Time-Invariant Rateless Codes for MIMO Channels

Maryam Modir Shanechi
Dept. EECS, MIT
Cambridge, MA
Email: shanechi@mit.edu

Uri Erez
Tel Aviv University
Ramat Aviv, Israel
Email: uri@eng.tau.ac.il

Gregory W. Wornell
Dept. EECS, MIT
Cambridge, MA
Email: gww@mit.edu

Kevin P. Boyle
Vanu, Inc.
Cambridge, MA
Email: kboyle@alum.mit.edu

Abstract— Two time-invariant rateless code constructions are developed for efficient communication over multi-input multi-output (MIMO) Gaussian channels. Both architectures employ layering, dithering, and repetition as key ingredients, and convert the MIMO channel into a scalar channel to which classical Gaussian base codes can be applied. Both constructions are convolutionally structured—one is based on faster-than-Nyquist (ftN) signaling, while the other on a diagonal layering (DL) structure. Moreover, both employ successive cancellation decoding. We show that ftN rateless codes are asymptotically capacity achieving at any signal-to-noise ratio (SNR) and induce a time-invariant scalar channel. We also show that DL codes are capacity achieving at any SNR, and induce a particular time-varying scalar channel to which standard LDPC base codes can be applied without significantly sacrificing performance.

I. INTRODUCTION

The design of practical and efficient codes for communicating over MIMO Gaussian channels when channel state information is not available a priori at the transmitter (or equivalently in a broadcast scenario where there are multiple receivers) is of significant interest in a variety of emerging wireless applications and standards. These include multi-antenna systems as well as orthogonal frequency-division multiplexing (OFDM) systems.

For such problems, a rateless approach is rather natural, whereby an encoder maps a message into an infinite-length codeword for transmission over the channel. The decoder attempts to decode from successive prefixes of its received sequence. When it succeeds, it sends a single-bit acknowledgment (ACK) to the encoder to terminate the transmission. As such, rateless codes are instances of hybrid ARQ (HARQ) protocols, the design of which for MIMO systems has received comparatively little attention to date in the literature.

Of particular interest are rateless codes that achieve rates close to capacity and require low decoding complexity. Examples of such codes are the well-known Raptor codes [1] for erasure channels, and constructions for the scalar Gaussian channel that involve the transmission of sequences of redundancy blocks developed in [2], [3].

In this paper, we develop efficient hybrid ARQ protocols, i.e., efficient rateless codes in the sense above, for the MIMO Gaussian channel. Moreover, in contrast to the approach in [3], we impose the additional constraint of time-invariant encoding, which means that message recovery can start with any redundancy block. This is accomplished by the use of an inherently more convolutionally-oriented code architecture than those in [3]–[5]. In our development, we also emphasize the finite SNR regime. As such, our approach differs from, e.g., those of [6], [7], which focus on high SNR analysis.

Our contribution consists of two time-invariant rateless code architectures. These architectures map the MIMO Gaussian channel into an equivalent scalar Gaussian channel to which a standard base code can be applied. The first construction is based on the application of faster-than-Nyquist (ftN) coding, and as such extends the scalar ftN codes in [2] to the MIMO channel. The second construction is based on a diagonal layering (DL) structure. In both cases, we combine the layering implicit in their construction with dithering and repetition techniques, as well as successive cancellation decoding, to produce efficient codes. In particular, we show that such codes are capacity-approaching when used in conjunction with suitable base codes.

II. CHANNEL MODEL AND PROBLEM FORMULATION

The channel model of interest takes the form[1]

$$\mathbf{y} = \mathbf{Hx} + \mathbf{w}, \tag{1}$$

where channel input $\mathbf{x}$ and output $\mathbf{y}$ at a particular time are N_t- and N_r-dimensional vectors, respectively, and where the associated noise $\mathbf{w}$ is $\mathcal{CN}(0, N_0\mathbf{I})$ and independent over time. The channel input is constrained to a total average power of $N_t P$.

The transmission is arranged in redundancy blocks of T symbols, which is assumed to be large but otherwise plays no role in the analysis. We use $\mathbf{x}_m$, $\mathbf{y}_m$, and $\mathbf{w}_m$ to denote the corresponding input, output, and noise for an arbitrary time in transmission of the mth redundancy block.

