Optimal SNR Exponent for Discrete-Input MIMO ARQ Block-Fading Channels

Allen Chuang*, Albert Guillén i Fàbregas†, Lars K. Rasmussen‡, Iain B. Collings§
*School of Electrical and Information Engineering, University of Sydney, Sydney NSW 2006, Australia
†Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom
‡Institute for Telecommunications Research, University of South Australia, Mawson Lakes SA 5095, Australia
§ICT Centre, CSIRO, Australia
Email: *a.chuang@ieee.org, †guillen@ieee.org, ‡lars.rasmussen@unisa.edu.au, §iain.collings@csiro.au

***Abstract*— In this paper, we consider an automatic-repeat-request (ARQ) retransmission protocol signaling over a block-fading multiple-input, multiple-output (MIMO) channel. In particular, we consider fixed rate codes constructed over discrete complex signal constellations. We show that the optimal signal-to-noise ratio (SNR) exponent is given by a modified Singleton bound, relating all the system parameters. To demonstrate the practical significance of the theoretical analysis, we present numerical results showing that practical Singleton-bound-achieving maximum distance separable codes achieve the optimal SNR exponent.**

I. Introduction

The block-fading channel model was introduced in [1], allowing for transmission extending over channels with multiple block-fading periods. Within a block-fading period, the fading channel coefficients remain constant, while between periods the channel coefficients change randomly according to a fading distribution. The block-fading channel model is a reasonable model for orthogonal frequency division multiplexing (OFDM) transmission over frequency-selective wireless channels. Despite its simplicity, the model captures important aspects of OFDM modulation over frequency-selective fading channels and it is useful for developing coding design criteria.

The seminal work of Teletar [2], and Foschini and Gans [3], has inspired a flurry of research in multiple-input, multiple-output (MIMO) channels. The fundamental tradeoff between diversity gain and multiplexing gain[1] for quasi-static MIMO channels is described in [4], assuming Gaussian distributed inputs. The fundamental tradeoff developed in [4] has become a benchmark for the performance evaluation of space-time coding schemes, and the corresponding framework has become a preferred approach for characterizing classes of MIMO channels [5, 6]. For fixed rate codes constructed over discrete signal constellations, the work in [7] reports the fundamental rate-diversity tradeoff over quasi-static MIMO channels.

**This work was supported by the Australian Research Council under ARC grants DP0558861 and RN0459498.

[1]The diversity gain (or signal-to-noise ratio (SNR) exponent) is defined as $d \triangleq -\lim_{\mathrm{SNR}\to\infty} \frac{\log P_e(\mathrm{SNR})}{\log \mathrm{SNR}}$, where $P_e(\mathrm{SNR})$ denotes the probability that the transmitted message is decoded incorrectly. The multiplexing gain is defined as $r_m \triangleq \lim_{\mathrm{SNR}\to\infty} \frac{R(\mathrm{SNR})}{\log \mathrm{SNR}}$, where $R(\mathrm{SNR})$ is the code rate. The multiplexing gain essentially quantifies how close the code rate is to the capacity of a single-input single-output link at high SNR [4].

The fundamental tradeoff for MIMO automatic-repeat-request (ARQ) block-fading channel is described in [5] under the Gaussian input assumption. In this work, the authors established the fundamental tradeoff between diversity gain, multiplexing gain and delay (i.e. maximum number of retransmissions) over MIMO ARQ channels. This result asserts ARQ delay as a potential source for diversity, even when the channel remains constant over the transmission of a given message.

In this paper, we consider an ARQ system signaling over a block-fading MIMO channel with L maximum allowable ARQ rounds and B fading blocks per ARQ round. We constrain the transmitter to fixed rate codes constructed over complex signal constellations. In particular, we examine the general case of average input-power-constrained constellations as well as the practically important case of finite discrete constellations. The receiver is able to generate a finite number of one-bit repeat-requests, subject to a latency constraint, whenever an error is detected in the decoded message. A maximum of L transmissions pertaining to each information message is allowed. The main focus of our work is to characterize the tradeoff between throughput, diversity gain and delay. Similar to [5], we demonstrate that while the optimal SNR exponent of the system is an increasing function of the maximum number of allowed ARQ rounds L, the throughput of the system is independent of L for sufficiently high SNR, and is determined by the rate of the first ARQ round. We therefore denote our main result as the *optimal throughput-diversity-delay tradeoff*. This result provides strong incentive to use ARQ as a way to increase reliability without suffering code rate penalties.

