MIMO ARQ with Multi-bit Feedback: Outage Analysis

Khoa D. Nguyen, Lars K. Rasmussen, Albert Guillén i Fàbregas and Nick Letzepis

Abstract

We study the asymptotic outage performance of incremental redundancy automatic-repeat-request (INR-ARQ) transmission over multiple-input multiple-output (MIMO) block-fading channels with discrete input constellations. We first show that transmission with random codes using a discrete signal constellation across all transmit antennas achieves the optimal outage diversity given by the Singleton bound. We then analyze the optimal SNR-exponent and outage diversity of INR-ARQ transmission over the MIMO block-fading channel. We show that a significant gain in outage diversity is obtained by providing more than one bit feedback at each ARQ round. Thus, the outage performance of INR-ARQ transmission can be remarkably improved with minimal additional overhead. A practical feedback-and-power-adaptation rule is proposed for MIMO INR-ARQ, demonstrating the benefits provided by multi-bit feedback. Although the rule is sub-optimal in terms of outage performance, it achieves the optimal outage diversity.

Khoa D. Nguyen and Nick Letzepis are with the Institute for Telecommunications Research, University of South Australia, Australia. email: dangkhoa.nguyen@postgrads.unisa.edu.au; nick.letzepis@unisa.edu.au.

Lars K. Rasmussen was with the Institute for Telecommunications Research, University of South Australia, Australia. He is now with the Communication Theory Laboratory, Royal Institute of Technology, and the ACCESS Linneaus Centre, Stockholm, Sweden. email: lars.rasmussen@ee.kth.se.

Albert Guillén i Fàbregas is with the Department of Engineering, University of Cambridge, Cambridge, UK. email: albert.guillen@eng.cam.ac.uk.

This work has been presented in part at the Australian Communication Theory Workshop, Sydney, Australia, 4–7 Feb. 2009 and the 2009 IEEE International Symposium on Information Theory, Seoul, South Korea, 28 June–3 July 2009.

This work was supported by the Australian Research Council under ARC grants RN0459498, DP0558861 and DP0881160, and by the Swedish Research Council under VR grant 621-2009-4666.

I. INTRODUCTION

In this paper, we take an information-theoretic approach to analyzing and designing multiple-input multiple-output (MIMO) transmission strategies for incremental redundancy (INR) automatic-repeat-request (ARQ) schemes over a block-fading channel. In particular, we propose the use of multi-bit feedback for power adaptation and study the outage diversity of the resulting protocol over the MIMO block-fading channel, which characterizes the slope of the outage probability curve at high signal-to-noise ratio (SNR) in log-log scale.

A. Prior Art

The block-fading channel [1, 2] is a useful mathematical model for many practical wireless communication scenarios. The channel consists of a finite number of consecutive or parallel transmission blocks, where each block is affected by an independent fading coefficient. The model approximates well the characteristics of delay-limited transmission over slowly varying channels, such as Orthogonal Frequency Division Multiplexing (OFDM) transmission over slowly-fading frequency-selective multipath channel, as well as narrowband transmission with frequency-hopping as encountered in the Global System for Mobile Communications (GSM) and the Enhanced Data rate for GSM evolution (EDGE) standards.

Due to the finite number of fading blocks, the information rate supported by the channel depends on the instantaneous channel realization, and therefore is a random variable. When the instantaneous mutual information is less than the transmission rate, transmission is in outage [2]. In this case, it follows from the strong converse theorem (see, e.g., [3–5]) that messages are decoded in error with probability one [6, 7]. Furthermore, it is shown in [4, 8] that the use of sufficiently long random codes achieves an average error rate equal to the outage probability. Therefore, the outage probability is a fundamental limit on the performance of block-fading channels.

MIMO transmission has revolutionized modern wireless communications, and is now a key technology used in most current standards, e.g. WiFi (IEEE 802.11) and WiMax (IEEE 802.16) [9, 10]. Moreover, due to the randomness of the communication rate supported by the channel, it is essential to use adaptive techniques to enable high-rate reliable communication, where the transmission rate and/or power is adjusted to the channel realization. The use of adaptive techniques depends strongly on the availability of channel state information (CSI) at the transmitter and the receiver. In most communication systems, CSI can be estimated at the receiver, while CSI is usually not directly available at the transmitter. The use of ARQ transmission techniques is therefore a powerful approach for providing transmitter CSI, which in turn can be used to significantly improve the performance over block-fading channels [11].

The optimal diversity-multiplexing tradeoff for a MIMO channel with optimal (Gaussian) input constellation has been characterized in [12]. For systems with discrete input constellations, the rank criterion for the optimal outage diversity was derived in [13] from a worst-case analysis of the pair-wise error probability (PEP). References [14, 15] establish the Singleton bound on the optimal SNR-exponent of quasi-static MIMO channels with discrete input constellations. The Singleton bound is achievable by a wide range of input constellations via a unified code construction method proposed in [15].

In an INR-ARQ scheme, transmission starts with a high-rate codeword, and additional redundancy is requested via a feedback link when the codeword is not successfully decoded. Transmission is in outage if the codeword is not decodable within the maximum delay constraint allowed by the system. Traditional INR-ARQ systems implement one-bit feedback from the receiver, indicating whether additional redundancy is required. However, due to the accumulative nature of INR-ARQ schemes, performance improvements are possible when additional information regarding the status of the current transmission is provided through the feedback link. Several multi-bit feedback INR-ARQ schemes have been proposed in the literature. In particular, reference [16] shows that the throughput performance of ARQ systems can be improved by multi-bit feedback prior to each transmission round. The proposed system is equivalent to a conventional ARQ system with quantized CSI at the transmitter (CSIT). For systems with no CSIT, references [17, 18] propose transmission using convolutional codes, while reference [19] proposes using a multi-layer broadcasting strategy for multi-bit feedback ARQ. Both approaches show that multi-bit feedback can significantly improve the throughput performance of ARQ transmission. There is, however, no unified approach for designing multi-bit feedback INR-ARQ transmission schemes.