The channel matrix realization, $\mathbf{H}$, is known to the receiver but not to the transmitter. Furthermore, our target capacity is the white input capacity

$$C_{\mathrm{MIMO}}(\mathbf{H}) = \log \det \left(\mathbf{I} + \frac{P}{N_0} \mathbf{HH}^\dagger \right) \tag{2}$$

The encoder generates from the message a series of incremental redundancy blocks and transmits them over space (the

---

[1]To simplify the notation, we keep $\mathbf{H}$ fixed. However, our constructions are also capacity-achieving for models with time-varying $\mathbf{H}$ that are the natural generalization of the SISO channels in [5].
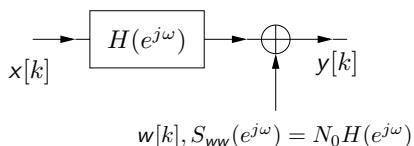
Fig. 1. Equivalent discrete-time channel model for faster-than-Nyquist communication.

Fig. 2. Equalizer structure for faster-than-Nyquist communication.

transmit antennas) and time. The receiver in turn collects as many redundancy blocks as needed to accumulate enough mutual information to decode the message. The number of blocks that need to be collected directly depends on $C_{\mathrm{MIMO}}(\mathbf{H})$. We use C^* to denote the *ceiling rate* of the code, i.e., a parameter specifying the highest rate at which the code can operate.

We require our rateless codes to be capacity-approaching in the following sense. For any channel matrix $\mathbf{H}$ that satisfies

$$C_{\mathrm{MIMO}}(\mathbf{H}) = C^*/m \qquad 1 \le m \le M, \qquad (3)$$

for some integer m, we require the message to be decoded after sending a sequence of m sets of N_{t} redundancy blocks. The parameter M is the *range* of the code, i.e., the maximum number of temporal redundancy blocks for which the code is designed.

Finally, note that when $\mathbf{H} = \mathrm{diag}(\alpha_1, \ldots, \alpha_N)$, the MIMO channel specializes to the parallel channel, in which case (2) becomes

$$C_{\|}(\mathbf{H}) = \sum_{n=1}^{N} \log \left(1 + \frac{P}{N_0} |\alpha_n|^2 \right). \qquad (4)$$

III. FASTER-THAN-NYQUIST RATELESS CODING

Faster-than-Nyquist coding is a symmetric, layered capacity-achieving communication strategy well suited for use in the construction of time-invariant rateless codes, as we now develop.

A. ftN Coding for Scalar Channels

Fig. 1 depicts the equivalent discrete-time model for ftN communication with oversignaling rate L over a scalar additive white Gaussian noise (AWGN) channel. In particular, the channel model takes the form

$$y[k] = \sum_i x[i] h[k-i] + w[k] \qquad (5)$$

where $h[k] = \mathrm{sinc}(k/L)$, and where $w[k]$ is the equivalent noise with power spectral density $S_{ww}(e^{j\omega}) = N_0 H(e^{j\omega})$. The associated power constraint on the input sequence is given by $E\left[|x[k]|^2\right] \le P/L$.

At the decoder, we exploit information-lossless successive cancellation in the form of (unbiased) minimum mean-square error decision feedback equalization (MMSE-DFE), in which the slicer performs decoding on an entire codeword. This
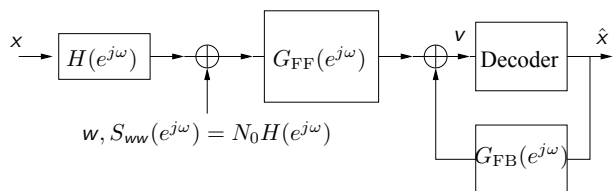
equalizer structure is depicted in Fig. 2, where [8]

$$G_{\mathrm{FF}}(z) = \frac{P/(LS_0)}{F^*(1/z^*)} \left(\frac{S_0}{S_0 - N_0} \right)$$

$$G_{\mathrm{FB}}(z) = 1 - \frac{S_0}{S_0 - N_0} \left(F(z) - \frac{N_0}{S_0} \right)$$

with $F(z)$ denoting the causal, monic and minimum phase solution to

$$S_0 F(z) F^*(1/z^*) = \frac{P}{L} H(z) + N_0 \qquad (6)$$

and S_0 the associated normalization constant.