The following notation is used in the paper. Sets are denoted by calligraphic fonts. The exponential equality $f(z) \doteq z^d$ indicates that $\lim_{z\to\infty} \frac{\log f(z)}{\log z} = d$. The exponential inequality $\dot{\leq}, \dot{\geq}$ are similarly defined. $\|\cdot\|_F$ is the Frobenius norm and vector/matrix transpose is denoted by $'$ (e.g. $\mathbf{v}'$). $\lceil x \rceil$ ($\lfloor x \rfloor$) denotes the smallest (largest) integer greater (smaller) than x.

II. System Model

Consider a block-fading MIMO ARQ system with N_t transmit antennas and N_r receive antennas. We investigate the use of a simple stop-and-wait ARQ protocol where the maximum number of ARQ rounds is denoted by L. Each ARQ round is subject to B independent block-fading periods, each of length T (coherence time/bandwidth) in channel uses. Hence each ARQ round spans BT channel uses.

The received signal at the bth block and ℓth ARQ round is

$$\mathbf{Y}_{\ell,b} = \sqrt{\frac{\rho}{N_t}}\mathbf{H}_{\ell,b}\mathbf{X}_{\ell,b} + \mathbf{W}_{\ell,b}, \quad (1)$$

where $\mathbf{X}_{\ell,b} \in \mathbb{C}^{N_t\times T}, \mathbf{Y}_{\ell,b}, \mathbf{W}_{\ell,b} \in \mathbb{C}^{N_r\times T}$ and $\mathbf{H}_{\ell,b} \in \mathbb{C}^{N_r\times N_t}$ denote the transmitted signal matrix, received signal matrix, the noise matrix and the channel fading gain matrix, respectively, while ρ denotes the average SNR per receive antenna. Both the elements of the channel fading gain matrix $\mathbf{H}_{\ell,b}$ and the elements of the noise matrix $\mathbf{W}_{\ell,b}$ are assumed i.i.d. zero mean circularly symmetric complex Gaussian with variance $\sigma^2 = 0.5$ per dimension. The channel coefficients are assumed to be perfectly known to the receiver. In addition, we consider two types of fading dynamics and obtain the *long-term static* model of [5] by letting $\mathbf{H}_{\ell,b} = \mathbf{H}_{\ell',b}$ for all $\ell \neq \ell'$ in (1). The *short-term static* model of [5] is obtained when the matrices $\mathbf{H}_{\ell,b}$ are i.i.d. for each block and ARQ round.

The receiver attempts to decode following the reception of an ARQ round. If the received codeword can be decoded, the receiver sends back a one-bit acknowledgement signal to the transmitter via a zero-delay and error-free feedback link. The transmission of the current codeword ends immediately following the acknowledgment signal and the transmission of the next message in the queue starts. If an error is detected in the received codeword before the Lth ARQ round, then the receiver requests another ARQ round by sending back a one-bit negative acknowledgment along the perfect feedback path. However, a decision must be made at the end of the Lth ARQ round regardless of whether errors are detected.

In general, the optimal ARQ decoder makes use of all available coded blocks and corresponding channel state information up to the current ARQ round in the decoding process. This leads to the concept of information accumulation, where individual ARQ rounds are combined, along with any other side information. We hence introduce the ARQ channel model *up to the ℓth ARQ round*, completely analagous to (1), but allowing for a more concise notation. In particular, we have

$$\widetilde{\mathbf{Y}}_\ell = \sqrt{\frac{\rho}{N_t}}\widetilde{\mathbf{H}}_\ell\widetilde{\mathbf{X}}_\ell + \widetilde{\mathbf{W}}_\ell, \quad (2)$$