An important performance measure for INR-ARQ transmission in the MIMO block-fading channel is the rate-diversity-delay tradeoff. This tradeoff has only been studied for INR-ARQ systems with one-bit ACK/NACK feedback in [20–22]. In particular, reference [20] characterizes the rate-diversity-delay tradeoffs of Gaussian input MIMO INR-ARQ systems with both short-term and long-term average power constraints. For systems with discrete input constellations, the optimal rate-diversity tradeoff for systems with short-term power constraints was characterized in [21, 23]. For ARQ systems with discrete input constellation and long-term power constraints, an optimal power allocation rule has been derived in

[23], providing significant improvement on outage performance. However, the rate-diversity tradeoff of the corresponding system was not studied.

B. Contributions

As a first contribution we consider fixed-rate transmission over the MIMO block-fading channel. We show that the outage diversity is given by the Singleton bound, and that it is achievable with random codes constructed over arbitrary discrete input constellations. This rigorously proves that the Singleton bound is the optimal SNR-exponent of MIMO transmission with discrete input constellations. The result will also prove instrumental in designing and analyzing INR-ARQ transmission over the MIMO block-fading channel.

As our main contribution we study the rate-diversity tradeoff of the MIMO ARQ system with multi-bit feedback under long-term power constraints. The analysis shows that multi-bit feedback and optimal power adaptation provide significant outage diversity gains for ARQ transmission over the block-fading channel. It is shown that a finite number of feedback bits is sufficient to achieve the maximal outage diversity. The optimal rate-diversity tradeoff for the one-bit feedback case is also presented, which characterizes the asymptotic gains provided by the optimal power allocation rule proposed in [23]. As a further contribution a practically feasible feedback-and-power-adaptive rule is proposed. Although the rule is sub-optimal in terms of outage performance, it can achieve the optimal outage diversity, thus clearly illustrating the benefits offered by multi-bit feedback.

C. Notation and Organization

The following notations are used in the paper. Boldface uppercase ($\boldsymbol{A}$) and lowercase ($\boldsymbol{a}$) variables correspondingly denote matrices and vectors; while scalar variables are denoted by lightface (a or A). Sets are denoted by calligraphic letters; while the sets of real and complex numbers are correspondingly denoted with $\mathbb{R}$ and $\mathbb{C}$. The mathematical expectation of a random variable is denoted by $\mathbb{E}[\cdot]$. Non-conjugate transpose of matrices are denoted by $(\cdot)'$. The operation $\lfloor\cdot\rfloor$ ($\lceil\cdot\rceil$) returns the maximum (minimum) integer smaller (larger) than a real number. For convenience, the physical meanings of commonly used parameters are summarized in Table I.

The remainder of the paper is organized as follows. Section II describes the MIMO block-fading channel model. Section III proposes the multi-bit feedback INR-ARQ system based on mutual information and information outage. Sections IV and V discuss system design and performance analysis. Finally, concluding remarks are given in Section VI and proofs are collected in the Appendices.

II. System Model

Consider INR-ARQ transmission over a MIMO block-fading channel with N_t transmit and N_r receive antennas. Each ARQ round is transmitted over B additive white Gaussian noise (AWGN) blocks of J channel uses each, where block b at ARQ round ℓ is affected by a flat fading channel gain matrix $\boldsymbol{H}_{\ell,b} \in \mathbb{C}^{N_r \times N_t}$. The baseband equivalent of the channel in the ℓ-th ARQ round is given by

$$\boldsymbol{Y}_\ell = \sqrt{\frac{P_\ell}{N_t}} \boldsymbol{H}_\ell \boldsymbol{X}_\ell + \boldsymbol{W}_\ell, \tag{1}$$

where P_ℓ is the transmit power in round ℓ, $\boldsymbol{X}_\ell \in \mathbb{C}^{BN_t \times J}$, $\boldsymbol{Y}_\ell, \boldsymbol{W}_\ell \in \mathbb{C}^{BN_r \times J}$ are correspondingly the transmitted signal, the received signal, and the additive noise; while $\boldsymbol{H}_\ell \in \mathbb{C}^{BN_r \times BN_t}$ is a block diagonal channel gain matrix at round ℓ with

$$\boldsymbol{H}_\ell = \mathrm{diag}(\boldsymbol{H}_{\ell,1}, \ldots, \boldsymbol{H}_{\ell,B}).$$

In the INR-ARQ scheme, the receiver attempts to decode at round ℓ based on the received signals collected in rounds $1, \ldots, \ell$. The entire channel after ℓ ARQ rounds is

$$\boldsymbol{Y}_{\overline{1,\ell}} = \boldsymbol{H}_{\overline{1,\ell}} \boldsymbol{X}_{\overline{1,\ell}} + \boldsymbol{W}_{\overline{1,\ell}}, \tag{2}$$

where

$$\begin{aligned}
\boldsymbol{Y}_{\overline{1,\ell}} &= [\boldsymbol{Y}'_1, \ldots, \boldsymbol{Y}'_\ell]' \\
\boldsymbol{X}_{\overline{1,\ell}} &= [\boldsymbol{X}'_1, \ldots, \boldsymbol{X}'_\ell]' \\
\boldsymbol{H}_{\overline{1,\ell}} &= \mathrm{diag}\left(\sqrt{\frac{P_1}{N_t}} \boldsymbol{H}_1, \ldots, \sqrt{\frac{P_\ell}{N_t}} \boldsymbol{H}_\ell \right) \\
\boldsymbol{W}_{\overline{1,\ell}} &= [\boldsymbol{W}'_1, \ldots, \boldsymbol{W}'_\ell]' .
\end{aligned}$$

We consider transmission where the entries of $\boldsymbol{X}_\ell$ are uniformly drawn from an input constellation $\mathcal{X} \subset \mathbb{C}$ of size 2^M, and assume that the constellation $\mathcal{X}$ has unit average energy, i.e., entries $x \in \mathcal{X}$ of $\boldsymbol{X}_\ell$ satisfy $\mathbb{E}\left[|x|^2\right] = 1$. We further assume that the entries of $\boldsymbol{H}_{\ell,b}$ and $\boldsymbol{W}_\ell$ are independently drawn from a unit-variance Gaussian complex distribution $\mathcal{N}_{\mathbb{C}}(0,1)$, and that $\boldsymbol{H}_{\ell,b}$ is available at the receiver. The average SNR at each receive antenna is then P_ℓ.

