Cooperative Lattice Coding and Decoding in Half-Duplex Channels

Arul Murugan, Kambiz Azarian, and Hesham El Gamal*

Abstract

We propose novel lattice coding/decoding schemes for half-duplex outage-limited cooperative channels. These schemes are inspired by the cooperation protocols of Azarian *et al.* and enjoy an excellent performance-complexity tradeoff. More specifically, for the relay channel, we first use our lattice coding framework to generalize Yang and Belfiore implementation of the non-orthogonal amplify and forward cooperation protocol. This generalization is shown to offer significant performance gains while keeping the decoding complexity manageable. We then devise a novel variant of the dynamic decode and forward protocol, along with a lattice-coded implementation, which enjoys a near-optimal diversity-multiplexing tradeoff with a low encoding/decoding complexity. Finally, for the cooperative multiple-access channel, we present a lattice-coded implementation of the non-orthogonal amplify and forward protocol and demonstrate its excellent performance-complexity tradeoff. Throughout the paper, we establish the performance gains of our proposed protocols via a comprehensive simulation study.

1 Keywords:

Lattice coding, lattice decoding, half-duplex channels, tree search, MMSE-Fano decoder, amplify and forward, decode and forward, cooperative multiple access.

*The authors are with the Department of Electrical Engineering, The Ohio State University, Columbus, OH. (e-mail : {palaniva,azariany,helgamal}@ece.osu.edu).

2 Introduction

Lately, cooperative communications has been the focus of intense research activities. As a consequence, we now have a wealth of results that covers a wide variety of system models and design criteria [1–12]. In this paper, we focus on two outage-limited, i.e., slow fading, cooperative channels, namely the relay and cooperative multiple-access (CMA) channels. We further impose the half-duplex constraint, limiting each node to either transmit or receive at any instant. The primary goal in this setting is to devise strategies that exploit the available *cooperative diversity*, at a *reasonable* decoding complexity. Our work here is inspired by the cooperation protocols of Azarian *et al.* in [1] and the lattice decoding framework of Murugan *et al.* in [13].

The design of cooperation protocols for half-duplex outage-limited channels was pioneered by the work of Laneman *et al.* [5]. Inspired by a high signal-to-noise ratio (SNR) analysis, Azarian *et al.* obtained more efficient protocols that achieve a superior diversity-multiplexing tradeoff (DMT) [1]. The practical implementation of Azarian *et al.* cooperation schemes was recently considered by Yang and Belfiore where they presented a low complexity scheme that achieves the optimal DMT of the amplify and forward relay channel. Here, building on our work in [1, 13], we construct low complexity lattice coding schemes which efficiently exploit the available cooperative diversity in the relay and CMA channels. In particular, our work considers the following distinct scenarios:

1. For the amplify and forward (AF) relay channel, we present a lattice-coded cooperation scheme which generalizes Yang and Belfior implementation [12] of the nonorthogonal amplify and forward (NAF) protocol [1], while keeping the decoding complexity relatively low. This generalization is shown to offer significant performance gains for moderate-to-large block lengths.

2. For the decode and forward (DF) relay channel, we first devise a novel variant of the dynamic decode and forward (DDF) protocol [1] which, through the judicious use of orthogonal space-time constellations, reduces the channel seen by the destination to a single-input single-output (SISO) time-selective channel. This variant achieves the excellent DMT of the DDF protocol, while reducing the decoding complexity at the destination. We further modify this variant by limiting the relay to start transmission only at a finite number of time instants. As argued in the sequel, this modification allows for a significant reduction in complexity while still achieving a near-optimal DMT. We then present a lattice-coded implementation of this variant and evaluate its performance through a comprehensive simulation study.

3. For the CMA channel, we present a lattice-coded implementation of the DMT-optimal nonorthogonal amplify and forward (CMA-NAF) protocol proposed in [1]. This implementation employs the minimum mean square error decision feedback equalizer (MMSE-DFE) Fano decoder of [13], to achieve near maximum-likelihood (ML) performance at a much lower complexity.

In summary, our contributions in this paper are twofold. First, we establish the utility of the information-theoretically optimal cooperation protocols proposed in [1]. This goal is accomplished through devising low-complexity implementations of the NAF, DDF and CMA-NAF protocols, and establishing their significant advantage over the recently proposed protocol of Yang and Belfiore. Second, we elucidate the underlying performance-complexity tradeoff in outage-limited cooperation channels.

The rest of the paper is organized as follows. Section 3 gives the notation and system model used throughout the paper. Section 4 briefly reviews the lattice coding framework adopted in our work. The proposed implementations for the NAF and DDF relay protocols are detailed in Section 5. Section 6 is devoted to our implementation of the CMA-NAF protocol. Finally, we offer some concluding remarks in Section 7.

3 Notation and System Model

Throughout the paper we use bold lowercase characters to denote vectors (e.g., $\mathbf{x}$) and bold uppercase characters to denote matrices (e.g., $\mathbf{H}$). We use superscript c to highlight complex vectors and matrices (e.g., $\mathbf{x}^c$ and $\mathbf{H}^c$). $\mathbb{Z}$, $\mathbb{Z}_Q$, $\mathbb{R}$ and $\mathbb{C}$ refer to the ring of integers, ring of integers modulo Q (with Q denoting a prime), field of real numbers, and field of complex numbers, respectively. We denote the identity matrix of dimension M by $\mathbf{I}_M$ and Kronecker product by $\otimes$. We also use $(x)^+$ to mean $\max\{x, 0\}$, $(x)^-$ to mean $\min\{x, 0\}$ and $\lceil x \rceil$ to mean nearest integer to x towards plus infinity.

Next we state the assumptions that apply to the two channels considered in this paper (i.e., relay and CMA channels). Assumptions pertaining to a specific channel will be given in the related section. All channels are assumed to be flat Rayleigh-fading and quasi-static, i.e., the channel gains remain constant during one codeword and change independently from one codeword to the next. Furthermore, the channel gains are mutually independent with unit variance. The additive noises at different nodes are zero-mean, mutually-independent, circularly-symmetric, and white complex-Gaussian. The variances of these noises are proportional to one another such that there are always *fixed* offsets between the different SNRs. In particular, we define the ratio of the destination noise variance to that of the relay (or cooperating node in

the CMA case) by c, i.e., $c \triangleq \sigma_v^2/\sigma_w^2$, where σ_v^2 denotes the noise variance at the destination and σ_w^2 the noise variance at the relay (or cooperating node). All nodes have the same power constraint, have a single antenna, and operate synchronously. Only the receiving node of any link knows the channel gain; no feedback to the transmitting node is permitted. All cooperating partners operate in the half-duplex mode, i.e., at any point in time, a node can either transmit or receive, but not both. This constraint is motivated by, e.g., the typically large difference between the incoming and outgoing signal power levels.

In the following we summarize some of the definitions that are used throughout the paper.

1. The SNR of a link, ρ, is defined as

$$\rho \triangleq \frac{E}{\sigma^2}, \tag{1}$$

where E denotes the average energy available for transmission of a symbol across the link and σ^2 denotes the variance of the noise observed at the receiving end of the link. We say that $f(\rho)$ is *exponentially equal to* ρ^b, denoted by $f(\rho)\dot{=}\rho^b$, when

$$\lim_{\rho\to\infty} \frac{\log f(\rho)}{\log \rho} = b. \tag{2}$$

In (2), b is called the *exponential order* of $f(\rho)$. $\dot{\leq}$ and $\dot{\geq}$ are defined similarly.

2. Assuming that g is a complex Gaussian random variable with zero mean and unit variance, the exponential order of $1/|g|^2$ is defined as

$$v \triangleq -\lim_{\rho\to\infty} \frac{\log |g|^2}{\log \rho}. \tag{3}$$

The probability density function (PDF) of v has the following property [14]

$$p_v \dot{=} \begin{cases} \rho^{-\infty} = 0, & \text{for } v < 0, \\ \rho^{-v}, & \text{for } v \geq 0 \end{cases}. \tag{4}$$

3. The input-output relationship of the cooperation schemes considered in our work will be re-casted, in the appropriate sections, as a generalized version of the point-to-point quasi-static flat Rayleigh-fading multiple-input multiple-output (MIMO) channel. For a MIMO point-to-point channel with M transmit and N receive antennas, the received signal at the destination is given by

$$\mathbf{y}_t^c = \sqrt{\frac{\rho}{M}}\mathbf{H}^c\mathbf{x}_t^c + \mathbf{z}_t^c, \quad t = 1, \cdots, T \tag{5}$$

where $\mathbf{H}^c \in \mathbb{C}^{N\times M}$ denotes the complex channel gain matrix with entries $h_{i,j} \sim \mathcal{N}_C(0,1)$, $\mathbf{X}^c = [\mathbf{x}_1^c, \cdots, \mathbf{x}_T^c]$ (in $\mathbb{C}^{M\times T}$) denotes the transmitted space-time codeword of block length T, $\mathbf{z}_t^c \sim \mathcal{N}_C(\mathbf{0}, \mathbf{I}_N)$ denotes the noise at the destination and ρ denotes the SNR observed at any of the receive antennas. The real representation of this MIMO channel is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{6}$$

where

$$\mathbf{y} \triangleq \sqrt{2} \times [\mathrm{Re}\{\mathbf{y}_1^c\}^T, \mathrm{Im}\{\mathbf{y}_1^c\}^T, \cdots, \mathrm{Re}\{\mathbf{y}_T^c\}^T, \mathrm{Im}\{\mathbf{y}_T^c\}^T]^T, \tag{7}$$

$$\mathbf{x} \triangleq [\mathrm{Re}\{\mathbf{x}_1^c\}^T, \mathrm{Im}\{\mathbf{x}_1^c\}^T, \cdots, \mathrm{Re}\{\mathbf{x}_T^c\}^T, \mathrm{Im}\{\mathbf{x}_T^c\}^T]^T, \tag{8}$$

and

$$\mathbf{H} \triangleq \mathbf{I}_T \otimes \sqrt{\frac{2\rho}{M}} \times \begin{bmatrix} \mathrm{Re}\{\mathbf{H}^c\} & -\mathrm{Im}\{\mathbf{H}^c\} \\ \mathrm{Im}\{\mathbf{H}^c\} & \mathrm{Re}\{\mathbf{H}^c\} \end{bmatrix}. \tag{9}$$

Notice that in (6), $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{H} \in \mathbb{R}^{n\times m}$, $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, where $n \triangleq 2NT$, $m \triangleq 2MT$.