To accommodate the requirement that the slicer operate on entire codewords, the message is divided into K independently coded submessages and the interleaved sequence [9], [10]

$$\underbrace{0, \cdots, 0}_{J}, \mathbf{x}_0(1), \cdots, \mathbf{x}_{K-1}(1), \underbrace{0, \cdots, 0}_{J}, \mathbf{x}_0(2),$$

$$\cdots, \mathbf{x}_{K-1}(2), \cdots, \underbrace{0, \cdots, 0}_{J}, \mathbf{x}_0(T), \cdots, \mathbf{x}_{K-1}(T) \quad (7)$$

is sent over the channel, where $\mathbf{x}_k(t)$ is the tth symbol in the kth subcodeword, and where the interleaver length is $J + K$ and J is a parameter determined by the effective length of the unit-sample response of the feedforward equalized channel.

With this interleaving structure and a sufficiently large J, the slicer input approaches the form

$$\mathbf{v}_k = \mathbf{x}_k + \left(\sum_{i=-\infty}^{-1} \tilde{h}[i] \mathbf{x}_{k-i} \right) + \tilde{\mathbf{w}}_k \qquad (8)$$

where

$$\tilde{h}[k] = g_{\mathrm{FF}}[k] * h[k] \qquad (9)$$

$$\tilde{\mathbf{w}}_k = g_{\mathrm{FF}}[k] * \mathbf{w}_k, \qquad (10)$$

with $\mathbf{w}_k$ denoting the noise vector corresponding to $\mathbf{x}_k$. We denote by $\mathrm{SNR_{MMSE-DFE-U}}$ the associated signal-to-interference+noise ratio (SINR).

With the use of the unbiased MMSE-DFE, it follows that the capacity of the equivalent ftN channel is independent of L, i.e.,

$$L \log \left(1 + \mathrm{SNR_{MMSE-DFE-U}} \right) = \log \left(1 + \mathrm{SNR} \right) \qquad (11)$$

where $\mathrm{SNR} = P/N_0$ is the SNR per symbol of the original channel, whence

$$\mathrm{SNR_{MMSE-DFE-U}} = \left(1 + \mathrm{SNR} \right)^{1/L} - 1. \qquad (12)$$

The attractiveness of ftN signaling is that higher oversignaling factors L leads to lower per-symbol SNRs. As we develop next, this allows us, with appropriate care, to exploit repetition redundancy as part of our rateless code design, since repetition incurs a vanishing loss in the low-SNR regime.

B. ftN-based Rateless Coding for Parallel Channels

A rateless code for the parallel channel is readily constructed by creating a collection of redundancy blocks in space and time that are dithered versions of the basic ftN-coded block, generalizing [2]. In particular, the transmitter sends the dithered codewords

$$\mathbf{x}_0 \odot \mathbf{d}_0^n(m), \quad \mathbf{x}_1 \odot \mathbf{d}_1^n(m), \quad \cdots, \quad \mathbf{x}_{K-1} \odot \mathbf{d}_{K-1}^n(m)$$

as the mth redundancy block on the nth subchannel, where $\mathbf{d}_j^n(m)$ is a symbol-by-symbol i.i.d. equiprobable ± 1 dither sequence for codeword j in the corresponding redundancy block, and where $\odot$ denotes component-wise multiplication. The dither sequences are known to both encoder and decoder.