We consider ARQ transmission with a long-term power constraint, where the average power is defined as the average energy normalized by J [23]. With long-term power constraint P,

$$\mathbb{E}_{\boldsymbol{H}_{\overline{1,L}}}\left[\sum_{\ell=1}^{L} P_\ell\right] \leq P, \tag{3}$$

where P_ℓ is adapted to $\boldsymbol{H}_{\overline{1,\ell-1}}$ through receiver feedback. Strictly speaking, the long-term average power is

$$P_{\text{av}} = \frac{\mathbb{E}_{\boldsymbol{H}_{\overline{1,L}}}\left[\sum_{\ell=1}^{L} P_\ell\right]}{\overline{L}}, \tag{4}$$

where $\overline{L}$ is the average number of transmission round per codeword, or equivalently the expected inter-renewal time [6]. For simplicity, we will study the system with power constraint given in (3). However, as will be shown, the analysis is also valid when the average power in (4) is constrained.

III. Preliminaries

A. Accumulated Mutual Information

Assuming that the realized channel matrix at round ℓ is $\boldsymbol{H}_\ell$, the input-output mutual information of the MIMO channel in round ℓ is

$$I_\ell\left(\sqrt{\frac{P_\ell}{N_t}}\boldsymbol{H}_\ell\right) = \frac{1}{B}\sum_{b=1}^{B} I_{\mathcal{X}}\left(\sqrt{\frac{P_\ell}{N_t}}\boldsymbol{H}_{\ell,b}\right), \tag{5}$$

where $I_{\mathcal{X}}\left(\sqrt{\frac{P_\ell}{N_t}}\boldsymbol{H}_{\ell,b}\right)$ is the input-output mutual information [5], measured in bits per channel use (bpcu), of an AWGN MIMO channel with input constellation $\mathcal{X}$ and channel matrix $\sqrt{\frac{P_\ell}{N_t}}\boldsymbol{H}_{\ell,b}$. More specifically,

$$I_{\mathcal{X}}\left(\boldsymbol{H}\right) = \mathbb{E}_{\boldsymbol{x},\boldsymbol{w}}\left[\log_2 \frac{e^{-\|\boldsymbol{w}\|^2}}{\sum_{\boldsymbol{x}'\in\mathcal{X}^{N_t}} \frac{1}{2^M} e^{-\|\boldsymbol{w}-\boldsymbol{H}(\boldsymbol{x}-\boldsymbol{x}')\|^2}}\right], \tag{6}$$

where $\boldsymbol{x}$ is uniformly drawn from $\mathcal{X}^{N_t}$ and the entries of $\boldsymbol{w} \in \mathbb{C}^{N_r}$ are i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$. The average input-output mutual information after ℓ ARQ rounds is given by $\frac{1}{\ell}\sum_{l=1}^{\ell} I_l$ bpcu. Let

$$I_{\overline{1,\ell}} \triangleq \sum_{l=1}^{\ell} I_l \tag{7}$$

be the *accumulated mutual information* after ℓ ARQ rounds. We now propose the multi-bit feedback INR-ARQ transmission scheme based on the accumulated mutual information $I_{\overline{1,\ell}}$.

B. Multi-Level Feedback

We consider an INR-ARQ system with a delay constraint of L ARQ rounds, where a feedback index $k \in \{0, \dots, K-1\}$ is delivered after each transmission round through a zero-delay error-free feedback channel. Power and rate adaptation are performed based on receiver feedbacks. The overall system model is illustrated in Figure 1.

1) Transmitter: Consider a code book $\mathcal{C}$ of rate $\frac{R}{L}$, $R \in (0, MN_t)$ bits per coded symbol, that maps a message $m \in \{1, \dots, 2^{RBJ}\}$ to a codeword $\boldsymbol{x}(m) \in \mathcal{X}^{N_t BJL}$. At transmission round ℓ, $N_t BJ$ of the coded symbols are formatted into $\boldsymbol{X}_\ell(m) \in \mathcal{X}^{BN_t \times J}$ and transmitted via the channel in (1) with power $P_\ell(\boldsymbol{k}_{\ell-1})$, where $\boldsymbol{k}_{\ell-1} = [k_1, \dots, k_{\ell-1}]$ is the vector of feedback indices collected from rounds $1, \dots, \ell-1$. The realized code rate of a single ARQ round is R bpcu, and the realized code rate after ℓ ARQ rounds is $\frac{R}{\ell}$ bpcu. If feedback $k_\ell = K-1$ (denoting positive acknowledgment (ACK)) is received after ℓ transmission rounds, the transmission is successful and transmission of the next message starts. Otherwise, the transmitter continues with new transmission rounds until feedback index $K-1$ is received or until L transmission rounds have elapsed.