4. Our work relies heavily on the notion of diversity-multiplexing tradeoff (DMT) posed by Zheng and Tse in [14]. In this formulation, we consider a family of codes (one for every SNR ρ), such that the code corresponding to ρ has a rate of $R(\rho)$ bits per channel use (BPCU) and error probability $P_e(\rho)$. For this family, the multiplexing gain r and the diversity gain d are defined as

$$r \triangleq \lim_{\rho\to\infty} \frac{R(\rho)}{\log\rho}, \qquad\qquad d \triangleq -\lim_{\rho\to\infty} \frac{\log P_e(\rho)}{\log\rho}. \tag{10}$$

For the point-to-point MIMO channel with M transmit and N receive antennas, Zheng and Tse showed that, for any $r \leq \min\{M,N\}$, the optimal diversity gain $d^*(r)$ is given by the piecewise linear function joining the (r,d) pairs $(k,(M-k)(N-k))$ for $k = 0, ..., \min\{M,N\}$, provided that the code-length T satisfies $T \geq M+N-1$ [14].

5. We say that protocol A *uniformly dominates* protocol B if, for any multiplexing gain r, $d_A(r) \geq d_B(r)$.

6. We say that protocol A is *Pareto optimal*, if there is no protocol B dominating protocol A in the Pareto sense. Protocol B is said to dominate protocol A in the Pareto sense if there is some r_0 for which $d_B(r_0) > d_A(r_0)$, but no r such that $d_B(r) < d_A(r)$.

4 Lattice Coding and Decoding

An m-dimensional lattice $\Lambda \subset \mathbb{R}^m$ is the set of points

$$\Lambda \triangleq \{\boldsymbol{\lambda} = \mathbf{G}\mathbf{u} \mid \mathbf{u} \in \mathbb{Z}^m\}, \tag{11}$$

where $\mathbf{G} \in \mathbb{R}^{m \times m}$ denotes the lattice generator matrix. Let $\boldsymbol{\eta} \in \mathbb{R}^m$ be a vector and $\mathcal{R}$ a measurable region in $\mathbb{R}^m$. The lattice code $\mathcal{C}(\Lambda, \eta, \mathcal{R})$ is defined as the subset of lattice translate $\Lambda + \boldsymbol{\eta}$, inside the *shaping region* $\mathcal{R}$ [15], i.e.,

$$\mathcal{C}(\Lambda, \boldsymbol{\eta}, \mathcal{R}) \triangleq (\Lambda + \boldsymbol{\eta}) \cap \mathcal{R}. \tag{12}$$

Here, we focus on construction-A lattice codes [16]. In this construction $\Lambda = C + Q\mathbb{Z}^m$, where $C \subseteq \mathbb{Z}_Q^m$ denotes a linear code over $\mathbb{Z}_Q$. The generator matrix of Λ is given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{P} & Q\mathbf{I} \end{bmatrix}, \tag{13}$$

where $[\mathbf{I}, \mathbf{P}^{\mathrm{T}}]^{\mathrm{T}}$ is the generator matrix of C, in a systematic form [16]. The shaping region $\mathcal{R}$ of the lattice code $\mathcal{C}(\Lambda, \boldsymbol{\eta}, \mathcal{R})$ can be 1) the m-dimensional sphere, 2) the fundamental Voronoi region of a sublattice $\Lambda' \subset \Lambda$, or 3) the m-dimensional hypercube. These three alternatives provide a tradeoff between performance and encoding complexity. The spherical shaping approach yields the maximum shaping gain but suffers from the largest encoding complexity. Encoding is typically done via look-up tables which limits this approach to short block lengths (since the number of entries in the table grows exponentially with the block length). Voronoi shaping offers a nice compromise between performance and complexity. In this approach, the encoding complexity is equivalent to that of m-dimensional vector quantization which, loosely speaking, ranges from polynomial to linear (in the block length), depending on the choice of the shaping lattice. On the other hand, with large enough m and an appropriate choice of Λ', the shaping gain approaches that of the sphere. The third alternative, i.e., hypercubic shaping, allows for the lowest encoding complexity, however, at the expense of a performance loss (i.e., the worst shaping gain among the three alternatives). As a side benefit, hypercubic shaping also has a low peak-to-average power ratio. The final ingredient in lattice coding is the translate vector $\boldsymbol{\eta}$ which is used to maximize the number of lattice points inside $\mathcal{R}$ and/or randomize the distribution of the codebook over $\mathcal{R}$ [15, 17]. In fact, Voronoi coding with $\boldsymbol{\eta}$ distributed uniformly over the Voronoi cell of Λ' corresponds to the mod-Λ approach of Erez and Zamir [17]. The proposed coding/decoding approaches in our work can be coupled with any shaping technique and any choice of the translate $\boldsymbol{\eta}$. The optimization of these parameters is beyond the scope of

this paper, and hence, will not be considered further. For simplicity of presentation and implementation, we focus in our simulation study on hypercubic shaping.

The primary appeal of lattice codes, in our framework, stems from their amenability to a low complexity decoding architecture, as argued in the sequel. For decoding purposes, we express $\mathcal{C}(\Lambda, \boldsymbol{\eta}, \mathcal{R})$ as the set of points $\mathbf{x}$ given by

$$\mathbf{x} = \mathbf{G}\mathbf{u} + \boldsymbol{\eta}, \qquad \mathbf{u} \in \mathcal{U} \tag{14}$$

where $\mathcal{U} \subset \mathbb{Z}^m$ is called the code *information set*. Assuming that the code is transmitted over a MIMO channel, the input-output relation becomes (recall (6))

$$\mathbf{y} = \mathbf{H}\mathbf{G}\mathbf{u} + \mathbf{H}\boldsymbol{\eta} + \mathbf{z}. \tag{15}$$

In the coherent paradigm where $\mathbf{H}$ is known to the receiver, the ML decoding rule reduces to

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u} \in \mathcal{U}} |\mathbf{y} - \mathbf{H}\boldsymbol{\eta} - \mathbf{H}\mathbf{G}\mathbf{u}|^2 \tag{16}$$

The optimization problem in (16) can be viewed as a *constrained* version of the closest lattice point search (CLPS) with lattice generator matrix given by $\mathbf{HG}$ and constraint set $\mathcal{U}$ [13]. This observation inspired the class of *sphere decoding* algorithms (e.g., [18]). More recently, a unified tree search decoding framework which encompasses the sphere decoders as special cases was proposed in [13]. Of particular interest to our work here is the (MMSE-DFE) Fano decoder discovered in [13]. In this decoder, we first *preprocess* the channel matrix via the feedforward filter of the MMSE-DFE, and then we apply the well-known Fano tree search algorithm to identify the closest lattice point. In particular, we attempt to approximate the optimal solution for

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u} \in \mathbf{Z}^m} |\mathbf{F}\mathbf{y} - \mathbf{R}\boldsymbol{\eta} - \mathbf{R}\mathbf{G}\mathbf{u})|^2, \tag{17}$$

where $\mathbf{F}$ and $\mathbf{R}$ are the feedforward and feedback filters of the MMSE-DFE, respectively. It is important to note the expanded search space in (17), i.e., $\mathcal{U}$ in (16) is replaced with $\mathbb{Z}^m$ in (17). This relaxation results in a significant reduction in complexity since, for non-trivial lattice codes, enforcing the boundary control $\mathbf{u} \in \mathcal{U}$ can be computationally intensive [13]. While the search space expansion is another source for suboptimality, it is shown in [13, 19] that the loss in performance is negligible *conditioned* that the MMSE-DFE preprocessing is employed (i.e., with $\mathbf{F} = \mathbf{I}$ one would see a significant performance loss). One of our main contributions is showing that the MMSE-DFE Fano decoder yields an excellent performance-vs-complexity tradeoff when appropriately used in cooperative channels. For more details about this decoder, the reader is referred to [13].

5 The Relay Channel

For exposition purposes, we limit our discussions below to the single relay scenario.

5.1 Amplify and Forward (AF) Cooperation

In [1], the non-orthogonal amplify and forward (NAF) strategy was shown to achieve the optimal DMT within the class of AF protocols. In NAF relaying, the source transmits on every symbol-interval in a cooperation-frame, where a cooperation-frame is defined as two consecutive symbol-intervals. The relay, on the other hand, transmits only once per cooperation-frame; it simply repeats the (noisy) signal it observed during the previous symbol-interval. It is clear that this design is dictated by the half-duplex constraint, which implies that the relay can repeat at most once per cooperation-frame. We denote the repetition gain by b and, for cooperation-frame k, the information symbols by $\{x_{j,k}\}_{j=1}^{2}$. The signals received by the destination during cooperation-frame k are given by:

$$y_{1,k} = g_1 x_{1,k} + v_{1,k}, \tag{18}$$

$$y_{2,k} = g_1 x_{2,k} + g_2 b(h x_{1,k} + w_{1,k}) + v_{2,k}. \tag{19}$$

Note that, in order to decode the message, the destination needs to know the relay repetition gain b, the source-relay channel gain h, the source-destination channel gain g_1, and the relay-destination channel gain g_2. The following result from [1] states the DMT achieved by this protocol.