where

$$\begin{aligned}
\widetilde{\mathbf{Y}}_\ell &= \left[\mathbf{Y}'_{1,1}, \dots, \mathbf{Y}'_{1,B}, \dots, \mathbf{Y}'_{\ell,1}, \dots, \mathbf{Y}'_{\ell,B}\right]',\\
\widetilde{\mathbf{X}}_\ell &= \left[\mathbf{X}'_{1,1}, \dots, \mathbf{X}'_{1,B}, \dots, \mathbf{X}'_{\ell,1}, \dots, \mathbf{X}'_{\ell,B}\right]',\\
\widetilde{\mathbf{W}}_\ell &= \left[\mathbf{W}'_{1,1}, \dots, \mathbf{W}'_{1,B}, \dots, \mathbf{W}'_{\ell,1}, \dots, \mathbf{W}'_{\ell,B}\right]',\\
\widetilde{\mathbf{H}}_\ell &= \text{diag}\left(\mathbf{H}_{1,1}, \dots, \mathbf{H}_{1,B}, \dots, \mathbf{H}_{\ell,1}, \dots, \mathbf{H}_{\ell,B}\right).
\end{aligned}$$

That is, $\widetilde{\mathbf{Y}}_\ell \in \mathbb{C}^{\ell BN_r\times T}, \widetilde{\mathbf{X}}_\ell \in \mathbb{C}^{\ell BN_t\times T}$ and $\widetilde{\mathbf{W}}_\ell \in \mathbb{C}^{\ell BN_r\times T}$ are simply collections of the received, code and noise matrices, respectively, available at the end of the ℓth ARQ round, concatenated into block column matrices. The new channel matrix $\widetilde{\mathbf{H}}_\ell \in \mathbb{C}^{\ell BN_r\times \ell BN_t}$ is a block diagonal matrix with the diagonal blocks composed of the respective channel state during each block-fading period up to ARQ round ℓ.

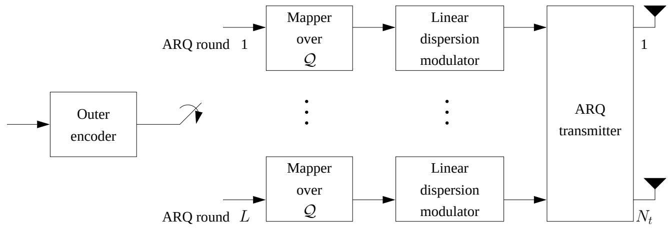

Fig. 1. Block diagram of the concatenated MIMO ARQ architecture.

A. Encoding

The information message to be transmitted is passed through a space-time coded modulation encoder with codebook $\mathcal{C} \subset \mathbb{C}^{LBN_t\times T}$ and code rate R_0, where $R_0 \triangleq \frac{R_1}{L}$ and $R_1 \triangleq \frac{1}{BT}\log_2|\mathcal{C}|$ is the code rate of the first ARQ round. Therefore, $|\mathcal{C}| = 2^{R_0LBT}$ and the message index $m \in \mathcal{M}$, where $\mathcal{M} \triangleq \{1, 2, \dots, 2^{R_0LBT}\}$ is the set of possible information message indices. We denote the codeword corresponding to information message m by $\mathbf{X}(m) \in \mathbb{C}^{LBN_t\times T}$. The rate R_0 codeword can be partitioned into a sequence of LB space-time coded matrices, denoted $\mathbf{X}_{\ell,b} \in \mathbb{C}^{N_t\times T}$. We consider a *short term* average power constraint where the transmitted codewords are normalized in energy such that $\forall\, \mathbf{X} \in \mathcal{C}, \frac{1}{LBT}\mathbb{E}[\|\mathbf{X}\|_F^2] = N_t$.

Let $\mathcal{C}$, illustrated in Figure 1, be obtained as the concatenation of a *classical* coded modulation scheme $\mathcal{C}_\mathcal{Q} \subseteq \mathcal{Q}^{LBTN_t}$ constructed over a complex-plane signal set $\mathcal{Q} = \{q_1, \dots, q_{|\mathcal{Q}|}\} \subset \mathbb{C}$ with a unit rate linear dispersion space-time modulator [8]. Furthermore, let $\mathbf{c}_\mathcal{Q} \in \mathcal{C}_\mathcal{Q}$ denote a codeword of $\mathcal{C}_\mathcal{Q}$ of length $LBTN_t$ and $Q = \log_2|\mathcal{Q}|$ the number of bits conveyed in one symbol of $\mathcal{Q}$, namely, $|\mathcal{Q}| = 2^Q$. Since the linear dispersion space-time modulator has unit rate we have that $0 \leq R_1 \leq LQN_t$.