A suitable receiver takes the form illustrated in Fig. 3 (for the case of $N = 2$ subchannels). In particular, the output of each subchannel n is processed by the appropriate feedforward equalizer filter $g_{\text{FF}}^n[k]$ determined by its channel gain α_n. Next, maximal ratio combining (MRC) is applied to the set of spatial and temporal (feedforward equalized) redundancy blocks $\tilde{\mathbf{v}}_k^n(m)$, generating

$$\tilde{\mathbf{v}}_k = \sum_{n=1}^N \lambda_n \frac{1}{m'} \sum_{m=1}^{m'} \frac{\mathbf{d}_k^n(m) \odot \tilde{\mathbf{v}}_k^n(m)}{\alpha_n} \tag{13}$$

for an appropriate choice of λ_n. Finally, the result is used in conjunction with the feedback filter

$$g_{\text{FB}}[k] = \sum_{n=1}^N \lambda_n g_{\text{FB}}^n[k], \tag{14}$$

where

$$\lambda_n = \frac{\text{SNR}_{\text{MMSE-DFE-U}}^{(n)}(L)}{\sum_{n'=1}^N \text{SNR}_{\text{MMSE-DFE-U}}^{(n')}(L)}$$

and [cf. (12)]

$$\text{SNR}_{\text{MMSE-DFE-U}}^{(n)}(L) = \left(1 + |\alpha_n|^2 \text{SNR}\right)^{1/L} - 1$$

to produce the slicer input, from which codeword decisions are produced.

With these choices, and provided that [cf. (4)] $C_\| = C^*/m'$, the slicer input takes the form

$$\mathbf{v}_k = \mathbf{x}_k + \sum_{i=-\infty}^{-1} \sum_{n=1}^N \tilde{h}_n[i] \lambda_n \frac{1}{m'} \sum_{m=1}^{m'} \tilde{\mathbf{x}}_{k,i}^n(m) + \tilde{\mathbf{w}}_k \tag{15}$$

where

$$\tilde{\mathbf{x}}_{k,i}^n(m) = \mathbf{d}_k^n(m) \odot \mathbf{d}_{k-i}^n(m) \odot \mathbf{x}_{k-i} \tag{16}$$

and

$$\tilde{\mathbf{w}}_k = \sum_{n=1}^N \lambda_n \frac{1}{m'} \sum_{m=1}^{m'} \frac{\mathbf{d}_k^n(m) \odot \tilde{\mathbf{w}}_k^n(m)}{\alpha_n}. \tag{17}$$
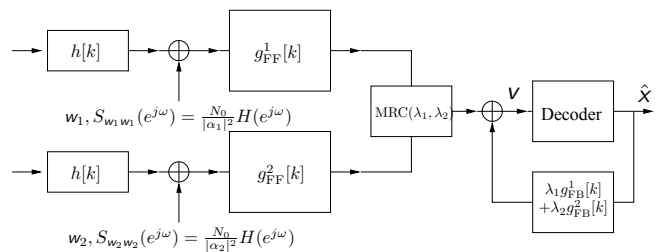

Fig. 3. Faster-than-Nyquist system model for the parallel channel with $N = 2$ subchannels.

Since the sequences $\tilde{\mathbf{x}}_{k,i}^n(m)$ are uncorrelated in m and n, the SINR at the slicer input for the kth subcodeword is

$$\text{SNR}_{\text{MRC}}(m') = m' \sum_{n=1}^N \text{SNR}_{\text{MMSE-DFE-U}}^{(n)}(L),$$

which can be bounded by

$$m' \sum_{n=1}^N \text{SNR}_{\text{MMSE-DFE-U}}^{(n)}(L)$$
$$\geq m' \sum_{n=1}^N \ln 2 \, \log(1 + \text{SNR}_{\text{MMSE-DFE-U}}^{(n)}(L))$$
$$= \frac{m'}{L} \sum_{n=1}^N \ln 2 \, \log(1 + |\alpha_n|^2 \text{SNR}) = \ln 2 \frac{C^*}{L},$$

where to obtain the last equality we have used $C_\| = C^*/m'$. Hence the achievable mutual information per combined coded symbol is bounded by

$$R_{\text{MRC}}(m') = \log\left(1 + m' \sum_{n=1}^N \text{SNR}_{\text{MMSE-DFE-U}}^{(n)}(L)\right)$$
$$\geq \log\left(1 + \ln 2 \frac{C^*}{L}\right) \tag{18}$$

Thus, given perfect base codes with rate (18), the message can be decoded, and the associated efficiency of the scheme is at least

$$\eta(L) \geq \frac{L \log\left(1 + \ln 2 \frac{C^*}{L}\right)}{C^*}$$

Evidently, the efficiency can be made arbitrarily close to unity by making the oversignaling rate L sufficiently large.