Theorem 1 *([1]) The DMT achieved by the NAF relay protocol is*

$$d(r) = 1 - r + (1 - 2r)^{+}. \tag{20}$$

In [1], the achievability of (20) is established through using an ensemble of long Gaussian codebooks that span infinitely many cooperation-frames (with the same channel coefficients). More recently, Yang and Belfiore proposed a novel scheme that achieves the tradeoff in (20) by only coding over two cooperation-frames. Yang and Belfiore design is inspired by the fact that the input-output relationship of the NAF, as given by (18) and (19), corresponds to the following 2×2 MIMO channel

$$\mathbf{y}_k^c = \mathbf{H}^c \mathbf{x}_k^c + \mathbf{z}_k^c, \tag{21}$$

where $\mathbf{y}_k^c \triangleq [y_{1,k}, \sqrt{c/(|g_2 b|^2+c)}y_{2,k}]^T$, $\mathbf{x}_k^c \triangleq [x_{1,k}, x_{2,k}]^T$, $\mathbf{z}_k^c \sim \mathcal{N}_C(\mathbf{0}, \sigma_v^2\mathbf{I}_2)$ and

$$H^c \triangleq \begin{bmatrix} g_1 & 0 \\ \sqrt{\frac{c}{|g_2 b|^2+c}}g_2 bh & \sqrt{\frac{c}{|g_2 b|^2+c}}g_1 \end{bmatrix}.$$

Thus the NAF input-output relation, this time over two consecutive cooperation frames, is given by

$$\mathbf{y}^c = (\mathbf{I}_2 \otimes \mathbf{H}^c)\mathbf{x}^c + \mathbf{z}^c, \tag{22}$$

where $\mathbf{y}^c \triangleq [(\mathbf{y}_k^c)^T, (\mathbf{y}_{k+1}^c)^T]^T$, $\mathbf{x}^c \triangleq [(\mathbf{x}_k^c)^T, (\mathbf{x}_{k+1}^c)^T]^T$ and $\mathbf{z}^c \sim \mathcal{N}_C(\mathbf{0}, \sigma_v^2\mathbf{I}_4)$. One can now use any of the 2×2 linear dispersion (LD) constellations [20] as a cooperative coding scheme. A 2×2 LD constellation point $\mathbf{x}^c \in \mathbb{C}^4$ is obtained by multiplying a 4×1 QAM vector $\mathbf{u}^c$ by a generator matrix $\mathbf{G}^c$, i.e., $\mathbf{x}^c = \mathbf{G}^c\mathbf{u}^c$. As shown in [12], setting the generator matrix to that of the so called Golden constellation, i.e.,

$$\mathbf{G}_{gc}^c = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha & \alpha\theta & 0 & 0 \\ 0 & 0 & i\bar{\alpha} & i\bar{\alpha}\bar{\theta} \\ 0 & 0 & \alpha & \alpha\theta \\ \bar{\alpha} & \bar{\alpha}\bar{\theta} & 0 & 0 \end{bmatrix} \tag{23}$$

where $\theta = \frac{1+\sqrt{5}}{2}, \bar{\theta} = 1-\theta, \alpha = 1+i\bar{\theta}$ and $\bar{\alpha} = 1+i\theta$, guarantees the achievability of (20). This is due to the non-vanishing determinant property of the Golden code. Thus

$$\mathbf{y}^c = (\mathbf{I}_2 \otimes \mathbf{H}^c)\mathbf{G}_{gc}^c\mathbf{u}^c + \mathbf{z}^c. \tag{24}$$

One can now use our lattice decoding framework by transforming this last expression to its real representation (recall (6)). This reduces (24) to our model in (15), where $\mathcal{U} = \mathbb{Z}_Q^8$ corresponds to a 4×1 Q^2-QAM input constellation.

In applications where the codeword is allowed to span multiple cooperation-frames, The Yang-Belfiore Golden constellation fails to exploit the long block length to improve the coding gain. By appealing to our lattice coding framework, however, one can improve the coding gain while still enjoying a low complexity decoder. For example, we can concatenate the inner Golden constellation with an outer trellis code whose constraint length is allowed to increase with the block length. The improved coding gain of this approach translates into lower frame error rates, as validated by the numerical results in Section 5.3. In our design, we use a systematic convolutional code (CC) over $\mathbb{Z}_Q$ as the outer trellis code. Assuming that a codeword spans N consecutive cooperation-frames (N even), the received vector can be expressed as (compare to

(22))

$$\mathbf{y}^c = (\mathbf{I}_N \otimes \mathbf{H}^c)\mathbf{x}^c + \mathbf{z}^c,$$

where $\mathbf{y}^c \triangleq [(\mathbf{y}_k^c)^T, \cdots, (\mathbf{y}_{k+N}^c)^T]^T$, $\mathbf{x}^c \triangleq [(\mathbf{x}_k^c)^T, \cdots, (\mathbf{x}_{k+N}^c)^T]^T$ and $\mathbf{z}^c \sim \mathcal{N}_C(\mathbf{0}, \sigma_v^2\mathbf{I}_{2N})$. In this expression, $\mathbf{x}^c$ represents the output of the concatenated code (inner Golden constellation with an outer CC), i.e.,

$$\mathbf{x}^c = (\mathbf{I}_{N/2} \otimes \mathbf{G}_{gc}^c)(\mathbf{G}_{cc}^c\mathbf{u}^c + \eta^c),$$

where $\mathbf{G}_{cc}^c\mathbf{u}^c + \eta^c$ denotes the output of the convolutional encoder. This is based on the fact that a CC can be viewed as a construction-A lattice code with hypercubic shaping and a generator matrix $\mathbf{G}_{cc}^c$ [13]. Now

$$\begin{aligned}\mathbf{y}^c &= (\mathbf{I}_N \otimes \mathbf{H}^c)(\mathbf{I}_{N/2} \otimes \mathbf{G}_{gc}^c)(\mathbf{G}_{cc}^c\mathbf{u}^c + \eta^c) + \mathbf{z}^c, \\ &= \tilde{\mathbf{H}}^c\mathbf{G}_{cc}^c\mathbf{u}^c + \tilde{\mathbf{H}}^c\eta^c + \mathbf{z}^c, \end{aligned} \tag{25}$$

where

$$\begin{aligned}\tilde{\mathbf{H}}^c &\triangleq (\mathbf{I}_N \otimes \mathbf{H}^c)(\mathbf{I}_{N/2} \otimes \mathbf{G}_{gc}^c) \\ &= \mathbf{I}_{N/2} \otimes \Big((\mathbf{I}_2 \otimes \mathbf{H}^c)\mathbf{G}_{gc}^c\Big).\end{aligned}$$

Again, our lattice decoding framework (refer to (15)) can be directly applied to (25), after transforming it to its real representation (recall (6)).

5.2 Decode and Forward (DF) Cooperation

To the best of our knowledge, within the class of DF strategies, the dynamic decode and forward (DDF) protocol achieves the best DMT [1] (the same protocol was independently discovered in different contexts [21, 22]). This motivates developing low complexity variants of this strategy that are particularly suited for implementation. In the sequel we take a step by step approach where, a number of lemmas that characterize the modifications needed for complexity reduction while maintaining a good performance, are derived. For the sake of completeness, we first describe the DDF protocol.

In the DDF protocol the source transmits data at a rate of R bits per channel use (BPCU) during every symbol-interval in the codeword. A codeword is defined as M consecutive sub-blocks, where each sub-block is composed of T symbol-intervals. All channel gains are assumed to remain fixed during the length

of a codeword. The relay, on the other hand, listens to the source for enough number of sub-blocks such that the mutual information between its received signal and source signal exceeds MTR. It then decodes and re-encodes the message using an independent codebook and transmits the encoded symbols for the rest of the codeword. We denote the signals transmitted by the source and relay by $\{x_k\}_{k=1}^{MT}$ and $\{\tilde{x}_k\}_{k=M'T+1}^{MT}$, respectively, where M' is the number of sub-blocks that the relay waits before starting transmission. Using this notation, the received signal (at the destination) can be written as

$$y_k = \begin{cases} g_1 x_k + v_k & \text{for} \quad M'T \geq k \geq 1 \\ g_1 x_k + g_2 \tilde{x}_k + v_k & \text{for} \quad MT \geq k > M'T \end{cases},$$

where g_1 and g_2 denote the source-destination and relay-destination channel gains, respectively. It is now clear that the number of sub-blocks that the relay listens, should be chosen according to

$$M' = \min \left\{ M, \left\lceil \frac{MR}{\log_2 \left(1 + |h|^2 c\rho\right)} \right\rceil \right\}, \tag{26}$$

where h is the source-relay channel gain. In this expression, $c = \sigma_v^2/\sigma_w^2$ denotes the ratio of the destination noise variance, to that of the relay. The following result from [1], describes the DMT achievable by the DDF protocol as $T \to \infty$ and $M \to \infty$.

Theorem 2 *([1]) The DMT achieved by the DDF protocol is given by*

$$d(r) = \begin{cases} 2(1-r) & \textit{if} \quad \frac{1}{2} \geq r \geq 0 \\ (1-r)/r & \textit{if} \quad 1 \geq r \geq \frac{1}{2} \end{cases}. \tag{27}$$

In [1], the achievability result in (27) was established using independent and random codebooks at the source and relay nodes. This approach may not be practically feasible due to the prohibitive decoding complexity required at the destination. Allowing the relay node to start transmission at the beginning of any sub-block (based on the instantaneous value of the source-relay channel gain), is another potential source for complexity. In practice, this requires the source to use a very high-dimensional constellation (with a very low code-rate) to ensure that the information stream is uniquely decodable, even after one sub-block, given that the source-relay channel is good enough. It also impacts the amount of overhead in the relay-destination packet, since the destination needs to be informed of the starting time of the relay. In the sequel, we introduce two modifications of the original DDF protocol that aim to lower the complexities associated with these two aspects.

1. After successfully decoding, the relay can correctly anticipate the future transmissions from the source (i.e., x_k for $MT \geq k \geq M'T+1$) since it knows the source codebook. Based on this knowledge, the relay implements the following scheme, i.e.

$$\tilde{x}_k = \begin{cases} x_{k+1}^* & \text{for} \quad k = M'T+1, M'T+3, \cdots \\ -x_{k-1}^* & \text{for} \quad k = M'T+2, M'T+4, \cdots \end{cases}, \text{ and} \tag{28}$$

which reduces the signal seen by the destination for $MT \geq k \geq M'T+1$ to an Alamouti constellation.