B. Decoding

We make use of the ARQ decoder proposed in [5], which behaves as a typical set decoder for the first $L-1$ ARQ round and finally performs ML decoding at the last ARQ round. Specifically, the decoding function at ARQ round ℓ, denoted $\psi_\ell(\widetilde{\mathbf{Y}}_\ell, \widetilde{\mathbf{H}}_\ell)$, outputs the message index $\hat{m} \in \mathcal{M}$ whenever the received vector can be decoded and $\psi_\ell(\widetilde{\mathbf{Y}}_\ell, \widetilde{\mathbf{H}}_\ell) = 0$ whenever errors are detected. On the last ARQ round, retransmissions are not allowed and the ARQ decoder always outputs the best estimate of the transmitted message.

III. ARQ Performance Metrics

For ease of notation, let

$$\mathcal{D}_\ell \triangleq \left\{\psi_1(\widetilde{\mathbf{Y}}_1, \widetilde{\mathbf{H}}_1) = 0, \dots, \psi_\ell(\widetilde{\mathbf{Y}}_\ell, \widetilde{\mathbf{H}}_\ell) = 0\right\}$$

denote the event of error detection up to and including ARQ round ℓ. The expected latency κ of the system is determined by the probability of error detection, and it is given by

$$\kappa = 1 + \sum_{\ell=1}^{L-1}\Pr\left(\mathcal{D}_\ell\right), \quad (3)$$

where κ is expressed in terms of number of ARQ rounds. The corresponding transmit throughput of the system in terms of the average effective code rate is simply obtained by [5]

$$\eta(R_1, L) = \frac{R_1}{1 + \sum_{\ell=1}^{L-1} \Pr(\mathcal{D}_\ell)}, \quad (4)$$

where $\eta(R_1, L)$ is expressed in bits per channel use.

IV. Information Accumulation

The instantaneous input-output mutual information of the channel (2) up to ARQ round ℓ, for the channel realization $\widetilde{\mathbf{H}}_\ell = \widetilde{\mathbf{G}}_\ell$ can be written as

$$I\big(\rho|\widetilde{\mathbf{G}}_\ell\big) \triangleq \frac{1}{T} I(\widetilde{\mathbf{X}}_\ell\,;\,\widetilde{\mathbf{Y}}_\ell \mid \widetilde{\mathbf{H}}_\ell = \widetilde{\mathbf{G}}_\ell) = \frac{1}{T}\sum_{k=1}^{\ell} I(\rho|\mathbf{G}_\ell) \quad (5)$$

where $I(\rho|\mathbf{G}_\ell)$ is the instantaneous input-output mutual information corresponding to ARQ round ℓ. Following (5) we will refer to $I\big(\rho|\widetilde{\mathbf{G}}_\ell\big)$ as the *accumulated* mutual information up to ARQ round ℓ. $I\big(\rho|\widetilde{\mathbf{G}}_\ell\big)$ measures the normalized mutual information between $\widetilde{\mathbf{Y}}_\ell$ and $\widetilde{\mathbf{X}}_\ell$, given $\widetilde{\mathbf{H}}_\ell = \widetilde{\mathbf{G}}_\ell$. Since $\widetilde{\mathbf{G}}_\ell$ is a random matrix, $I\big(\rho|\widetilde{\mathbf{G}}_\ell\big)$ is a non-negative random variable. Further, from (5) it is clear that the accumulated mutual information is an increasing function of the ARQ round index ℓ, for a given realization of $\widetilde{\mathbf{G}}_\ell$.