2) Receiver: Upon receiving round ℓ, the receiver attempts to decode the transmitted message from the received signals collected from rounds 1 to ℓ. The receiver employs a decoder with error detection capabilities as described in [6]. The decoder outputs $\hat{m} \in \{1, \dots, 2^{RBJ}\}$ if there exists a unique message $\hat{m}$ such that $\boldsymbol{X}_{\overline{1,\ell}}(\hat{m})$ and $\boldsymbol{Y}_{\overline{1,\ell}}$ are jointly typical conditioned on $\boldsymbol{H}_{\overline{1,\ell}}$ [5]; then an ACK is delivered to the transmitter via feedback index $k_\ell = K-1$. Otherwise, a quantization of the *accumulated mutual information* $I_{\overline{1,\ell}}$ is delivered via feedback index k_ℓ satisfying $I_{\overline{1,\ell}} \in \big[\ \overline{I}([\boldsymbol{k}_{\ell-1}, k_\ell]), \overline{I}([\boldsymbol{k}_{\ell-1}, k_\ell+1])\big)$, with predefined quantization thresholds $\overline{I}(\boldsymbol{k}_\ell), \boldsymbol{k}_\ell \in \{0, \dots, K-2\}^\ell$, and $\overline{I}([\boldsymbol{k}_{\ell-1}, K-1]) = \infty$ for $\ell = 1, \dots, L-1$. An example of the feedback thresholds for the first two rounds of an ARQ system with $K = 4$ is illustrated in Figure 2. Feedback index 3 is used to denote successful transmission. At the first ARQ round, the leftmost set of feedback thresholds is used; while at the second ARQ round, one of the three sets of feedback thresholds on the right is employed, depending on which feedback index was delivered in the first round. Noting that $I_{\overline{1,\ell+1}} \geq I_{\overline{1,\ell}}$, the feedback thresholds in round $\ell+1$ should be designed such that $\overline{I}(\boldsymbol{k}_\ell) = \overline{I}([\boldsymbol{k}_\ell, 0]) < \ldots < \overline{I}([\boldsymbol{k}_\ell, K-2])$. Thus, the set of quantization thresholds is completely defined by $\overline{I}(\boldsymbol{k}_{L-1})$ for practical purposes.

3) Power constraint: The probability of having feedback vector $\boldsymbol{k}_\ell$ at round ℓ, denoted as $q(\boldsymbol{k}_\ell)$, is recursively expressed as

$$q(\boldsymbol{k}_0) = 1 \tag{8}$$

$$q([\boldsymbol{k}_{\ell-1}, k]) = \Pr\left\{k_\ell = k | \boldsymbol{k}_{\ell-1}\right\} q(\boldsymbol{k}_{\ell-1}), \tag{9}$$

$$\Pr\left\{k_\ell = k \middle| \boldsymbol{k}_{\ell-1}\right\} = \Pr\left\{I_{\overline{1,\ell-1}} + I_\ell \in \left[\ \overline{I}([\boldsymbol{k}_{\ell-1}, k]), \overline{I}([\boldsymbol{k}_{\ell-1}, k+1])\right) \middle| \boldsymbol{k}_{\ell-1}\right\},$$

where I_ℓ is given by (5) with $P_\ell = P_\ell(\boldsymbol{k}_{\ell-1})$. Noting that $k_{\ell-1} = K-1$ denotes a successful decoding at round $\ell - 1$, the power constraint in (3) can be written as

$$\mathbb{E}_{\boldsymbol{H}_{\overline{1,L}}}\left[\sum_{\ell=1}^{L} P_\ell\right] = P_1 + \sum_{\ell=2}^{L} \sum_{\boldsymbol{k}_{\ell-1}\in\{0,\ldots,K-2\}^{\ell-1}} q(\boldsymbol{k}_{\ell-1}) P_\ell(\boldsymbol{k}_{\ell-1}) \le P; \tag{10}$$

while the average power in (4) is equivalent to

$$P_{\text{av}} = \frac{P_1 + \sum_{\ell=2}^{L} \sum_{\boldsymbol{k}_{\ell-1}\in\{0,\ldots,K-2\}^{\ell-1}} q(\boldsymbol{k}_{\ell-1}) P_\ell(\boldsymbol{k}_{\ell-1})}{1 + \sum_{\ell=2}^{L} \sum_{\boldsymbol{k}_{\ell-1}\in\{0,\ldots,K-2\}^{\ell-1}} q\left(\boldsymbol{k}_{\ell-1}\right)}. \tag{11}$$

C. *Information Outage*

After ℓ ARQ rounds, the input-output mutual information is $\frac{I_{\overline{1,\ell}}}{\ell}$ and the realized code rate is $\frac{R_M N_t}{\ell} = \frac{R}{\ell}$ (bpcu). Hence, transmission is in outage at round ℓ if $I_{\overline{1,\ell}} < R$. The probability of having an outage at round ℓ is then given by

$$p(\ell) \triangleq \Pr\left\{I_{\overline{1,\ell}} < R\right\}. \tag{12}$$

With an optimal coding scheme, and in the limit of the number of channel uses $J \to \infty$, the codeword is correctly decoded whenever $I_{\overline{1,\ell}} > R$; otherwise, an error is detected [6]. Therefore, the outage probability $p(\ell)$ is an achievable lower bound on the word error probability at round ℓ. For INR-ARQ transmission with delay constraint L, the overall outage probability is $p(L)$.

IV. Asymptotic Analysis

Consider a power adaptation rule $P_\ell = P_\ell(\boldsymbol{k}_{\ell-1})$ satisfying the power constraint in (10). We prove that for large P, the optimal outage probability at round ℓ behaves like

$$p(\ell) \doteq P^{-d_\ell(R)}, \tag{13}$$

where $d_\ell(R)$ is the outage diversity at round ℓ and the exponential equality ($\doteq$) indicates [12]

$$d_\ell(R) = \lim_{P\to\infty} \frac{-\log p(\ell)}{\log P}. \tag{14}$$

It is trivial that the optimal power allocation rule satisfies (10) with equality, and consequently, the average transmit power in (11) satisfies $\frac{P}{L} \leq P_{\text{av}} \leq P$. Therefore,

$$\lim_{P_{\text{av}}\to\infty} \frac{-\log p(\ell)}{\log P_{\text{av}}} = d_\ell(R), \tag{15}$$

and thus $d_\ell(R)$ is also the optimal outage diversity with respect to the average power P_{av}.

Subsequently, we determine the optimal rate-diversity-delay tradeoff $d_\ell(R)$ of ARQ systems with K levels feedback and prove that the optimal outage diversity is achievable.

A. MIMO Block-Fading without ARQ

In order to characterize the outage diversity or achievable SNR-exponent for the MIMO INR-ARQ channel, we first need to study the corresponding limits for fixed-rate transmission over the MIMO block-fading channel. These results are key elements to proving our main results for multi-bit ARQ.