Of course, the preceding analysis neglects the loss due to the zero-padding during interleaving. However, the efficiency loss due to zero-padding is $J/(J + K)$, which can be made negligible by choosing $K \gg J$.

C. ftN-based Rateless Coding for MIMO Channels

To obtain a ftN-based rateless code for general MIMO channels, it suffices to exploit the D-BLAST architecture and associated MMSE-SIC (successive interference cancellation) receiver [11] to transform a MIMO channel into a parallel channel. This results in a concatenated code where D-BLAST
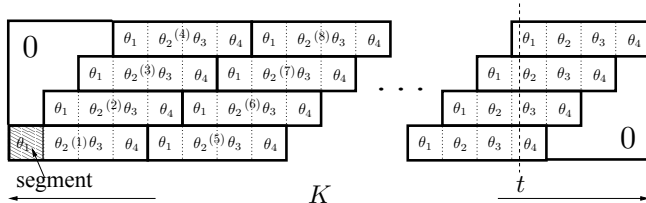
Fig. 4. Encoding structure for diagonally layered communication. In this example, $L = 4$. Each "brick" represents a subcodeword, and its associated parenthesized index indicates the order in which decoding proceeds. The phases θ_l applied within the subcodewords are parameters exploited to develop a rateless code from DL encoding.

is the inner code, and our ftN-based scheme for the parallel channel is the outer code.

In the D-BLAST architecture, the ftN scheme submessages are each further divided into U further submessages, which are independently coded. These messages are then transmitted over the N_t channel inputs according to the following encoding of length $N_t + U - 1$

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{x}_1^{N_t} \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ \mathbf{x}_1^{N_t-1} \\ \mathbf{x}_2^{N_t} \end{bmatrix}, \cdots, \begin{bmatrix} \mathbf{x}_1^1 \\ \vdots \\ \mathbf{x}_{N_t-1}^{N_t-1} \\ \mathbf{x}_{N_t}^{N_t} \end{bmatrix}, \cdots, \begin{bmatrix} \mathbf{x}_U^1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $\mathbf{x}_u^n$ denotes the ftN encoding of the uth set of K submessages for the nth channel input.

Since the D-BLAST architecture is capacity achieving when combined with MMSE-SIC decoding, the overall ftN-based rateless code inherits the same efficiency characteristics as that for the parallel channel. Such analysis neglects the further efficiency loss due to zero-padding in the D-BLAST architecture, which can be made sufficiently small by choosing $U \gg N_t$.

As is apparent from the preceding development, to achieve high efficiencies with the ftN-based rateless code architecture can require large delays. Since many applications are more sensitive to delay characteristics, we now turn to a different but related rateless construction for MIMO channels.

IV. DIAGONALLY LAYERED RATELESS CODING

A time-invariant rateless code for MIMO channels can be constructed via a diagonal layering approach, as we now develop.

A. DL Coding for Scalar Channels

With DL coding, the message is divided into LK submessages, each of which is independently coded into a corresponding subcodeword. These subcodewords are then superimposed in a staggered (diagonal) manner, where the offset with respect to the previous subcodeword is fixed and determined by the choice of L. The resulting equivalent layered structure of the encoding is as depicted in Fig. 4, where we see that there are K subcodewords per layer, L layers, and L segments per subcodeword which see interference from different sets of subcodewords.

Such an encoding can be decoded using successive cancellation. In particular, as shown in Fig. 4, we decode starting from the lower left in a cyclic manner, subtracting the effect of subcodewords as they are decoded, and treating undecoded layers as noise. With this approach, the effective channel for all subcodewords are identical. Thus, using perfect base codes, the resulting DL encoding is capacity achieving when the same rate and power is used for each subcodeword (provided K is chosen large enough that the efficiency loss due to zero-padding in the construction is negligible). This symmetry makes the DL architecture well-suited for the development of our rateless codes.