2. We allow the relay to transmit only after the codeword is halfway through, i.e., we replace the rule in (26) with

$$M' = \min\left\{M, \max\left\{\frac{M}{2}, \left\lceil \frac{MR}{\log_2\left(1+|h|^2 c\rho\right)} \right\rceil\right\}\right\}, \tag{29}$$

Fortunately, these modifications do not entail any loss in performance (at least from the DMT perspective) as formalized in the following lemma.

Lemma 3 *The modified DDF protocol (with modifications given by* (28) *and* (29)*), still achieves the DMT of Theorem 2.*

Proof: To prove the first part of the lemma, let us denote the destination received signals by $\{y_k\}_{k=1}^{MT}$. Then

$$y_k = \begin{cases} g_1 x_k + v_k & \text{for} \quad k = 1, \cdots, M'T \\ g_1 x_k + g_2 x_{k+1}^* + v_k & \text{for} \quad k = M'T+1, M'T+3, \cdots \\ g_1 x_k - g_2 x_{k-1}^* + v_k & \text{for} \quad k = M'T+2, M'T+4, \cdots \end{cases}$$

Now, through linear processing of $\{y_k\}_{k=1}^{MT}$, the destination derives $\{\tilde{y_k}\}_{k=1}^{MT}$ such that

$$\tilde{y}_k = \begin{cases} g_1 x_k + \tilde{v}_k & \text{for} \quad k = 1, \cdots, M'T \\ \sqrt{|g_1|^2+|g_2|^2}\, x_k + \tilde{v}_k & \text{for} \quad k = M'T+1, \cdots, MT \end{cases}, \tag{30}$$

with $\tilde{v_k}$ being statistically identical to v_k. Using (30), it is straightforward to see that destination pairwise error probability, averaged over the ensemble of Gaussian codes (used by the source) and conditioned on a certain channel realization, is given by [1]

$$P_{pe|g_1,g_2,h} \leq \left(1+\frac{1}{2}|g_1|^2\rho\right)^{-M'T}\left(1+\frac{1}{2}(|g_1|^2+|g_2|^2)\rho\right)^{-(M-M')T}.$$

This last expression, however, is identical to that of the original DDF protocol (refer to [1]). This means that the DDF strategy, modified according to (28), achieves the DMT of Theorem 2 as $T \to \infty$ and $M \to \infty$. This completes the proof of the first part.

To prove the second part, we notice that the effect of constraining the relay to start transmission only after the codeword is halfway through, i.e., adopting (29), is to replace equation (52) in [1] with

$$\inf_{O^+,f} u = \begin{cases} 0 & \text{if} \quad f = \frac{1}{2} \\ (1 - \frac{r}{f})^+ & \text{if} \quad 1 \geq f > \frac{1}{2} \end{cases}, \tag{31}$$

where u denotes the exponential order of $1/|h|^2$. Now, since (31) is different from (52) in [1] only for $f = 1/2$, for which $\inf(v_1 + v_2)$ (v_1 and v_2 denote the exponential orders of $1/|g_1|^2$ and $1/|g_2|^2$, respectively) is already equal to the optimal value $2(1 - r)$ (refer to (49) in [1]), we conclude that this restriction does not affect the DMT achieved by the protocol. □

As indicated by (30), the channel seen by the destination in the modified DDF protocol is a time-selective SISO. This greatly reduces the decoding complexity at the destination, as it facilitates leveraging standard SISO decoding architectures (e.g., belief propagation, Viterbi/Fano decoders). In addition, restricting the relay to transmit only after $M' \geq M/2$ implies that the constellation size can be chosen such that the information stream is uniquely decodable only after $M' = M/2$.

The next result investigates the effect of limiting the relay to start transmission only at a finite number of time-instants. These time-instants partition the code word into $N+1$ segments which are not necessarily equal in length. We let the j-th segment ($N + 1 \geq j \geq 1$) span sub-blocks $M_{j-1} + 1$ through M_j, with $M_0 \triangleq 0$ and $M_{N+1} \triangleq M$. We further define the set of waiting fractions $\{f_j\}_{j=0}^{N+1}$ by $f_j \triangleq \frac{M_j}{M}$. Thus

$$f_0 = 0 < f_1 < \cdots < f_N < f_{N+1} = 1.$$

The question now is how to choose $\{f_j\}_{j=1}^{N}$, for a finite N, such that the protocol achieves the *optimal* DMT. The following lemma shows that this problem does not have a uniformly optimal solution and characterizes a Pareto optimal set of waiting fractions.

Lemma 4 *For the DDF protocol with a finite N,*

1. *there exists no* uniformly *dominant set of fractions* $\{f_j^u\}_{j=1}^{N}$.

2. let $f_1^p = \frac{1}{2}$ *and*

$$f_j^p = \frac{1 - f_{j-1}^p}{2 - (1 + \frac{1}{f_N^p}) f_{j-1}^p}, \quad \textit{for} \quad N \geq j > 1 \tag{32}$$

then the set of fractions $\{f_j^p\}_{j=1}^N$ *is Pareto optimal, with*

$$d^p(r) = 1 - r + (1 - \frac{r}{f_N^p})^+. \tag{33}$$

Proof: To prove the first part, we note that the outage set for the DDF protocol with a general set of waiting fractions, $\{f_j\}_{j=1}^N$, is still given by equation (49) in [1], i.e.

$$O^+ = \{(v_1, v_2, u) \in \mathbb{R}^{3+} | f(1 - v_1)^+ + (1 - f)(1 - \min\{v_1, v_2\})^+ \leq r\}.$$

The only difference is that f is now given by (compare with (52) of [1])

$$f = \begin{cases} f_1 & \text{if} \quad 1 - \frac{r}{f_1} > u \geq 0 \\ f_j & \text{if} \quad 1 - \frac{r}{f_j} > u \geq (1 - \frac{r}{f_{j-1}})^+ \\ 1 & \text{if} \quad u \geq (1 - \frac{r}{f_N})^+ \end{cases}. \tag{34}$$

Next, we split O^+ such that

$$O^+ = \cup_{j=1}^{N+1} O_j^+, \quad \text{where} \quad O_j^+ = \{(v_1, v_2, u) \in O^+ | f = f_j\}. \tag{35}$$

Now, from (34) and (35) we get

$$\inf_{O_j^+} u = \begin{cases} 0 & \text{for} \quad j = 1 \\ (1 - \frac{r}{f_{j-1}})^+ & \text{for} \quad N + 1 \geq j > 2 \end{cases}.$$

Also, since $f_j \geq \frac{1}{2}$, we have $\inf_{(v_1, v_2) \in O_j^+} (v_1 + v_2) = (1 - r)/f_j$ (refer to (55) in [1]). Thus

$$\begin{aligned} d_j(r) &\triangleq \inf_{O_j^+} (v_1 + v_2 + u), \\ d_j(r) &= \frac{1 - r}{f_j} + (1 - \frac{r}{f_{j-1}})^+. \end{aligned} \tag{36}$$

But, (35) along with (36) results in

$$\begin{aligned} d(r) &= \min_{f_j \geq r} d_j(r), \\ d(r) &= \min_{f_j \geq r} \frac{1 - r}{f_j} + (1 - \frac{r}{f_{j-1}})^+. \end{aligned} \tag{37}$$

Now let us assume that the set of waiting fractions $\{f_j^u\}_{j=1}^N$ is uniformly optimal. Pick $\{f_j\}_{j=1}^N$ such that $f_N^u < f_N < 1$. Then from (37) we conclude that for any $f_N^u < r < f_N$, $d^u(r) = 1 - r$ and $d(r) = 1 - r + 1 - r/f_N$. Thus $d^u(r) < d(r)$, which is in contradiction with the uniform optimality assumption of $\{f_j^u\}_{j=1}^N$. This completes the proof of the first part of the lemma.

To prove the second part, we observe that for $N \geq j \geq 1$, $\{f_j^p\}_{j=1}^N$ as given by (32) results in

$$d_{N+1}^p(r) < d_j^p(r), \text{ for } f_j \geq r \geq 0, r \neq f_{j-1} \tag{38}$$

and

$$\begin{aligned} d_{N+1}^p(f_{j-1}) &= d_j^p(f_{j-1}), \text{ or} \\ d_{N+1}^p(f_{j-1}) &= \frac{1 - f_{j-1}}{f_j}. \end{aligned} \tag{39}$$

Now (38), along with (37) and (36) proves that $\{f_j^p\}_{j=1}^N$ achieves (33). The only thing left is to show that $d^p(r)$ is indeed Pareto optimal, i.e., no other set $\{f_j\}_{j=1}^N$ dominates $\{f_j^p\}_{j=1}^N$, in the Pareto sense. To do so, we assume such a set exists and observe that since $f_0 = f_0^p = 0$ and $f_{N+1} = f_{N+1}^p = 1$, there should be $N + 1 \geq i \geq 1$ and $N + 1 \geq \ell \geq 1$ such that

$$f_{i-1} \leq f_{\ell-1}^p < f_\ell^p < f_i, \text{ or} \tag{40}$$

$$f_{i-1} < f_{\ell-1}^p < f_\ell^p \leq f_i. \tag{41}$$

Now, if (40) is true, then we observe from (37) that

$$d(f_{\ell-1}^p) \leq d_i(f_{\ell-1}^p) = \frac{1 - f_{\ell-1}^p}{f_i} < \frac{1 - f_{\ell-1}^p}{f_\ell^p} = d^p(f_{\ell-1}^p),$$

where we have used (39) in deriving the last step. This, however is in contradiction with the Pareto dominance of $\{f_j\}$, since

$$d(f_{\ell-1}^p) < d^p(f_{\ell-1}^p).$$

On the other hand, if (41) is true, then

$$d(f_{i-1}) \leq d_i(f_{i-1}) = \frac{1 - f_{i-1}}{f_i} < 2 - (1 + \frac{1}{f_N^p}) f_{i-1} = d^p(f_{i-1}),$$

or

$$d(f_{i-1}) < d^p(f_{i-1}),$$

which again is in contradiction with Pareto dominance of $\{f_j\}$. This establishes the Pareto optimality of $d^p(r)$, and thus completes the proof of the second part. $\square$