Theorem 1: Consider fixed-rate transmission ($L = 1$) with rate R and power P over the MIMO block-fading channel in (1) using constellation $\mathcal{X}$ of size 2^M and the transmission scheme described in Section III-B. Let $I = I_1\left(\sqrt{\frac{P}{N_t}}\boldsymbol{H}_1\right)$ be the realized input-output mutual information as defined in (5). For large P, we have that

$$\Pr\{I < R\} \doteq P^{-d(R)}, \tag{16}$$

$$\Pr\{I \leq R\} \doteq P^{-\underline{d}(R)}, \tag{17}$$

where $d(R)$ is bounded by $\underline{d}(R) \leq d(R) \leq \overline{d}(R)$, and

$$\overline{d}(R) \triangleq N_r\left(1 + \left\lfloor B\left(N_t - \frac{R}{M}\right)\right\rfloor\right) \tag{18}$$

$$\underline{d}(R) \triangleq N_r\left\lceil B\left(N_t - \frac{R}{M}\right)\right\rceil. \tag{19}$$

Furthermore, $\underline{d}(R)$ is the SNR-exponent achieved by using random codes with rate R, where the code symbols are drawn uniformly from $\mathcal{X}$.

Proof: See Appendix A[1]. ∎

[1] A more general result of the theorem, which deals with power allocation for block-fading channels with mismatched channel state information, was derived in [24] after the submission of this paper. The proof given here is simpler and forms a basis for the result in [24].

To the best of our knowledge, this is the first rigorous proof for the outage diversity of a MIMO block-fading channel with a general discrete input constellation. The results of [13, 15] establish $\overline{d}(R)$ as an upper bound for the outage diversity for the quasi-static fading channel. Code design techniques in [15] show that $\overline{d}(R)$ can be achieved by specifically constructed input constellations. As a generalization, Theorem 1 shows that $\overline{d}(R)$ is the outage diversity for MIMO block-fading channels with any input constellation of size 2^M (except when $\frac{BR}{M}$ is an integer). Furthermore, Theorem 1 shows that $\overline{d}(R)$ is achievable by using random codes when $\frac{BR}{M}$ is non-integer, which is essential for analyzing the performance of INR-ARQ systems.

B. *Multi-bit MIMO ARQ*

We now consider ARQ transmission over the block-fading channel in (1) using input constellation $\mathcal{X}$ as described in Section III-B1. Using Theorem 1, the optimal rate-diversity-delay tradeoff of the MIMO INR-ARQ scheme with multi-bit feedback is characterize as follows.

Theorem 2: Consider INR-ARQ transmission over the MIMO block-fading channel in (1) using constellation $\mathcal{X}$ of size 2^M and the transmission scheme described in Section III-B, where a codeword is considered successfully delivered at round ℓ if $I_{\overline{1,\ell}} \geq R$. Assume that the number of feedback levels is $K \geq \left\lceil \frac{BR}{M} \right\rceil + 1$. Subject to the power constraint in (10), the optimal rate-diversity-delay tradeoff is given by

$$d_\ell(R) = (1 + BN_tN_r)^{\ell-1} \left(\overline{d}(R) + 1\right) - 1 \tag{20}$$

when $\frac{BR}{M}$ is not an integer, where $\overline{d}(R)$ is given in Theorem 1.

Proof: See Appendix B for a proof. ■

Theorem 2 only gives the optimal outage diversity when $\frac{BR}{M}$ is non-integer. When $\frac{BR}{M}$ is an integer, the bounds for $d(R)$ in Theorem 1 does not coincide, thus a definite value of $d_\ell(R)$ is not known. It can be shown that the optimal outage diversity is bounded by

$$(1 + BN_tN_r)^{\ell-1} \left(\underline{d}(R) + 1\right) - 1 \leq d_\ell(R) \leq (1 + BN_tN_r)^{\ell-1} \left(\overline{d}(R) + 1\right) - 1. \tag{21}$$

An intuitive explanation for the outage diversity gains offered by multi-bit feedback is given as follows. At round $\ell + 1$, the feedback vector $\boldsymbol{k}_\ell$ provides the transmitter with the past channel realizations. This allows raising the transmit power in round $\ell + 1$ by a factor of

$\frac{1}{q(\boldsymbol{k}_\ell)}$ without violating the long-term power constraint. In the limit of large power constraint, the optimal transmit power in round $\ell+1$ satisfies

$$P_{\ell+1} \doteq P_\ell^{\overline{d}\left(\overline{I}(\boldsymbol{k}_\ell)-\overline{I}(\boldsymbol{k}_{\ell-1})\right)}, \tag{22}$$

where $\overline{d}(R)$ is given by Theorem 1. Since only the exponent is significant in diversity analysis, the maximum outage diversity can be achieved if there are sufficient thresholds to feedback $\overline{d}\,(I)$ for $I \le R$ ($I > R$ implies successful transmission). For MIMO block-fading channels with discrete input constellation, the rate-diversity tradeoff is a stair-case function; therefore a finite number of feedback levels is sufficient to achieve the maximum outage diversity. Meanwhile, for systems with $\frac{BR}{M} < 1$, $\overline{d}(I)$ is a constant for $I < R$, especially systems with Gaussian input constellations,. Therefore, no gains in outage diversity can be obtained by multi-bit feedback. Conversely, in multiplexing-diversity tradeoff analysis [12], the outage diversity is a continuous, decreasing function of the multiplexing gain, and thus an infinite number of feedback levels is required to achieve the optimal outage diversity.

Remark 1: The proof of Theorem 2 also gives the following guidelines to designing the feedback and power allocation rules.

