hat{\mathbf{W}}_1 + \frac{N_0}{2T}\mathbf{R}^{-1} = \mathbf{J}^{\mathrm{T}}\mathbf{J}, \tag{14}$$

where $\mathbf{J}$ is lower triangular; $\mathbf{J}$ is invertible. Another change of variables with $\mathbf{s} = \mathbf{J}\mathbf{k}$ gives that

$$\text{SIR}_1 = \frac{\mathbf{s}^{\mathrm{T}}(\mathbf{J}^{\mathrm{T}})^{-1}\hat{\mathbf{W}}_{2:M}\mathbf{J}^{-1}\mathbf{s}}{\mathbf{s}^{\mathrm{T}}\mathbf{s}}. \tag{15}$$

This is maximized when $\mathbf{s}$ is any eigenvector corresponding to the largest eigenvalue of $(\mathbf{J}^{\mathrm{T}})^{-1}\hat{\mathbf{W}}_{2:M}\mathbf{J}^{-1}$, a diagonal matrix with only the 1,1 element being non-zero. The maximizing $\mathbf{s}$ is any non-zero multiple of the first column of the identity matrix. Hence we have that

$$\mathbf{f}_{\text{opt}} = \alpha\mathbf{R}^{-1}\text{col}_1\left(\mathbf{J}^{-1}\right), \tag{16}$$

where α is any non-zero constant, and that the resulting maximum SIR for user 1 is

$$\mathrm{SIR}_1 = \frac{w_1}{\mathbf{J}_{11}^2} = w_1 \left(\mathbf{S}^{-1}\right)_{11}. \tag{17}$$

The second equality follows from the relationship of the Cholesky decomposition to the Schur complement.

These results are extended in Appendix A to cover the case where $\mathbf{R}$ is only positive semi-definite. In Appendix B it is shown that, remarkably, this capacity-maximizing DFR (CM-DFR) is also a minimum mean-squared error (MMSE) DFR. That is, any $\mathbf{f}$ satisfying (16) also minimizes the MSE given by $\mathrm{E}\left\{(\mathbf{f}^{\mathrm{T}}\mathbf{R}\mathbf{b} + \tilde{\mathbf{n}}_1 - \mathbf{b}_1)^2\right\}$, and vice versa. A nice result of this equivalence is that the MMSE yields a very simple representation of the optimum feedforward and feedback filters (e.g., [7]),

$$\begin{aligned} \mathbf{F} &= \left(\mathbf{\Phi}^{\mathrm{T}}\right)^{-1} \mathbf{W}^{1/2} \\ \mathbf{B} &= \left(\mathbf{\Phi} - \mathrm{diag}\left(\mathbf{\Phi}\right)\right) \mathbf{W}^{-1/2}, \end{aligned} \tag{18}$$

where $\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} = \mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2} + \frac{N_0}{2T}\mathbf{I}$, and $\mathbf{\Phi}$ is lower triangular.

3.5 Total Capacity of DFRs Compared

We now illustrate how the DFRs perform relative to each other. When the signature sequences are non-orthogonal and linearly independent, we can compare the total capacities of SOOP, the C-DFR, and the CM-DFR as the noise power varies. Examples can be seen in Figures 1 and 2, where the time interval, T, is chosen for each case so the the EW-RMS bandwidth [4] is 100 Hz. The correlation matrices for the two cases are

$$\mathbf{R} = \begin{bmatrix} 1.0 & 0.8 & 0.6 \\ 0.8 & 1.0 & 0.4 \\ 0.6 & 0.4 & 1.0 \end{bmatrix}, \quad \mathbf{R} \begin{bmatrix} 1.0 & 0.9 & 0.9 \\ 0.9 & 1.0 & 0.9 \\ 0.9 & 0.9 & 1.0 \end{bmatrix}. \tag{19}$$

In the first case all of the users have unit power, while in the second they have powers of 4, 2 and 1. The SNR is defined to be the SNR of the weakest user, i. e., $10 \log \left(\frac{w_M}{N_0}\right)$. Note that at very high SNRs the performances of the CM-DFR and the C-DFR coincide. This happens because the noise power is very small compared to the power of interfering users, so the CM-DFR removes all interfering users by using a decorrelating feedforward filter. Similarly, at low SNRs the CM-DFR and SOOP have similar performances because allowing significant interference from other users degrades the SIR very little (since the SIR is dominated by the noise term) but also enables the effective power of the user of interest to be maximized. In these extreme regions it should be noticed that, in general, either the C-DFR or SOOP will not perform as well as the other. There are also regions in between the extremes where the CM-DFR performs significantly better than either the C-DFR or SOOP.