Following [9, Lemma 1], we get that for $|\mathcal{M}| = 2^{R_1 BT}$, there exists a codebook $\mathcal{C}$ such that the conditional probability of error $P_e(\rho|\widetilde{\mathbf{G}}_\ell) < \epsilon$ for any $\epsilon > 0$ whenever $I\big(\rho|\widetilde{\mathbf{G}}_\ell\big) \geq R_1$ for any $\ell = 1, \ldots, L$, provided that the block length ℓBT is sufficiently large. We hence define information outage as the event that occurs when the accumulated mutual information is below R_1, namely

$$\mathcal{O}_\ell \triangleq \left\{\widetilde{\mathbf{G}}_\ell \in \mathbb{C}^{\ell BTN_r \times \ell BTN_t} : I\big(\rho|\widetilde{\mathbf{G}}_\ell\big) < R_1\right\}. \quad (6)$$

For any finite B and L, the channel defined in (2) is not information stable and the channel capacity in the strict Shannon sense is zero, since the probability of the outage event is nonzero. The corresponding outage probability is defined as

$$P_{\text{out}}(\rho, \ell, R_1) = \Pr\left(I\big(\rho|\widetilde{\mathbf{G}}_\ell\big) < R_1\right). \quad (7)$$

V. Throughput-Diversity-Delay Tradeoff

We now present the main results of this paper concerning the optimal SNR exponent of ARQ systems.

Theorem 1: Consider the channel model (2) with input constellation satisfying the short term average power constraint $\frac{1}{LBT}\mathbb{E}[\|\mathbf{X}\|_F^2] \leq N_t$. The optimal SNR exponent $d^\star(R_1)$ is given by

$$d^\star(R_1) = \begin{cases} N_t N_r LB & \text{for short-term static fading} \\ N_t N_r B & \text{for long-term static fading} \end{cases} \quad (8)$$

Further, this is achieved by Gaussian random codes of rate $R_1 > 0$, provided that the block length is sufficiently long.

Proof: Theorem 1 follows immediately as a corollary of [5, Theorem 2] after taking into account the introduction of B in the system. ■

Theorem 1 states that Gaussian codes achieve maximal diversity gain for any positive rate. As we show in the following, this is not the case with discrete input signal constellations.

Theorem 2: Consider the channel model (2) satisfying the short term average power constraint $\frac{1}{LBT}\mathbb{E}[\|\mathbf{X}\|_F^2] \leq N_t$, with discrete input signal constellations of cardinality 2^{QN_t}. The optimal SNR exponent is given by

$$d_D^\star(R_1) = N_t N_r \left(1 + \left\lfloor LB\left(1 - \frac{R_1}{LQN_t}\right)\right\rfloor\right) \quad (9)$$

for short-term static fading and

$$d_D^\star(R_1) = N_t N_r \left(1 + \left\lfloor B\left(1 - \frac{R_1}{LQN_t}\right)\right\rfloor\right) \quad (10)$$

for long-term static fading over the full range of $0 \leq R_1 \leq LQN_t$, where (9) and (10) are continuous.

Proof (Sketch): We first prove the converse and show that the diversity gain $d_D^\star(R_1)$ is upper-bounded by (9) and (10). We can use Fano's inequality to show that the outage probability $P_{\text{out}}(\rho, \ell, R_1)$ lower-bounds the error probability $P_e(\rho)$ for a sufficiently large block length. Then we bound the maximum SNR exponent by considering the diversity gain of the outage probability. For large SNR, the instantaneous mutual information is either zero or QN_t bits per channel use, corresponding to when the channel is in deep fade and when the channel is not in deep fade, respectively [10]. Achievability is proved by bounding the error probability of the typical set decoder [5] for ARQ rounds $\ell = 1, \ldots, L-1$, and that of the ML decoder at round L, using the union Bhattacharyya bound on a random coded modulation scheme over $\mathcal{Q}$ concatenated with linear dispersion space-time modulation. For finite T, we obtain similar conditions to those in [10]. Finally, as $T \to \infty$, we show that the SNR exponent of random codes is given by the bounds (9) and (10) for all values of R_1 where they are continuous. Further details of the proof are in [11]. ■

Theorem 2 states that optimal diversity gain of $N_t N_r LB$ and $N_t N_r B$ for short- and long-term models, respectively, can also be achieved by discrete signal sets coupled with linear dispersion space-time modulators, spreading the symbols of $\mathcal{Q}$ over B fading blocks at each ARQ round. Under this scenario, full diversity is maintained for all rates $0 \leq R_1 \leq QN_t$ over short-term static fading channels, and all rates $0 \leq R_1 \leq LQN_t$ over long-term static fading channels. This result demonstrates the utility of ARQ: over short-term static fading channels, ARQ can be used to increase reliability, while over long-term static fading channels, ARQ can be used to increase the range of supported transmission rates.