- The optimal outage diversity of INR-ARQ systems is achievable with $\lceil \frac{BR}{M} \rceil + 1$ feedback levels, where the feedback thresholds of each round are fixed at $\hat{I}_t = \frac{Mt}{B}, t = 0, \ldots, \lfloor \frac{BR}{M} \rfloor$. Therefore, for systems with $K \ge \lceil \frac{BR}{M} \rceil + 1$, the optimal outage diversity is achievable if for $\ell = 1, \ldots, L$, $\{\hat{I}_t : R \ge \hat{I}_t \ge \overline{I}(\boldsymbol{k}_{\ell-1})\} \subseteq \{\overline{I}(\boldsymbol{k}_\ell), \boldsymbol{k}_\ell \in \{1, \ldots, K-1\}^\ell\}$.
- Furthermore, the outage probability in round $\ell+1$ is dominated by the events with $I_{\overline{1,\ell}} \in \left[0, \frac{M}{B}\right) \cup \left[\hat{I}_\tau, R\right)$, where $\tau = \lfloor \frac{BR}{M} \rfloor$. Therefore, after placing $\lceil \frac{BR}{M} \rceil + 1$ thresholds at $\hat{I}_t$, the remaining feedback thresholds (for systems with $K > \lceil \frac{BR}{M} \rceil + 1$) should give higher priority to quantizing the aforementioned region to improve outage performance.
- With the optimal feedback rule, the optimal outage diversity can be achieved with power allocation satisfying $q(\boldsymbol{k}_{\ell-1})P_\ell(\boldsymbol{k}_{\ell-1}) = \alpha P$ for all $\boldsymbol{k}_{\ell-1}, \ell = 1, \ldots, L$, where α is a constant chosen to satisfy the power constraint (10).

We now prove that the rate-diversity-delay tradeoff $d_\ell(R)$ is achievable by using random codes, as given by the following theorem.

Theorem 3: Consider INR-ARQ transmission over the MIMO block-fading channel in (1) using constellation $\mathcal{X}$ of size 2^M and the transmission scheme described in Section III-B with power constraint P given in (10). Assume that the number of feedback levels is $K \ge \lceil \frac{BR}{M} \rceil + 1$. With random-coding schemes and $J \to \infty$, for large P, the word error probability $P_e(\ell)$ at round ℓ satisfies $P_e(\ell) \doteq P^{-d_\ell^{(r)}(R)}$, where

$$d_\ell^{(r)}(R) = (1 + BN_tN_r)^{\ell-1} \left(\underline{d}(R) + 1\right) - 1 \tag{23}$$

is the achievable SNR-exponent and $\underline{d}(\ell)$ is given in Theorem 1.

Proof: With a random coding scheme and $J \to \infty$, the codeword is correctly decoded with probability one at round ℓ if $I_{\overline{1,\ell}} > R$ [6, 25], in which case, the receiver feeds back an ACK (in contrast to the outage case, where an ACK is fed back if $I_{\overline{1,\ell}} \geq R$). The proof then follows similar arguments as the proof of Theorem 2, noting from Theorem 1 that $\Pr\{I_\ell \leq I\} \doteq P_\ell^{-\underline{d}(I)}$. ■

Theorem 3 shows that the rate-diversity-delay tradeoff $d_\ell(R)$ stated in Theorem 2 is achievable with random codes using the transmission scheme described in Section III-B when $\frac{BR}{M}$ is not an integer; and then, the optimal rate-diversity-delay tradeoff is given by (23). Furthermore, the optimal outage diversity and SNR-exponent of INR-ARQ transmission with delay constraint L is similarly characterized by $d_L(R)$ and $d_L^{(r)}(R)$ given in (20) and (23), respectively.

C. *One-bit MIMO ARQ*

In an INR-ARQ system with one-bit ACK/NACK feedback (classical INR-ARQ), the optimal rate-diversity-delay tradeoff is given by the following.

Theorem 4: Consider INR-ARQ transmission over the MIMO block-fading channel in (1) using constellation $\mathcal{X}$ of size 2^M and the transmission scheme described in Section III-B, where a codeword is considered successfully delivered at round ℓ if $I_{\overline{1,\ell}} \geq R$. Assume that the number of feedback levels is $K = 2$. Subject to the power constraint in (10), the optimal rate-diversity-delay tradeoff is given by

$$\hat{d}_1(R) = \overline{d}(R) \tag{24}$$

$$\hat{d}_\ell(R) = BN_tN_r\left(\ell - 1 + \sum_{l=1}^{\ell-2} \hat{d}_l(R)\right) + (1 + \hat{d}_{\ell-1}(R))\hat{d}_1(R), \quad \ell \geq 2. \tag{25}$$

for all R such that $\hat{d}_1(R)$ is continuous. Furthermore, the rate-diversity-delay tradeoff $\hat{d}_\ell(R)$ is achievable when $\frac{BR}{M}$ is not an integer.

Proof: The proof follows the same arguments as that of Theorems 2 and 3, with only two feedback levels at 0 and R, respectively. ■

Theorem 4 characterizes the optimal outage diversity for INR-ARQ systems with $K = 2$ when $\frac{BR}{M}$ is non-integer. When $\frac{BR}{M}$ is integer, the outage diversity at round ℓ is upper bounded by

$\hat{d}_\ell(R)$ given in (25). A lower bound on the outage diversity is given by the recursive formula in (25) with $\hat{d}_1(R) = \underline{d}(R)$.

D. Numerical Results

We numerically compare the optimal rate-diversity-delay tradeoff of INR-ARQ systems with $K \geq \lceil \frac{BR}{M} \rceil + 1$, and with $K = 2$ as well as the optimal tradeoff of an INR-ARQ system with constant transmit power. The optimal rate-diversity-delay tradeoff $d_L(R)$ and $\hat{d}_L(R)$ for INR-ARQ transmission with $L = 1, 2, 3$ over the MIMO block-fading channel with $N_t = N_r = B = 2$ are illustrated in Figure 3(a).