Figure 1 shows the DMT for Pareto optimal DDF protocols with $N = 1(\{\frac{1}{2}\})$, $N = 2(\{\frac{1}{2}, \frac{2}{3}\})$ and $N = \infty$. As evident from this figure, the set of waiting fractions $\{f_j^p\}$ (given by (32)), is *not* asymptotically optimal, i.e., the DMT of the Pareto optimal DDF protocol with $N = \infty$ differs from that of the original DDF. In contrast, the set of fractions $\{f_j^a\}$ given by

$$f_j^a \triangleq \frac{1}{2} + \frac{j-1}{2N}, \quad N \geq j \geq 1,$$

is asymptotically optimal. This can easily be verified by realizing that as N grows to infinity, the DDF with $\{f_j^a\}$ becomes identical to the original DDF, modified according to (29). We recall from Lemma 3 that this modification does not affect the achieved DMT. One of the reasons why we choose $\{f_j^p\}$ over, e.g., $\{f_j^a\}$, despite its lack of asymptotic optimality, is that it admits a simple closed-form DMT (refer to (33)). Another reason for preferring $\{f_j^p\}$ over, e.g, $\{f_j^a\}$, is that unlike $\{f_j^a\}$, by increasing N, $\{f_j^p\}$ provides a new DDF protocol which *uniformly* dominates the previous one, i.e.,

$$d_{N_1}^p(r) \geq d_{N_2}^p(r), \ \ \forall r, N_1 \geq N_2.$$

In this expression, $d_N^p(r)$ denotes the DMT of the Pareto optimal DDF protocol with $N+1$ segments. Another interesting observation regarding Figure 1 is that, the performance of the the DDF protocol is most enhanced with the first few added segments. This is true in terms of both the DMT and outage probability (refer to Figure 2). Finally, we observe that the modified DDF protocol effectively transforms the relay channel into a time-selective SISO (recall (30)) which makes the application of our lattice decoding framework straightforward. Hence, the corresponding details are omitted for the sake of brevity.

5.3 Numerical Results

Throughout this section, we consider construction-A lattice codes obtained from systematic convolutional codes (CCs) with generator polynomials over $\mathbb{Z}_Q$. The generator polynomials are of constraint length 4 and chosen at random (the optimization of generator polynomials is beyond the scope of this work). Unless otherwise stated, we choose the SNR level of the source-relay channel to be 3 dB higher than the SNR at the destination. In all scenarios, the MMSE-DFE Fano decoder uses a bias $b_F = 1.2$ and a step-size $\Delta = 5$ [13].

First, we demonstrate the performance improvements gained through augmenting the Golden constellation of [12] with lattice codes constructed from systematic CCs of rate $1/2$. The generator polynomials for these CC's are $(1\ 2\ 3\ 4\ 4;\ 1\ 0\ 0\ 0\ 0)$, $(1\ 9\ 11\ 15\ 13;\ 1\ 0\ 0\ 0\ 0)$ and $(1\ 49\ 1\ 30\ 66;\ 1\ 0\ 0\ 0\ 0)$, over $\mathbb{Z}_5$, $\mathbb{Z}_{17}$ and $\mathbb{Z}_{67}$, respectively. In both schemes, the frame length is chosen to be 128 symbols. We assume that the source and the relay transmit with equal power, and that the source transmission power is fixed throughout the entire frame. Figure 3 shows the frame error rate (FER) of the two coding schemes as a function of the average received SNR at the destination. In the Golden constellation scenario, modulating the information symbols using $4-$QAM, $16-$QAM and $64-$QAM leads to transmission rates of $2, 4$ and 6 BPCU, respectively. To achieve comparable (slightly higher) transmission rates in the augmented code scenario, we use primes $Q = 5, 17$ and 67, respectively. From Figure 3, we see that augmenting the Golden constellation with CC's provides a performance improvement of about $1 - 1.5$ dB. The inferior performance of the augmented code at 2 BPCU is due to the fact that its effective transmission rate, i.e., $\log_2 5 = 2.32$ BPCU, is significantly higher than the 2 BPCU offered by the Golden constellation. This choice is dictated by the MMSE-DFE Fano decoder's instrumental need for Q to be a prime, so that the received signal has a nice lattice representation. Figure 4 compares the complexity of the MMSE-DFE Fano decoder for the two schemes.

Next, we proceed to the low-complexity DDF protocol proposed in Section 5.2. At the source, the information stream is appended with 16 bits of cyclic redundancy check (CRC), and the resulting vector is encoded using a rate $1/4$ systematic CC. The frame length for the coded bit-stream is 128, as in the case of the NAF relay protocol. The generator polynomials for the CC over Z_{17} (2 BPCU) are chosen to be $(1\ 10\ 6\ 12\ 1;\ 1\ 0\ 16\ 11\ 6;\ 1\ 9\ 8\ 14\ 14;\ 1\ 0\ 0\ 0\ 0)$, while for the CC over Z_{67} (3 BPCU), we choose $(1\,65\,53\,4\,37;\ 1\,14\,66\,19\,58;\ 1\,20\,52\,19\,31;\ 1\,0\,0\,0\,0)$. The relay attempts decoding at the end of the M_j-th sub-block, where M_j is the smallest allowed number of sub-blocks such that the mutual information at the relay exceeds the received rate. In order to avoid error propagation at the destination, the decoded vector is checked for validity using the CRC bits. If the decoded vector is valid, i.e., it satisfies the CRC check, then the relay starts re-encoding and transmission. If the decoded stream at the relay does not satisfy the CRC check, then the relay attempts decoding after it receives the next allowed number of sub-blocks (M_{j+1}), and so on. At the receiver, we use the MMSE-DFE Fano decoder (note that the MMSE-DFE part is now a trivial scaling since the channel seen by the destination is a SISO). We also assume that the destination knows, via overhead bits, the time at which the relay starts transmitting. For the range of transmission rates considered in the sequel, it turns out that increasing the number of segments beyond 3 provides negligible

increase in performance. Figure 2 shows the outage probability of the low-complexity DDF variant, when the codeword is partitioned into $2, 3$ and 4 segments. As seen from the figure, the gap between the outage curves is negligible. Therefore, in the sequel, we consider only the variant of DDF relay protocol with 3 segments. Moreover, we choose the waiting fractions $\{f_j^p\}_{j=1}^2$ according to Lemma 4, i.e., $\{\frac{1}{2}, \frac{2}{3}\}$. Figure 5 compares the FER of the DDF variant with that of the NAF protocol for 2 and 3 BPCU. As seen from this figure, the proposed DDF strategy offers a gain of about 4 dB (and about 6 dB) over the NAF scheme, for 2 (and 3) BPCU. Note, however, that the DDF variant entails more complexity at the relay compared to the NAF protocol, since the relay needs to decode the information stream. Finally, we notice that although there are several implementations of AF and DF protocols in the literature (e.g., [7], [9], [8]), they all consider low transmission rates. It should be noted, however, that the difference in performance between the protocols that are spectrally efficient and those that are not, reveals itself only at high transmission-rate scenarios. Therefore, in this paper we focus on high transmission rates and do not compare against these works.

6 The Cooperative Multiple Access (CMA) Channel

In this section, we implement the CMA-NAF protocol for two users. We start with describing the protocol. In the CMA-NAF protocol, each of the two sources transmits once per cooperation-frame, where a cooperation-frame is defined by two consecutive symbol-intervals. Each source, when active, transmits a linear combination of the symbol it intends to send and the (noisy) signal it received from its partner during the last symbol-interval. For source j and cooperation-frame k, we denote the broadcast and repetition gains by a_j and b_j, respectively, the symbol to be sent by $x_{j,k}$, and the transmitted signal by $t_{j,k}$. At startup the transmitted signals will take the form

$$t_{1,1} = a_1 x_{1,1} \tag{42}$$

$$t_{2,1} = a_2 x_{2,1} + b_2(h t_{1,1} + w_{2,1}) \tag{43}$$

$$t_{1,2} = a_1 x_{1,2} + b_1(h t_{2,1} + w_{1,1}) \tag{44}$$

$$t_{2,2} = a_2 x_{2,2} + b_2(h t_{1,2} + w_{2,2}) \tag{45}$$

where h denotes the inter-source channel gain and $w_{j,k}$ the noise observed by source j during the cooperation-frame k. (We assume that $w_{j,k}$ has variance σ_w^2.) The corresponding signals received by the destination

are

$$y_{1,1} = g_1 t_{1,1} + v_{1,1} \tag{46}$$

$$y_{2,1} = g_2 t_{2,1} + v_{2,1} \tag{47}$$

$$y_{1,2} = g_1 t_{1,2} + v_{1,2} \tag{48}$$

$$y_{2,2} = g_2 t_{2,2} + v_{2,2} \tag{49}$$

where g_j is the gain of the channel connecting source j to the destination and $v_{j,k}$ the destination noise of variance σ_v^2. Note that, as mandated by our half-duplex constraint, no source transmits and receives simultaneously. The broadcast and repetition gains $\{a_j, b_j\}$ are (experimentally) chosen to minimize outage probability at the destination. As a consequence of symmetry, a_1 and a_2, as well as b_1 and b_2, will have the same optimal value. Thus, we assume that broadcast and repetition gains are the same at each source and omit the subscripts, yielding $\{a, b\}$. We assume the codewords of each source to be of length N. Notice that it takes $2N$ symbol-intervals, or equivalently N cooperation frames for the two sources to transmit their codewords. The following result from [1] gives the DMT achieved by this protocol.