Summary

We have proposed a decision-feedback receiver structure that incorporates a low- complexity detector in conjunction with decision-feedback decoding for multiuser Gaussian

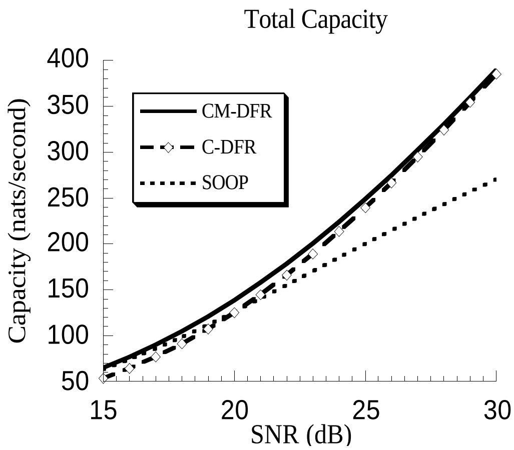

Figure 1: EW-RMS bandwidth is 100 Hz.

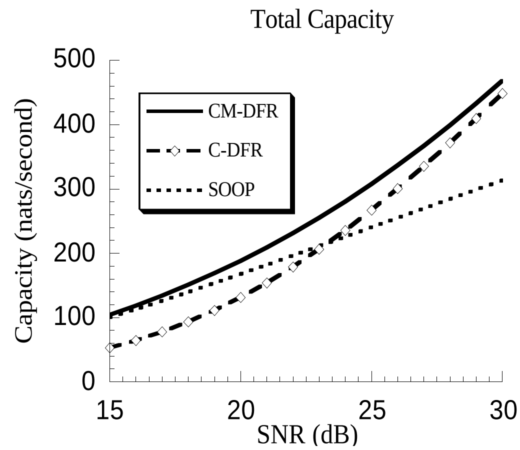

Figure 2: EW-RMS bandwidth is 100 Hz.

channels. Such receivers enable the users to incorporate error-control coding independently of each other using single-user codes. Single-output onion peeling (here applied to the general GMAC) and the C-DFR, two receivers in the literature that also use decision-feedback decoding, are special cases within this class of receivers. For a given correlation matrix we have derived the receiver from this class that simultaneously maximizes the capacity of each user, i. e., the CM-DFR.

Appendix

A Optimum DFR When R is Singular

Lemma 1 *Let* $\mathbf{R}$ *be a real,* $M \times M$*, normalized correlation matrix so that* $\mathbf{R}$ *is positive semi-definite and* $\mathrm{diag}(\mathbf{R}) = \mathbf{I}$. $\mathbf{R}$ *can be factored as*

$$\mathbf{R} = \mathbf{U}\mathbf{R}_{\mathrm{red}}\mathbf{U}^{\mathrm{T}} \triangleq \left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \\ & \mathbf{V} & & \end{array} \right] \mathbf{R}_{\mathrm{red}} \left[\begin{array}{cc} 1 & \\ 0 & \\ \vdots & \mathbf{V}^{\mathrm{T}} \\ 0 & \end{array} \right], \tag{20}$$

where $\mathbf{V}$ *is an* $M-1 \times n$ *matrix and* $\mathbf{R}_{\mathrm{red}}$ *(the subscript* red *denotes the word reduced) is a* $n \times n$ *matrix with* n *representing the rank of* $\mathbf{R}$*. In addition, the ranks of* $\mathbf{U}$ *and* $\mathbf{R}_{\mathrm{red}}$ *are both* n*.*

Proof: A singular-value decomposition of the real symmetric $\mathbf{R}$ allows us to write

$$\mathbf{R} = \mathbf{U_1}\mathbf{\Lambda}\mathbf{U_1}^{\mathrm{T}}, \tag{21}$$

with the dimensions of $\mathbf{U_1}$ and $\mathbf{\Lambda}$ being $M \times n$ and $n \times n$ respectively; both have a rank of n. Let $\mathbf{u}^{\mathrm{T}}$ be the first row of $\mathbf{U_1}$. We know that it cannot be the zero vector since the 1,1 element of $\mathbf{R}$ is unity. We let

$$\mathbf{A} = \left[\begin{array}{cc} \mathbf{u} & \mathbf{P} \left[\begin{array}{c} \mathbf{0}^{\mathrm{T}} \\ \mathbf{I} \end{array} \right] \end{array} \right], \tag{22}$$