For an INR-ARQ system with delay constraint L and constant transmit power (short-term power constraint), the outage probability $p(L)$ is the same as that obtained by transmission with rate $\frac{R}{L}$ over a block-fading channel with BL fading blocks [21]. From Theorem 1, the optimal outage diversity $\underline{d}_L(R)$ is given by[2]

$$\underline{d}_L(R) = N_r \left(1 + \left\lfloor BL \left(N_t - \frac{R}{LM}\right)\right\rfloor\right), \tag{26}$$

and is achievable by random codes for all rates R such that $\underline{d}_L(R)$ is continuous. The rate-diversity-delay tradeoff of the INR-ARQ system with constant transmit power is plotted in Figure 3(b). Figure 3 shows an order-of-magnitude improvement in outage diversity of INR-ARQ when a long-term power constraint is allowed. Furthermore, significant gains in outage diversity are provided by multi-bit feedback, especially at transmission rates R close to $N_t M$. Since high R is particularly relevant in ARQ systems, the result suggests that multi-bit feedback will give significant gains in practical implementations.

V. Power Adaptation and Feedback Design

The design of optimal feedback and transmission rules for an ARQ system with multi-bit feedback includes joint optimization of the overall set of quantization thresholds $\{\overline{I}(\boldsymbol{k}_{L-1}), \boldsymbol{k}_{L-1} \in \{0, \ldots, K-2\}^{L-1}\}$ and the corresponding power adaptive rule $P_\ell(\boldsymbol{k}_{\ell-1})$. The optimal feedback and power adaptation rule is obtained by minimizing

$$\sum_{\boldsymbol{k}_{L-1}} q(\boldsymbol{k}_{L-1}) p(L|\boldsymbol{k}_{L-1}) \tag{27}$$

[2]The rate-diversity-delay tradeoff of [21] is larger than that given in (26) since it is obtained with rotations, which increase the constellation size, complexity and peak-to-average power ratio.

subject to the power constraint in (10). To the best of our knowledge, the optimization problem is not analytically tractable. We therefore propose to partition the design problem into two steps.

Step 1: At round ℓ, determine a set of feedback thresholds $\overline{I}([\boldsymbol{k}_{\ell-1}, k])$ for every feedback vector $\boldsymbol{k}_{\ell-1} \in \{1, \ldots, K-2\}^{\ell-1}$.

Step 2: Given the set of feedback thresholds in Step 1, determine the corresponding transmit power rule, minimizing the outage probability.

The above procedure sub-optimally partitions the joint optimization problem into two sequential problems. Moreover, in the following, each individual problem is also sub-optimally solved. Nevertheless, this design procedure leads to a practically implementable algorithm that achieves the optimal diversity derived in the previous section.

A. Selecting the Set of Feedback Thresholds

From the observations in Remark 1, we propose the following choice of feedback thresholds. Consider the feedback levels at round ℓ for a given feedback vector $\boldsymbol{k}_{\ell-1}$. Let $\tau \triangleq \left\lfloor \frac{BR}{M} \right\rfloor$, $\hat{I}_t = \frac{Mt}{B}$ and $t' \triangleq \left\lfloor \frac{B\overline{I}(\boldsymbol{k}_{\ell-1})}{M} \right\rfloor$. The feedback thresholds in round ℓ, given $\boldsymbol{k}_{\ell-1}$ is then determined as follows.

1) Place a threshold at $\overline{I}([\boldsymbol{k}_{\ell-1}, 0]) = \overline{I}(\boldsymbol{k}_{\ell-1})$, and at $\overline{I}([\boldsymbol{k}_{\ell-1}, K-1]) = R$;
2) Place $\tau - t'$ thresholds at $\hat{I}_t, t = t'+1, \ldots, \tau$;
3) Place the remaining $K - 2 - \tau + t'$ thresholds sequentially within

$$\left(\hat{I}_\tau, R\right), \left(\overline{I}(\boldsymbol{k}_{\ell-1}), \hat{I}_{t'+1}\right), \left(\hat{I}_{\tau-1}, \hat{I}_\tau\right), \left(\hat{I}_{t'+1}, \hat{I}_{t'+2}\right), \ldots$$

until no more thresholds are left to place, and such that the thresholds uniformly partition each region.

The procedure for choosing the thresholds $\overline{I}(\boldsymbol{k}_\ell)$, given the feedback vector $\boldsymbol{k}_{\ell-1}$, is illustrated in Figure 4. More particularly, the feedback thresholds for INR-ARQ transmission over the block-fading channel with $N_t = N_r = 1$, $B = 2$, $K = 4$, $L = 2$, and $R = 3.5$ using 16-QAM constellations are illustrated in Figure 2, where $\overline{I}(\boldsymbol{k}_{\ell-1}) = \overline{I}([\boldsymbol{k}_{\ell-1}, 0])$, and the values of $\overline{I}(\boldsymbol{k}_2)$ are reported in Table II.

B. Power Adaptation

The sub-optimal power adaptation rule is obtained from the following simplifications.

- We consider a power constraint more stringent than the constraint in (10),

$$\sum_{\boldsymbol{k}_\ell \in \{0,\dots,K-1\}^\ell} q(\boldsymbol{k}_\ell) P_{\ell+1}(\boldsymbol{k}_\ell) \leq \frac{P}{L}, \tag{28}$$

  for $\boldsymbol{k}_\ell \in \{0,\dots,K-1\}^\ell, \ell = 0,\dots,L-1$, where $q(\boldsymbol{k}_0) = 1$ by definition.
- When feedback $\boldsymbol{k}_{\ell-1}$ is received, we have that $I_{\overline{1,\ell-1}} \geq \overline{I}(\boldsymbol{k}_{\ell-1})$. Then, the feedback probability is approximated from (9) by replacing $I_{\overline{1,\ell-1}}$ with $\overline{I}(\boldsymbol{k}_{\ell-1})$; and the outage probability can be upper bounded as

$$\hat{p}(\ell|\boldsymbol{k}_{\ell-1}) \triangleq \Pr\left\{ I_\ell + \overline{I}(\boldsymbol{k}_{\ell-1}) < R \right\}, \tag{29}$$

  where I_ℓ is given by (5) with $P_\ell = P_\ell(\boldsymbol{k}_{\ell-1})$.
- To further simplify the problem, we consider minimizing $\hat{p}(\ell), \ell = 1,\dots,L$ sequentially.