The upper bounds (9) and (10) are also applicable to any systems using block codes over LB independent block-fading periods. The significance of the ARQ framework is that it provides a way of achieving the optimal SNR exponent attained by a block code with LB coded blocks, without always having to transmit all LB code blocks. Following [5], it is possible to show that the throughput approaches the rate of a single ARQ round asymptotically, i.e.

$$\eta(R_1, L) \doteq R_1. \quad (11)$$

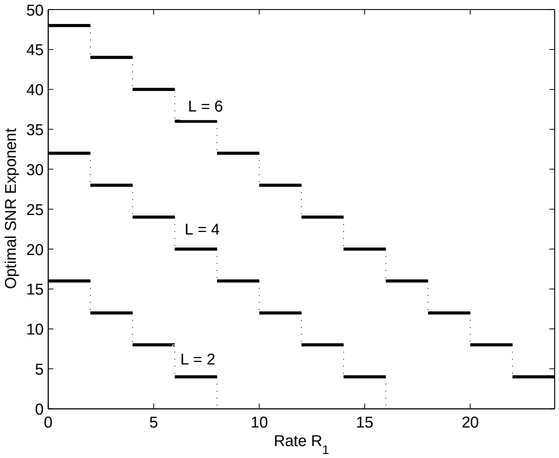

Fig. 2. Optimal diversity tradeoff curve corresponding to $B = 2, Q = 2$ for a short-term static 2×2 MIMO channel.

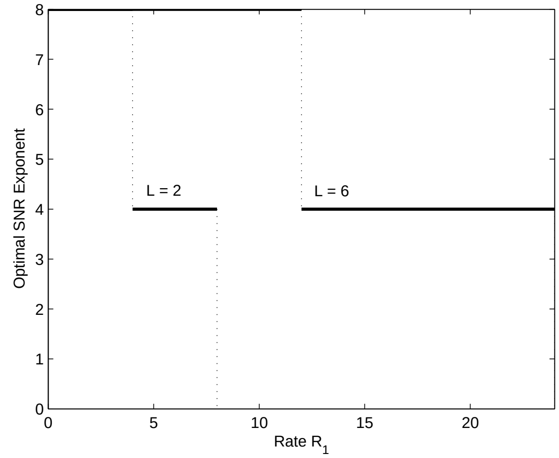

Fig. 3. Optimal diversity tradeoff curve corresponding to $B = 2, Q = 2$ for a long-term static 2×2 MIMO channel.

In other words, provided the SNR is sufficiently high, ARQ systems which send *on average* B coded blocks (i.e. one single ARQ round) can achieve the same diversity gain as that achieved by a block code system which sends LB coded blocks *every time*. This is because in the high SNR regime, most frames can be decoded correctly with high probability based only on the first transmitted code block. ARQ retransmissions are used to correct the rare errors which occur almost exclusively whenever the channel is in outage. While the throughput $\eta(R_1, L)$ is a function of L at mid to low SNR, it converges towards R_1 independent of L at sufficiently high SNR. Since the optimal diversity gain is an increasing function of L, this behavior can be exploited to increase reliability without suffering code rate losses. However, as noted in [5], this behavior is exhibited only by decoders capable of near perfect error detection (PED). Therefore, the performance of practical error detection schemes can be expected to significantly influence the throughput of ARQ systems.