Theorem 5 *([1]) The CMA-NAF protocol achieves the optimal DMT of the channel, i.e., for two users*

$$d(r) = 2(1 - r). \tag{50}$$

In order to apply our lattice decoding framework to the CMA-NAF protocol, we need to write its input-output relation in the form of (15). Towards this end, we exploit the linearity of the CMA-NAF protocol over the field of complex numbers, to describe the joint effect of lattice coding at the sources, and cooperation among them, by one extended generator matrix. This results in a typical setting in which the MMSE-DFE Fano decoder is expected to be efficient in recovering the two information streams jointly at the destination (this expectation is validated by our numerical results). In particular, by examining (42) through (49), we realize that the received signal at the destination can be written as

$$\tilde{\mathbf{y}}^c = \tilde{\mathbf{H}}^c \mathbf{x}^c + \mathbf{B}^c \mathbf{w}^c + \mathbf{v}^c, \tag{51}$$

where $\tilde{\mathbf{y}}^c \triangleq [y_{1,1}, y_{2,1}, \cdots, y_{1,N}, y_{2,N}]^T$ denotes the vector of received signals at the destination, and $\mathbf{x}^c \triangleq [x_{1,1}, x_{2,1}, \cdots, x_{1,N}, x_{2,N}]^T$ denotes the vector formed by multiplexing the two sources' codewords (i.e., $\mathbf{x}_j^c \triangleq [x_{j,1}, \cdots, x_{j,N}]^T, j \in \{1, 2\}$) in an alternate fashion. $\mathbf{w}^c \sim \mathcal{N}_C(\mathbf{0}, \sigma_w^2 \mathbf{I}_{2N-1})$ and $\mathbf{v}^c \sim \mathcal{N}_C(\mathbf{0}, \sigma_v^2 \mathbf{I}_{2N})$ denote noise vectors observed by the two sources and the destination, respectively. Finally, matrices

$\tilde{\mathbf{H}}^c \in \mathbb{C}^{2N\times 2N}$ and $\mathbf{B}^c \in \mathbb{C}^{2N\times(2N-1)}$ are given by

$$\tilde{\mathbf{H}}^c = a\mathbf{G}^c \begin{bmatrix} 1 & & & \\ bh & 1 & & \\ \vdots & \vdots & \ddots & \\ (bh)^{2N-1} & (bh)^{2N-2} & \cdots & 1 \end{bmatrix}, \tag{52}$$

and

$$\mathbf{B}^c = b\mathbf{G}^c \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ \vdots & \vdots & \ddots & \\ (bh)^{2N-3} & (bh)^{2N-4} & \cdots & 0 \\ (bh)^{2N-2} & (bh)^{2N-3} & \cdots & 1 \end{bmatrix}, \tag{53}$$

with $\mathbf{G}^c \triangleq \mathbf{I}_N \otimes \text{diag}(g_1, g_2)$. The noise component in (51), i.e. $\tilde{\mathbf{z}}^c \triangleq \mathbf{B}^c\mathbf{w}^c + \mathbf{v}^c$, is colored and needs to be whitened before our lattice decoding framework can be applied. Let us denote the covariance matrix of $\tilde{\mathbf{z}}^c$ by $\Sigma_{\tilde{\mathbf{z}}^c}$. Then

$$\Sigma_{\tilde{\mathbf{z}}^c} = \sigma_w^2\mathbf{B}^c\mathbf{B}^{cH} + \sigma_v^2\mathbf{I}.$$

Now, whitening is accomplished by multiplying both sides of (51) by $\Sigma_{\tilde{\mathbf{z}}^c}^{-\frac{1}{2}}$ (which can be computed using any of the standard methods, e.g., singular value decomposition, Cholesky decomposition), i.e.

$$\mathbf{y}^c = \mathbf{H}^c\mathbf{x}^c + \mathbf{z}^c. \tag{54}$$

In (54), $\mathbf{y}^c \triangleq \Sigma_{\tilde{\mathbf{z}}^c}^{-\frac{1}{2}}\tilde{\mathbf{y}}^c$, $\mathbf{H}^c \triangleq \Sigma_{\tilde{\mathbf{z}}^c}^{-\frac{1}{2}}\tilde{\mathbf{H}}^c$ and the noise vector $\mathbf{z}^c \triangleq \Sigma_{\tilde{\mathbf{z}}^c}^{-\frac{1}{2}}\tilde{\mathbf{z}}^c$ is now white with unit variance, i.e., $\mathbf{z}^c \sim \mathcal{N}_C(\mathbf{0}, \mathbf{I}_{2N})$. The next step in reducing (54) to the form of (15), is to establish the relation between the two sources' information streams $\mathbf{u}_j^c \triangleq [u_{j,1}, \cdots, u_{j,N}]^T, j \in \{1, 2\}$ and $\mathbf{x}^c$. For this purpose, let us denote the generator matrix and translate vector used by source j by $\mathbf{G_j^c}$ and η_j^c, i.e., $\mathbf{x}_j^c = \mathbf{G}_j^c\mathbf{u}_j^c + \eta_j^c$. To facilitate the *joint* decoding of the two sources, we define the joint information stream $\mathbf{u}^c$ in the same manner as we defined $\mathbf{x}^c$, i.e., $\mathbf{u}^c \triangleq [u_{1,1}, u_{2,1}, \cdots, u_{1,N}, u_{2,N}]^T$. We can now write

$$\mathbf{x}^c = \mathbf{G}_{eq}^c\mathbf{u}^c + \eta_{eq}^c, \tag{55}$$

where $\mathbf{G}_{eq}^c$ and η_{eq}^c are the *equivalent* lattice generator matrix and translate vector, as seen by the destination. $\mathbf{G}_{eq}^c$ and η_{eq}^c may be constructed through appropriate combining of $\{\mathbf{G}_j^c\}_{j=1}^2$ and $\{\eta_j^c\}_{j=1}^2$, respectively. Now plugging (55) in (54) and using its real representation (recall (6)), yields the desired relation (in the form of (15)), suitable for use in our lattice decoding framework.

6.1 Numerical Results

In this section, we demonstrate the excellent performance of the lattice-coded CMA-NAF protocol, through comparing it against other schemes. Furthermore, we elucidate the performance-complexity tradeoff achieved by our devised MMSE-DFE Fano decoder. As before, we take construction-A lattice codes (obtained from systematic CC's with generator polynomials over $\mathbb{Z}_Q$), as our coding scheme. The frame length is set to $N = 128$. At the destination, MMSE-DFE Fano decoding is used to jointly decode the two messages. We define the frame error event as $\{(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) \neq (\mathbf{u}_1, \mathbf{u}_2)\}$, where $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ refer to the decoded information streams. Figures 6 and 7 compare the FER performance of the lattice coded CMA-NAF protocol and the NAF relay protocol, for 2 BPCU and 4 BPCU, respectively. For the NAF relay protocol, the coding scheme achieving the better performance (among Golden constellation alone and Golden constellation augmented by a CC) is chosen. The figures also show the performance when the CMA-NAF protocol is used with uncoded QAM transmission. Both coded and uncoded transmission with the CMA-NAF protocol perform significantly better than the NAF-relay protocol. The performance gap between the two schemes widens as the transmission rate increases. This can be explained by the superior DMT of the CMA-NAF protocol, compared to the NAF-relay protocol. For comparison purposes, figures 6 and 7 also give the performance curves when the two sources do not cooperate, i.e., each source transmits independently. Here too, we consider both coded and uncoded transmission. From these figures, we observe that both coded and uncoded non-cooperative scheme exhibit poor performance at higher SNRs, as expected. Similar to Figure 3, the inferior performance of the lattice-coded CMA-NAF (2 BPCU), compared to the uncoded CMA-NAF protocol, can be explained by its significantly higher transmission rate ($Q = 5$). Figure 8 shows the bit error rate (BER) performance of the uncoded CMA-NAF protocol versus NAF-relay protocol with Golden constellation. Again, CMA-NAF protocol with uncoded transmission shows significant BER improvement of about 3 dB at 2 BPCU, and about 5 dB at 4 BPCU. Finally, figures 9 and 10 give the performance-complexity tradeoff achieved by the MMSE-DFE Fano decoder, in the lattice-coded CMA-NAF setup ($N = 64$). As it is obvious from these figures, one can reduce the complexity of the Fano decoder by increasing the bias value b_F. This complexity reduction, however, comes at the expense of a performance loss. In our simulations, we use $b_F = 1.2$ to achieve good performance with reasonable complexity.

7 Conclusions

We presented novel lattice-coded protocols for the half-duplex outage-limited cooperative channels. The proposed protocols exhibit attractive performance-complexity tradeoffs. In particular, for the AF relay channel, we generalized Yang and Belfiore implementation of the NAF protocol and demonstrated the resulting performance gain for moderate-to-large block lengths. For the DF relay channel, we first devised a novel variant of the DDF protocol which enjoys a comparable DMT to the original DDF with a much lower complexity. We then presented a lattice-coded implementation of this variant and evaluated its performance through simulation. For the CMA channel, we presented a lattice-coded implementation of the CMA-NAF protocol and demonstrated its excellent performance-complexity tradeoff. Our results establish the natural matching between the optimal CMA-NAF protocol and the MMSE-DFE Fano decoder.

References

[1] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4152–4172, Dec. 2005.

[2] K. Azarian, H. E. Gamal, and P. Schniter, "On the optimality of ARQ-DDF protocols," *submitted to the IEEE Trans. Inform. Theory*, Jan 2006.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part I. system description," *IEEE Trans. Communications*, Nov. 2003.

[4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part II. implementation aspects and performance analysis," *IEEE Trans. Communications*, Nov. 2003.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Info. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[6] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Communications*, Oct. 2003.

[7] A. Stefanov and E. Erkip, "Cooperative space-time coding for wireless networks," *IEEE Trans. Communications*, vol. 53, pp. 1804–1809, Nov. 2005.