where $\mathbf{I}$ is $n-1 \times n-1$, and $\mathbf{P}$ is a $n \times n$ permutation matrix that moves the all-zero row of the matrix it permutes to a row, say row k, where $\mathbf{u}_k \neq 0$. This is always possible since $\mathbf{u}$ is not the zero vector. Clearly, $\mathbf{A}$ is nonsingular and $\mathbf{A}^{-1}\mathbf{u} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^{\mathrm{T}}$. Finally,

$$\mathbf{R} = \left(\mathbf{U_1}\left(\mathbf{A}^{\mathrm{T}}\right)^{-1}\right)\left(\mathbf{A}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{A}\right)\left(\mathbf{A}^{-1}\mathbf{U_1}^{\mathrm{T}}\right) \tag{23}$$

gives the result. □

Lemma 2 *When the correlation matrix is singular, let* $\mathbf{R} = \mathbf{U}\mathbf{R}_{\mathrm{red}}\mathbf{U}^{\mathrm{T}}$ *be a factorization resulting from* Lemma 1 . *Then the* SIR *of the first user can be expressed as*

$$\mathrm{SIR}_1 = w_1 \left(\mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_1\mathbf{U} + \frac{N_0}{2T}\mathbf{R}_{\mathrm{red}}^{-1}\right)^{-1}_{11}. \tag{24}$$

Proof: Substitute the factorization of $\mathbf{R}$ into (12) to get that

$$\mathbf{f}_{\mathrm{opt}} \in \arg\max_{\mathbf{f}} \left\{ \frac{\mathbf{f}^{\mathrm{T}}\mathbf{U}(\mathbf{R}_{\mathrm{red}}\mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_{2:M}\mathbf{U}\mathbf{R}_{\mathrm{red}})\mathbf{U}^{\mathrm{T}}\mathbf{f}}{\mathbf{f}^{\mathrm{T}}\mathbf{U}(\mathbf{R}_{\mathrm{red}}\mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_1\mathbf{U}\mathbf{R}_{\mathrm{red}} + N_0/(2T)\mathbf{R}_{\mathrm{red}})\mathbf{U}^{\mathrm{T}}\mathbf{f}} \right\}. \tag{25}$$

The change of variables of $\mathbf{y} = \mathbf{R}_{\mathrm{red}}\mathbf{U}^{\mathrm{T}}\mathbf{f}$ leads to

$$\mathbf{y}_{\mathrm{opt}} \in \arg\max_{\mathbf{y}} \left\{ \frac{\mathbf{y}^{\mathrm{T}}(\mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_{2:M}\mathbf{U})\mathbf{y}}{\mathbf{y}^{\mathrm{T}}(\mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_1\mathbf{U} + N_0/(2T)\mathbf{R}_{\mathrm{red}}^{-1})\mathbf{y}} \right\}. \tag{26}$$

Even though the transformation by $\mathbf{U}$ is not invertible, it does not matter because this simply means that the set of $\mathbf{f}_{\mathrm{opt}}$ solutions is given by more than one basis vector.

Now factor $\mathbf{S}_{\mathrm{red}} \triangleq \mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_1\mathbf{U} + N_0/(2T)\mathbf{R}_{\mathrm{red}}^{-1}$ as $\mathbf{J}_{\mathrm{red}}^{\mathrm{T}}\mathbf{J}_{\mathrm{red}}$ where $\mathbf{J}_{\mathrm{red}}$ is lower triangular. Noting that the first column of $\mathbf{U}$ is all zeros save its first element which is unity, and the fact that $\hat{\mathbf{W}}_{2:M}$ has only its 1,1 element being non-zero, we see that $\mathbf{U}^{\mathrm{T}}\hat{\mathbf{W}}_{2:M}\mathbf{U} = \hat{\mathbf{W}}_{2:M}$. We can now follow the same reasoning that led to (17). when $\mathbf{R}$ was non-singular, since here we have $\mathbf{R}_{\mathrm{red}}$ which is non-singular. □