Since equation (11) relates the asymptotic throughput with the coding parameter R_1, the optimal SNR exponent given by (9) and (10) gives the *optimal throughput-diversity-delay* tradeoff of MIMO ARQ block-fading channels. Examining the optimal discrete throughput-diversity-delay tradeoff (9) and (10) in more detail, we first note that $\frac{R_1}{N_t L Q} = \frac{R_0}{Q N_t} = r$ is the code rate of a binary code. i.e. $0 \leq r \leq 1$, as if the coded modulation scheme $\mathcal{C}_Q$ was obtained itself as the concatenation of a binary code of rate r and length $N_t LQBT$. Expressions (9) and (10) imply that the higher we set the target rate R_1 (equivalently, R_0), the lower the achievable diversity order. In particular, *uncoded* sequences (i.e. $R_1 = QN_tL$) achieve optimal diversity gain of $N_t N_r$, while any code with non-zero $R_1 \leq QN_tL$ will achieve optimal diversity gain less than or equal to $N_t N_r LB$ or $N_t N_r B$ in the short- and long-term static models, respectively. This is an intuitively satisfying result as LB and B are precisely the number of independent fading periods in the short- and long-term static models, respectively, each with inherent diversity $N_t N_r$.

Figures 2 and 3 illustrate the effect of the maximum number of allowed ARQ rounds L on the diversity of the system over short- and long-term static channels, respectively. It is clear from the plot that in the short-term static case the effect of L is to simply shift tradeoff curves upwards. This is also intuitively satisfying, since each additional ARQ round represents incremental redundancy, which can be considered as a form of advanced repetition coding. Each additional ARQ round contains B additional independent fading blocks and hence the diversity gain with L ARQ rounds is simply the diversity gain with $L - 1$ rounds plus B. In the case of long-term static fading, since each ARQ round uses the same channel realization, larger L implies a broader range of R_1 for which maximum diversity can be achieved.

The diversity tradeoff functions (9) and (10) can be viewed as modified versions of the Singleton bound with the diversity gain corresponding to the Hamming distance of our code $\mathcal{C}$, viewed as a code of length LB constructed over an alphabet of size 2^{QN_tT}. Therefore, Singleton-bound achieving maximum distance separable (MDS) codes are optimal for the discrete-input MIMO ARQ block-fading channel.

VI. Numerical Results

In this section, we show some examples of MDS codes for MIMO ARQ block-fading channels with $N_t = N_r = 2$. In particular, the first system has a maximum number of ARQ rounds of $L = 2$, $B = 1$, and is using the terminated 4-state $[5, 7]_8$ convolutional code, while the second system has a maximum number of ARQ rounds of $L = 4$, $B = 1$, and is using the terminated 4-state $[5, 5, 7, 7]_8$ convolutional code. The outer convolutional codes are divided into blocks, interleaved and modulated using the complex 2×2 threaded algebraic space-time (TAST) modulator proposed in [12]. The two systems are investigated for their performance over the short-term static fading channel.

In this example, the channel coherence time is $T = 32$ channel uses and the mapper over $\mathcal{Q}$ is set to 4QAM. In this case, ML decoding becomes impractical and we therefore

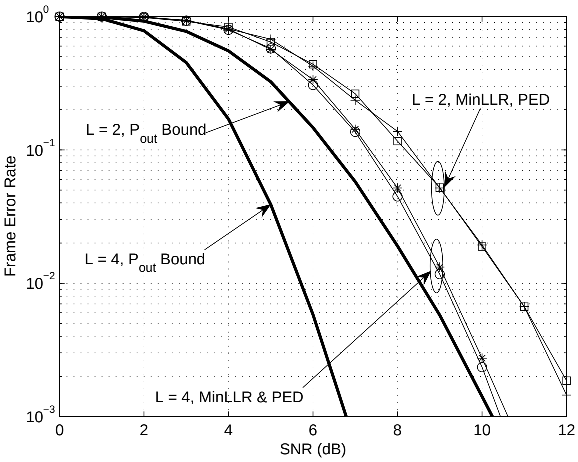

Fig. 4. FER with MDS convolutional code over a short-term static 2×2 MIMO channel corresponding to $B = 1, Q = 2$ and $T = 32$.