[8] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *IEEE Trans. Signal Processing*, vol. 52, pp. 362–371, Feb. 2004.

[9] Z. Zhang and T. M. Duman, "Capacity-approaching turbo coding and iterative decoding for relay channels," *IEEE Trans. Communications*, vol. 53, pp. 1895–1905, Nov. 2005.

[10] U. Mitra and A. Sabharwal, "On achievable rates for complexity constrained relay channels," *Proc. 41st Allerton conf. on Communication, control and computing*, Oct. 2003.

[11] A. Host-Madsen, "A new achievable rate for cooperative diversity based on generalized writing on dirty paper," *presented at ISIT'03*, 2003.

[12] S. Yang and J. C. Belfiore, "Optimal space-time codes for the MIMO amplify and forward cooperative channel," *submitted to IEEE Trans. Info. Theory*, Sep. 2005.

[13] A. Murugan, H. E. Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding:Rediscovering the sequential decoder," *to appear in the IEEE Trans. Info. Theory*, March 2006.

[14] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Info. Theory*, vol. 49, pp. 1073 –1096, May 2003.

[15] G. D. F. Jr., "Coset codes-part i: Introduction and geometrical classification," *IEEE Trans. Info. Theory*, vol. 34, pp. 1123–1151, Sept. 1988.

[16] H.-A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.

[17] U. Erez and R. Zamir, "Lattice decoding can achieve $\frac{1}{2}\log(1+snr)$ on the awgn channel using nested codes," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2293–2314, Oct. 2004.

[18] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Info. Theory*, vol. 48, pp. 2201–2214, Aug. 2002.

[19] H. E. Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-vs-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Theory*, June 2004.

[20] B. Hassibi and B. Hochwald, “High rate codes that are linear in space and time,” *IEEE Trans. Info. Theory*, vol. 48, pp. 1804–824, July 2002.

[21] M. Katz and S. Shamai, “Transmitting to colocated users in wireless ad-hoc and sensory networks,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 3540–3563, Oct. 2005.

[22] P. Mitran, H. Ochiai, and V. Tarokh, “Space time diversity enhancements using collaborative communications,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 2041–2057, June 2005.

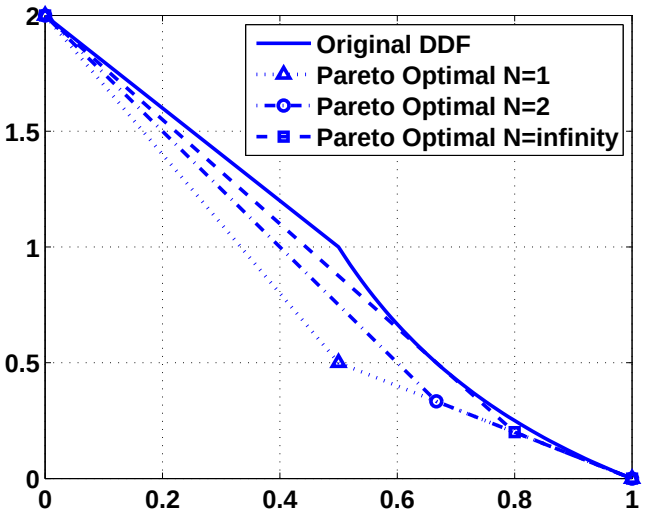

Figure 1: DMT of Pareto optimal DDF protocol with $N = 1, 2$ and ∞.

8 Biography

8.1 Arul Murugan

Arul D. Murugan received the B. Tech Degree in Electrical Engineering from Indian Institute of Technology, Madras in August 2001 and Ph. D. degree in Electrical Engineering from The Ohio State University, Columbus, Ohio, in June 2006. He is currently working as a senior design engineer at Marvell Tech. His research interests include channel coding and wireless communications.

8.2 Kambiz Azarian

Kambiz Azarian received the B.S. degree from Shahid Beheshti University, Tehran, Iran in 1996, the M.S. degree from Amirkabir University of Technology, Tehran, Iran in 1999, and Ph. D. from The Ohio State University, Columbus, Ohio, in 2006, all in Electrical Engineering. He is currently a post-doctoral fellow at the University of Notredame, Indiana. His areas of research include wireless communications and network information theory.

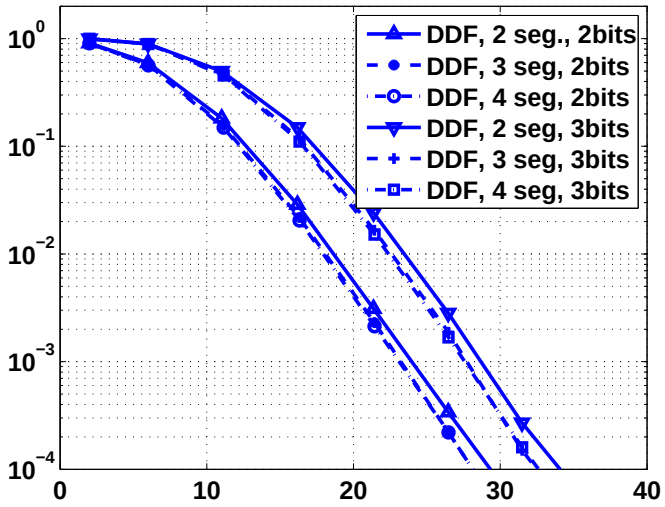

Figure 2: Outage probability of Pareto optimal DDF protocol with $2, 3$ and 4 segments.

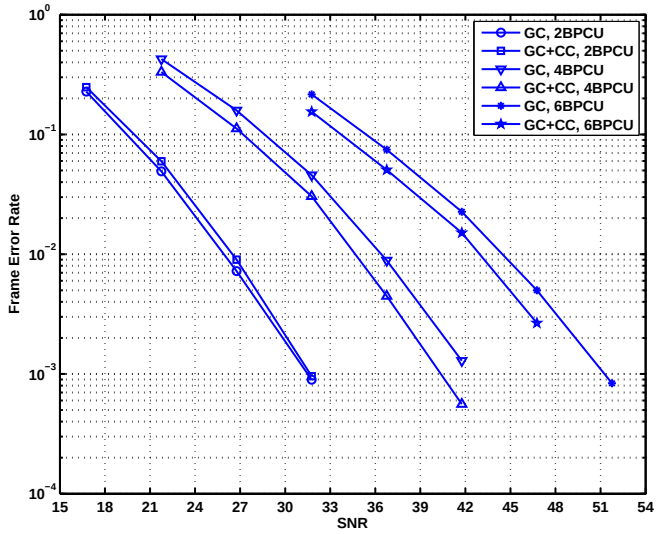

Figure 3: FER of NAF relay with Golden constellation, with and without CC outer code (denoted by GC+CC and GC).

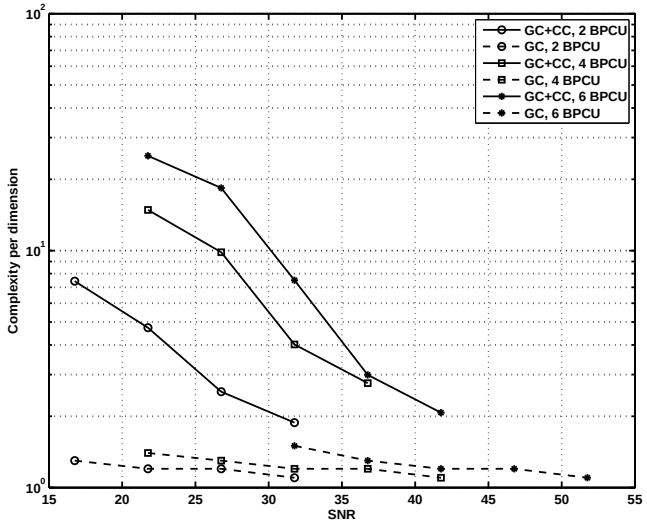

Figure 4: Complexity comparison of MMSE-DFE Fano decoder for the NAF relay with Golden constellation, with and without CC outer code (denoted by GC+CC and GC). The bias value is $b_F = 1.2$.

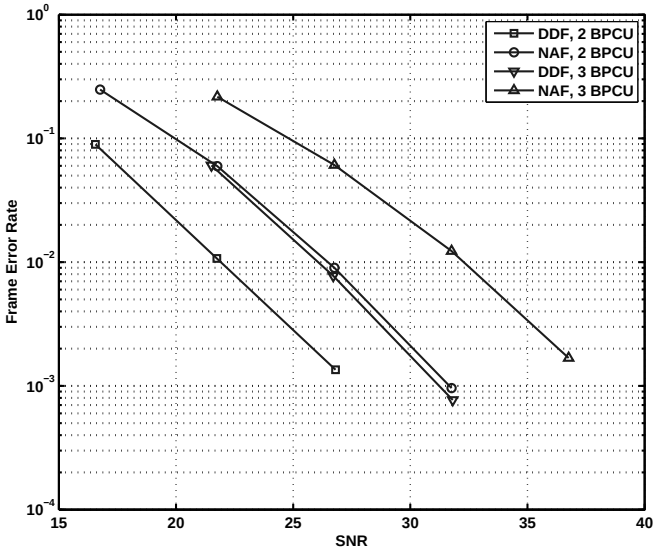

Figure 5: FER of DDF relay protocol (with 3 segments) versus NAF relay protocol.