B. DL-based Rateless Coding for MIMO Channels

A rateless code for the MIMO channel is constructed by creating a collection of redundancy blocks in space and time that are distinct linear transformed versions of the basic DL-coded block. In particular, as Fig. 4 reflects, prior to superposition in the DL encoding, the lth segment in each subcodeword is multiplied by $e^{j\theta_l}$, where θ_l is a phase parameter. These phases must be chosen for each of the M sets of N_t incremental redundancy blocks in the rateless code.

The redundancy block at an arbitrary time t (see Fig. 4) can be expressed in the form

$$\begin{bmatrix} \mathbf{x}_1^T & \cdots & \mathbf{x}_M^T \end{bmatrix}^T = \sqrt{\frac{P}{L}} \mathbf{G} \begin{bmatrix} c_1 & \cdots & c_L \end{bmatrix}^T \quad (19)$$

where c_l is an arbitrary element of a (normalized) subcodeword in layer l, $\mathbf{x}_m$ is the N_t-dimensional channel input during incremental redundancy interval m, and $\mathbf{G}$ is a $N_t M \times L$ matrix whose (i,l)th component $e^{j\theta_{i,l}}$ describes the phase for the segment at time t in the lth layer of the redundancy block indexed by i. Note that as Fig. 4 reflects, we can express the redundancy block in any other segment by keeping $\mathbf{G}$ unchanged but using the appropriate cyclic shift of $c_1, \ldots, c_L$.

The receiver collects as many spatial and/or temporal redundancy blocks as it needs to accumulate enough mutual information to recover the message. The receiver combines and decodes these redundancy blocks using the information-lossless MMSE-SIC front end, creating an equivalent scalar channel.

The choice of $\mathbf{G}$ determines the performance of the rateless code. We assert the following proposition, whose straightforward proof we omit due to space constraints.

Proposition 1: Provided perfect base codes of rate C^*/L are used, the DL-based rateless coding architecture is perfect, i.e., capacity achieving for any channel realization $\mathbf{H}$ such that $C_{\text{MIMO}}(\mathbf{H}) = C^*/m$ for every $m \in \{1, \ldots, M\}$ if and only if $L \geq N_t M$ and $\mathbf{G}/\sqrt{L}$ has orthonormal rows.

The constraints of Proposition 1 are straightforward to satisfy. For example, it is sufficient to take $\mathbf{G}$ to be the $N_t M \times N_t M$ discrete Fourier transform (DFT) matrix, or, when it exists, the Hadamard matrix.

C. DL-coding with AWGN Base Codes

The symmetry inherent in the DL-based encodings makes them well suited for rateless constructions. However, this
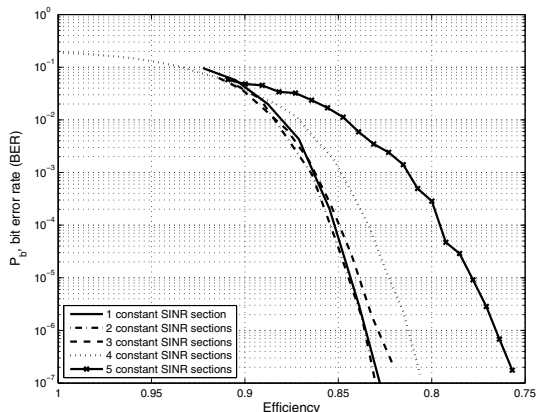
Fig. 5. Bit error rate as a function of efficiency and for different numbers L of constant SINR sections, for a fixed rate $1/4$ b/1D base code.

comes with a price. Specifically, the effective channel is time-varying—having the property that the SINR is fixed in each of the L segments comprising a subcodeword, and changes from segment to segment. The choice of base code and its decoding should, in general, take this into account. In this section, we investigate, by simulation, the performance of standard AWGN base codes in this application.

Motivated by the reported success of using LDPC codes on a variety of low SNR time-varying channels [12], [13], we investigate the performance of the rate 1/4 b/1D LDPC base code of length 64,800 designed for the DVB-S2 standard [14] in our application. The worst-case MIMO channel for such an application is that yielding the most pronounced SNR variation within subcodewords. In general this worst-case channel is difficult to determine. However, if we restrict our attention to channels without interference (i.e., diagonal **H**), corresponding to the subset of parallel channels, the worst among such channels has a single non-zero link. We use this as the nominal channel in our evaluation.