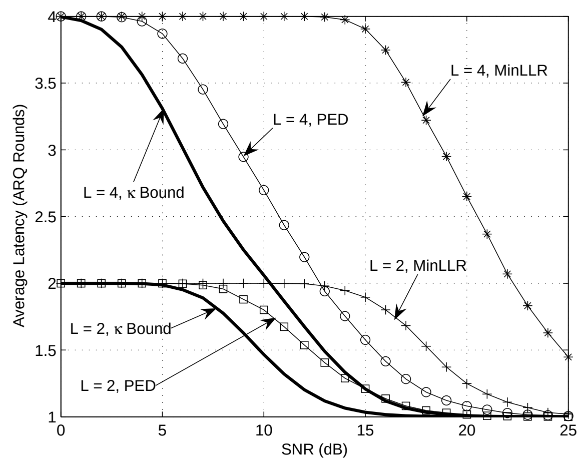

Fig. 5. Average latency for MDS convolutional code over a short-term static 2×2 MIMO channel corresponding to $B = 1, Q = 2$ and $T = 32$.

resort to a sub-optimal iterative error detection and decoding scheme. The iterative scheme is based on the max-log APP detector proposed in [13], recursively exchanging code symbol extrinsics with an outer APP decoder, thus generating estimates of the information sequence. Note that the max-log APP detector used here considers all possible input vectors, but it generates an approximation of the code symbol extrinsic using two maximization operations. At each ARQ round, we run the accumulated received signal through six iterations of the detection and decoding algorithm before examining the decoder output. Errors are detected in this system by examining the soft output of the decoder at each ARQ round, namely, the minimum bit-reliability criterion (MinLLR) proposed in [14].

Figure 4 compares the error rate performance of the $L = 2$ system and $L = 4$ systems under the short-term fading dynamics. For each system, we plot three curves, corresponding to the lower outage probability bound, the PED performance, as well as the MinLLR performance. We notice that additional retransmissions lead to an appreciable decrease in error rates, and, equally important, the MinLLR criterion performs virtually as good as perfect error detection.

Figure 5 compares the average latency (measured in ARQ rounds) of the two ARQ systems under the short-term fading scenario. Again, we plot three curves per system, corresponding to the lower bound of expected latency, as well as the PED and MinLLR performances. In this case, we observe that the cost of using the MinLLR criterion is mainly an increase in latency, caused by requesting superfluous retransmissions.

VII. Conclusion

In this paper, we derived expressions for the optimal ARQ SNR reliability function over the block-fading channel. The discrete reliability functions (9) and (10) characterize the tradeoff between diversity gain, throughput, signal set and delay. We showed that ARQ transmissions can significantly increase the level of diversity in the system. Further, the additional diversity gain due to ARQ comes with no throughput or delay penalty at high SNR. We recognized the optimal SNR reliability function as the Singleton bound in a modified form, which lead us to conclude the optimality of MDS codes. Finally, we showed via simulation that practical MDS codes can achieve the optimal SNR reliability function with low-complexity decoders on the ARQ block-fading channel.

References

[1] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[4] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[5] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: Diversity-multiplexing-delay tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3601–3621, Aug. 2006.

[6] T. T. Kim and M. Skoglund, "On the diversity-multiplexing tradeoff in multiantenna channels with resolution-constrained feedback," *Submitted to IEEE Trans. Inf. Theory*, 2006.

[7] R. Liu and P. Spasojevic, "On the rate-diversity function for MIMO channels with a finite input alphabet," in *Proc. Allerton Conf. Commun., Control and Computing*, Monticello, IL, Sep. 2005.

[8] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[9] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.

[10] A. Guillén i Fàbregas and G. Caire, "Coded modulation in the block-fading channel: Coding theorems and code construction," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 91–114, Jan. 2006.

[11] A. Chuang, A. Guillén i Fàbregas, L. K. Rasmussen, and I. B. Collings, "Optimal throughput-diversity-delay tradeoff in MIMO ARQ block-fading channels," *Submitted to IEEE Trans. Inf. Theory*, 2007. Available from `http://arxiv.org/abs/cs.IT/0701126`.

[12] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1097–1119, May 2003.

[13] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[14] A. Matache, S. Dolinar, and F. Pollara, "Stopping rules for turbo decoders," TMO Progress Rep. 42-142, JPL, Aug. 2000.