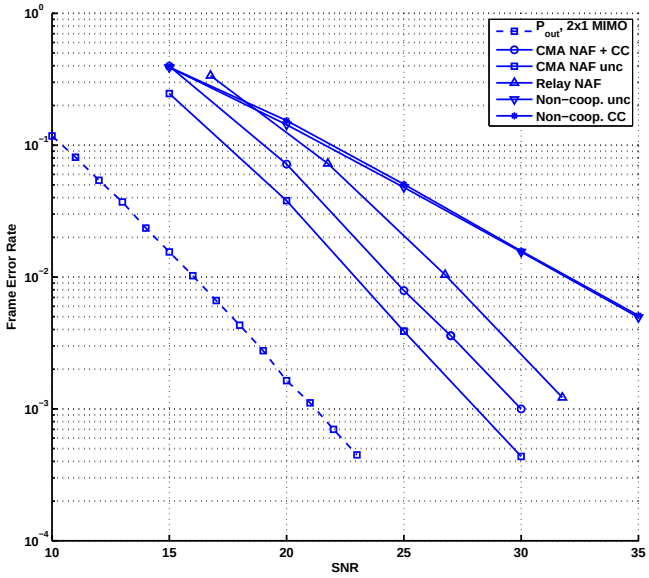

Figure 6: FER of CMA-NAF, Relay NAF and non-cooperative protocols (2 BPCU).

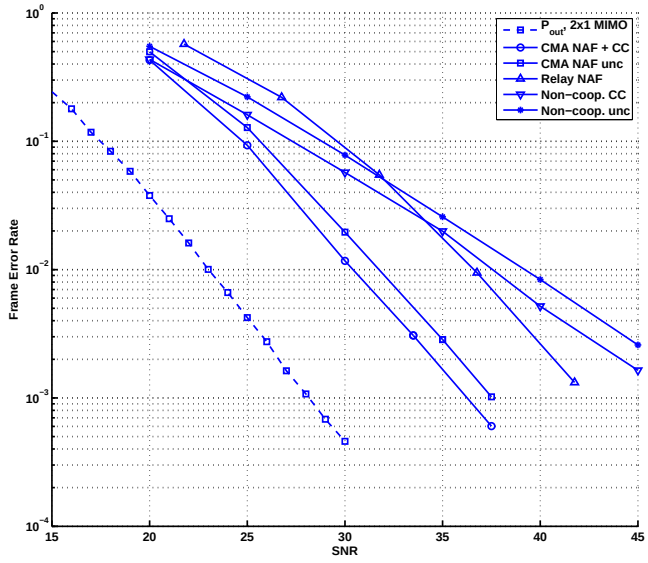

Figure 7: FER of CMA-NAF, Relay NAF and non-cooperative protocols (4 BPCU).

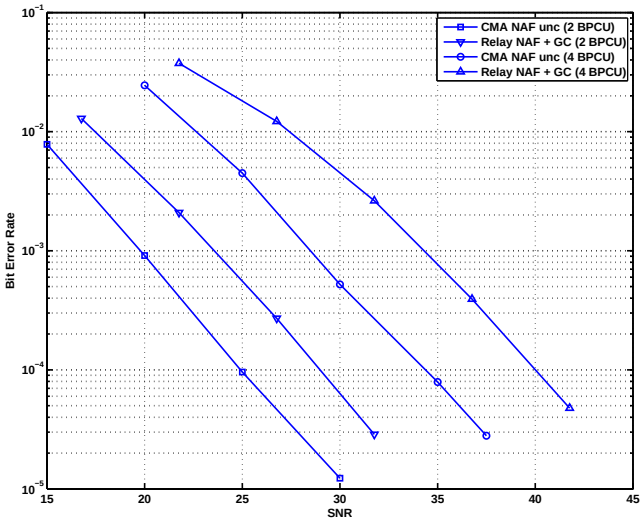

Figure 8: BER of CMA-NAF protocol with uncoded QAM versus Relay NAF protocol with Golden constellation.

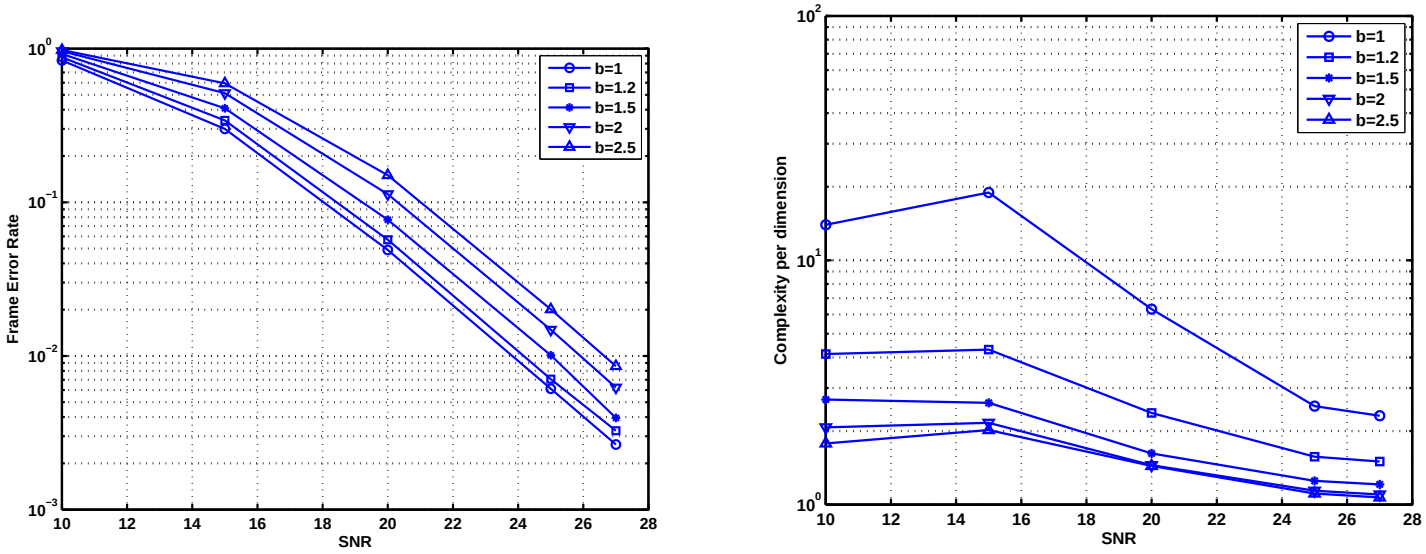

Figure 9: Performance-complexity tradeff of MMSE-DFE Fano decoder in the CMA-NAF setup (2 BPCU).

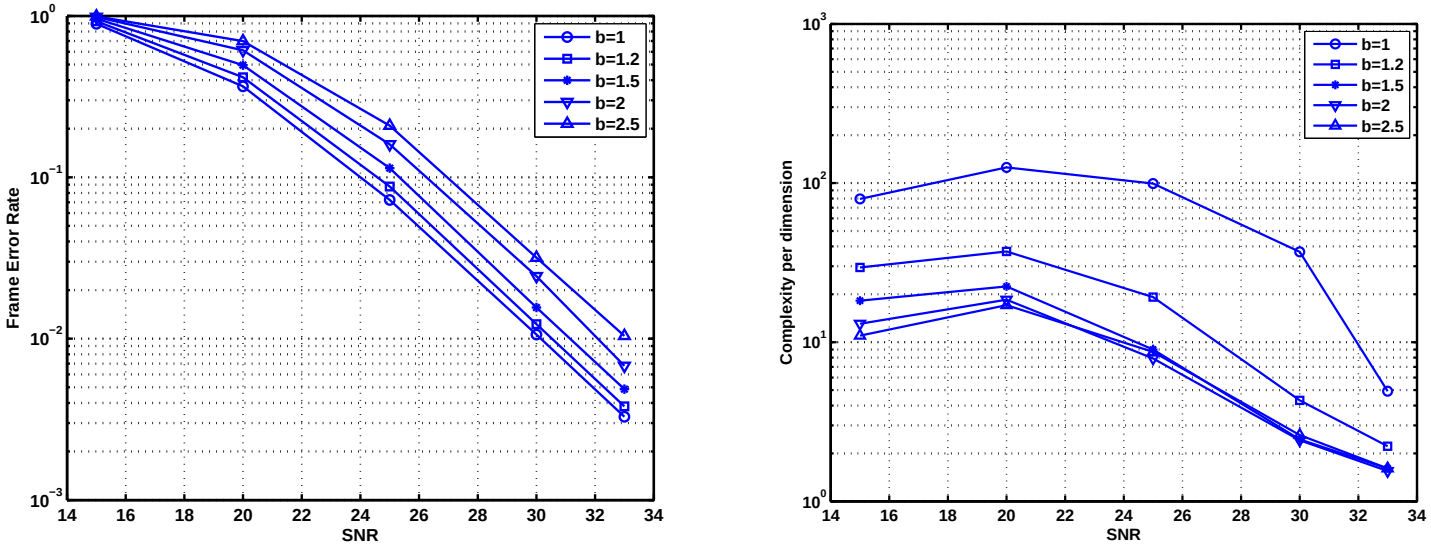

Figure 10: Performance-complexity tradeoff of MMSE-DFE Fano decoder in CMA-NAF setup (4 BPCU).

8.3 Hesham El Gamal

Hesham El Gamal received the B.S. and M.S. degrees in Electrical Engineering from Cairo University, Cairo, Egypt, in 1993 and 1996, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Maryland at College Park, MD, in 1999. From 1993 to 1996, he served as a Project Manager in the Middle East Regional Office of Alcatel Telecom. From 1996 to 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, the University of Maryland at College Park, MD. From February 1999 to December 2000, he was with the Advanced Development Group, Hughes Network Systems (HNS), Germantown, MD, as a Senior Member of the Technical Staff. In the Fall of 1999, he served as a lecturer at the University of Maryland at College Park. In January 2001 he joined the ECE Department at the Ohio State University where he is now an Associate Professor. He held visiting appointments at UCLA (Fall 2002, Winter 2003) and Institut Eurecom (Summer 2003).

He is a recipient of the HNS Annual Achievement Award (2000), the OSU College of Engineering Lumley Research Award (2003), the OSU Electrical Engineering Department FARMER Young Faculty Development Fund (2003-2008), and the National Science Foundation CAREER Award (2004). He holds 5 U.S. patents and has eight more patent applications pending. He is a Senior Member of the IEEE and currently serves as an Associate Editor for “Space-Time Coding and Spread Spectrum” for the IEEE Transactions on Communications.