Since the number of layers L determines the number of constant-SINR sections, and since the base code is of fixed rate, increasing L proportionally increases C^*. Thus, we plot error rate as a function of efficiency $\eta = R/\overline{I}$, where R is the rate of the base code and

$$\overline{I} = \frac{1}{L} \sum_{i=1}^{L} \log \left(1 + \frac{\text{SNR}}{(i-1)\text{SNR} + 1} \right) \qquad (20)$$

is the average mutual information achievable per symbol, with SNR denoting the SNR per coded symbol.

Fig. 5 shows the resulting bit-error rate (BER) curves as a function of efficiency for various numbers of fixed SNR segments with our fixed base code. We see that at a BER of 10^{-5}, the performance of the code is not significantly affected for up to $L = 4$ segments. Beyond $L = 4$, the efficiency drops, most likely due to the fact that the binary base code alphabet becomes limiting in the first segment where the SNR is high.

Using such results, we verify that the DL-based rateless codes perform significantly better on the 2×2 MIMO channel than the baseline rateless code one would get, for example, by using standard Alamouti code [15] with the redundancy blocks generated by repetition. In particular, a DL-based code consisting of 2 sets of 2 redundancy blocks, and $L = 4$ layers, used with the rate 1/4 b/1D base code above, achieves an efficiency of 82% on our nominal channel at a BER of 10^{-5}, neglecting error propagation, and corresponds to $C^* = 4 \times 1/2 \times 1/0.82 = 2.4$ b/s/Hz. By contrast, even with perfect base codes the Alamouti-with-repetition scheme can be shown to achieve an efficiency of 67%. Moreover, this scheme requires good (non-binary) base codes of rate $2.4 \times 0.67/2 = 0.8$ b/1D, which are more challenging to design. Thus we conclude that as a practical MIMO rateless code, the DL-based construction shows considerable promise.

REFERENCES

[1] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2551–2567, June 2006.

[2] U. Erez, G. W. Wornell, and M. D. Trott, "Faster-than-Nyquist coding," in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, Illinois, 2004.

[3] U. Erez, M. Trott, and G. Wornell, "Rateless coding for Gaussian channels," *IEEE Trans. Inform. Theory*, 2007, submitted for publication. Also available on arXiv; see http://arxiv.org/abs/0708.2575v1.

[4] M. M. Shanechi, U. Erez, and G. W. Wornell, "On universal coding for parallel Gaussian channels," in *Proc. Int. Zurich Seminar Commun.*, Zurich, Switzerland, Mar 2008.

[5] U. Erez, M. D. Trott, and G. W. Wornell, "An efficient ARQ scheme with SNR feedback," in *Proc. Int. Zurich Seminar Commun.*, Zurich, Switzerland, Mar 2008.

[6] L. Zhao and S.-Y. Chung, "Perfect rate-compatible codes for the Gaussian channel," in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, Illinois, Sep. 2007.

[7] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3233–3258, July 2006.

[8] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and J. G. D. Forney, "MMSE decision-feedback equalizers and coding—part I: Equalization results," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2582–2594, Oct 1995.

[9] T. Guess and M. K. Varanasi, "An information-theoretic framework for deriving canonical decision-feedback receivers in Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 173–187, Jan 2005.

[10] J. G. David Forney, "Shannon meets Wiener II: On MMSE estimation in successive decoding schemes," in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, Illinois, Oct 2004.

[11] G. J. Foschini, "Layered space-time architectures for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[12] C. R. Jones, J. V. T. Tian, and R. D. Wesel, "The universal operation of LDPC codes over scalar fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 122–132, Jan 2007.

[13] C. Jones, T. Tian, A. Matache, R. D. Wesel, and J. Villasenor, "Robustness of LDPC codes on periodic fading channels," in *Proc. GLOBECOM*, Taipei, Taiwan, Nov 2002, pp. 1282–1284.

[14] "European Telecommunications Standards Institute. ETSI EN 302 307 v1.1.2 (2006-06): Digital Video Broadcasting (DVB); Second generation framing structure, channel coding, and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications."

[15] S. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451–1458, Oct 1998.