B Equivalence of the CM-DFR and the MMSE-DFR

Consider the MSE of the output of a linear feedforward filter for user i when the interference of all previous users has been removed. To show equivalence to the CM-DFR (in the sense that they can use the same feedforward and feedback filters) it is sufficient to show equivalence for the first user, since after the interference of users 1 to $i-1$ have been removed then user i looks like a first user. The MSE of the first user is

$$\mathrm{MSE}_1 = E\left\{(\mathbf{f}^{\mathrm{T}}\mathbf{R}\mathbf{b} + \tilde{\mathbf{n}}_1 - \mathbf{b}_1)^2\right\}, \tag{27}$$

where $\tilde{\mathbf{n}}_1$ has a variance of $N_0/(2T)\mathbf{f}^{\mathrm{T}}\mathbf{R}\mathbf{f}$. When $\mathbf{R}$ is invertible then we can let $\mathbf{k} = \mathbf{R}\mathbf{f}$. This change of variables allows one to show that the MSE is minimized by $\mathbf{k} = w_1\mathrm{col}_1(\mathbf{S}_{\mathrm{mmse}}^{-1})$ where $\mathbf{S}_{\mathrm{mmse}} = \mathbf{W} + N_0/(2T)\mathbf{R}^{-1}$. If $\mathbf{J}_{\mathrm{mmse}}$ is such that $\mathbf{J}_{\mathrm{mmse}}^{\mathrm{T}}\mathbf{J}_{\mathrm{mmse}} =$

$\mathbf{S}_{\text{mmse}}$ and $\mathbf{J}_{\text{mmse}}$ is lower triangular, then the optimizing $\mathbf{k}$ is $\beta \text{col}_1(\mathbf{J}_{\text{mmse}}^{-1})$ where β is a non-zero constant. So it can be stated that

$$\mathbf{f}_{\text{mmse}} = \beta \mathbf{R}^{-1} \text{col}_1(\mathbf{J}_{\text{mmse}}^{-1}) \tag{28}$$

is the best linear feedforward filter if it is desired to minimize the MSE. From the development of the CM-DFR it was found in (16) that the best feedforward filter is

$$\mathbf{f}_{\text{opt}} = \alpha \mathbf{R}^{-1} \text{col}_1(\mathbf{J}^{-1}), \tag{29}$$

where $\mathbf{J}^{\text{T}}\mathbf{J} = \hat{\mathbf{W}}_1 + N_0/(2T)\mathbf{R}^{-1}$ and α is any non-zero constant. Since $\mathbf{W}$ and $\hat{\mathbf{W}}_1$ are identical except for their $1,1$ elements, $\mathbf{J}$ and $\mathbf{J}_{\text{mmse}}$ differ only in their $1,1$ elements. Knowing this it is then easy to verify that the first column of the inverse of $\mathbf{J}$ is linearly dependent on the first column of the inverse of $\mathbf{J}_{\text{mmse}}$. Hence any feedforward filter that maximizes the SIR will also minimize the MSE and vice versa.

For the case when $\mathbf{R}$ is singular, the quadratic form of the MSE can be equivalently expressed by combining linearly dependent users as was done in Appendix A. Once this transformation has been made, we will have a reduced size correlation matrix (given by the $\mathbf{R}_{\text{red}}$ of Lemma 1) that is invertible so that the optimum feedfoward filter can be found. Once again it will be a non-zero multiple of the feedfoward filter that maximizes the SIR.

References

[1] Cover, T. M., Thomas, J. A., *Elements of Information Theory.* New York: Wiley, 1991.

[2] Gallager, R. G., "A Perspective on Multiaccess Channels," *IEEE Trans. Inform. Theory*, vol. IT-31, No. 2, pp. 124-142, March 1985.

[3] Verdú, S., "Capacity Region of Gaussian CDMA channels: The Symbol- Synchronous Case," in *Proc. Twenty-fourth Allerton Conf. Commun., Contr. and Comput.*, Allerton, IL, Oct. 1986, pp.1025-1034.

[4] Parsavand, D., Varanasi, M. K., "RMS Bandwidth Constrained Signature Waveforms that Maximize the Total Capacity of PAM-Synchronous CDMA Channels," *IEEE Trans. Commun. Theory*, vol. COM-44, pp. 65-75, Jan 1996.

[5] Guess, T. R., Varanasi, M. K., "Onion Peeling for CDMA: Symmetric Capacity Under RMS-Bandwidth Constraints," *Proc. Thirtieth Annual Conference on Information Sciences and Systems*, Princeton University, Mar. 1996, pp. 584-589.

[6] Cheng, R. S., Verdú, S., "Capacity of Root-Mean-Square Bandlimited Gaussian Multi-user Channels," *IEEE Trans. Inform. Theory*, vol. IT-37, No. 3, pp. 453-465, May 1991.

[7] Duel-Hallen, A., "Equalizers for Multiple Input/Multiple Output Channels and PAM Systems with Cyclostationary Input Sequences,"*IEEE Journal on Selected Areas in Communications*, Vol. 10, No. 3, April 1992, pp. 630-639.