Based on the simplifications, the corresponding power adaptation rule $P_\ell(\boldsymbol{k}_{\ell-1})$ is obtained by solving

$$\begin{cases} \text{Minimize} & \sum_{\boldsymbol{k}_{\ell-1}} q(\boldsymbol{k}_{\ell-1}) \hat{p}(\ell|\boldsymbol{k}_{\ell-1}) \\ \text{Subject to} & \sum_{\boldsymbol{k}_{\ell-1}} q(\boldsymbol{k}_{\ell-1}) P_\ell(\boldsymbol{k}_{\ell-1}) \leq \frac{P}{L}. \end{cases} \tag{30}$$

The optimization problem is separable, and thus can be solved via a branch-and-bound simplex algorithm using piece-wise linear approximation [26]. For single-input multiple-output (SIMO) channels, the probabilities $q(\boldsymbol{k}_{\ell-1})$ and $\hat{p}(\ell|\boldsymbol{k}_{\ell-1})$ in (30) can be approximated numerically by shifting the outage probability bounds in [27] according to the gap between the bound and the corresponding simulation curve at high SNR. For MIMO channels, solving (30) requires tabulating the probabilities $q(\boldsymbol{k}_{\ell-1})$ and $\hat{p}(\ell|\boldsymbol{k}_{\ell-1})$, which can be obtained from Monte-Carlo simulations.

C. Numerical Results

First consider SISO ($N_t = N_r = 1$) INR-ARQ transmission with $L = 2$ at rate $R = 3.5$ over the block-fading channel in (1) with $B = 2$ using 16-QAM input constellations. The outage performance of systems with $K = 2, 3, 8, 16$ is illustrated in Figure 5. We observe that the outage diversity achieved by constant transmit power and by power adaptation for $K = 2$ is 3 and 4 as given in (26) and (25), respectively. For $K \geq 3$, the outage diversity is 5 as predicted from (20). This leads to significant improvement in outage performance for power adaptive ARQ transmission with multi-bit feedback at high P. Particularly, 2 dB gain in power is observed at outage probability 10^{-6} for $K \geq 8$. Note that at low P, the outage performance of systems with $K = 2$ is outperformed by system with constant transmit power due to the simplifying assumption (28).

The outage performance of MIMO INR-ARQ transmission over the block-fading channel in (1) with $N_t = 2, N_r = 1, B = 1, R = 7.5$ using 16-QAM input constellations is illustrated in Figures 6 and 7, where Figure 6 shows the simulation results, and Figure 7 presents the upper bound obtained from (30). The simulation results in Figure 6 have yet to show the correct outage diversity ($d_L(R) = 5$ for $K \geq 3$ and $d_L(R) = 4$ for $K = 2$). However, they follow the bounds from (30), which approach the optimal outage diversity at higher SNR as shown in Figure 7. Figure 6 shows that systems with power allocation significantly outperform that with constant transmit power. Moreover, allowing additional feedback levels ($K \geq 3$) provides further gains in outage diversity and thus significant gains in outage performance at high SNR.

In both cases, the simulation results suggest that increasing K beyond 8 does not substantially improve the outage performance; and thus, even for $K = 3$, the sub-optimal choice of feedback thresholds in Section V-A performs within 1dB of systems with large K and optimal thresholds.

VI. Conclusions

We have studied the outage performance of MIMO block-fading channels with and without employing the INR-ARQ strategy. An information-theoretic multi-bit feedback INR-ARQ scheme is proposed based on the accumulative mutual information, which potentially improves the performance of INR-ARQ transmission with minimal extra overhead requirement compared to classical INR-ARQ. The study on power adaptation has revealed large gains in outage diversity provided by multi-bit feedback in INR-ARQ systems with a long-term power constraint. More generally, the multi-bit feedback INR-ARQ based on accumulated mutual information may prove useful in obtaining the fundamental limit of multi-bit feedback INR-ARQ systems. Furthermore, since the proposed scheme is a generalization to that in [17] and [19], it promises further gain from the throughput performance obtained in [17, 19].

Appendix A
Proof of Theorem 1

We first assume a genie-aided receiver that perfectly eliminates the interference between the transmit antennas. This results in N_t parallel SIMO block-fading channels, each with N_r

receive antennas. Let $I^{\rm ga}$ be the realized input-output mutual information of the genie-aided channel, then $I^{\rm ga} \geq I$. Furthermore, from the analysis in [25, 27, 28], we have that

$$\Pr\left\{I^{\rm ga} < R\right\} \doteq P^{-\overline{d}(R)}. \tag{31}$$

Therefore,

$$\Pr\left\{I < R\right\} \dot{\geq} P^{-\overline{d}(R)}. \tag{32}$$

The proof is thus completed by proving that

$$\Pr\left\{I \leq R\right\} \doteq P^{-\underline{d}(R)}. \tag{33}$$

Following the arguments in [25, 27, 28], we have that

$$\Pr\left\{I^{\rm ga} \leq R\right\} \doteq P^{-\underline{d}(R)} \tag{34}$$

and therefore,

$$\Pr\left\{I \leq R\right\} \dot{\geq} P^{-\underline{d}(R)}. \tag{35}$$

We now prove that $\Pr\left\{I \leq R\right\} \dot{\leq} P^{-\underline{d}(R)}$. Considering transmission over the block-fading channel in (1) with random codes of rate R, where the JBN_t coded symbols in $\boldsymbol{x}$ are drawn uniformly random from the constellation $\mathcal{X}$. Let $P_e^{(r)}$ be the word error probability achieved by random coding. We have from the random-coding achievability and the strong converse theorem [3–5] that for a channel realization $\boldsymbol{H}$,

$$P_e^{(r)}(\boldsymbol{H}) = \begin{cases} 1 & \text{if } I < R \\ 0 & \text{if } I > R \end{cases} \tag{36}$$

when $J \to \infty$. Therefore, the word error probability of random codes satisfies

$$P_e^{(r)} = \Pr\left\{I \leq R\right\}. \tag{37}$$

We now prove that $P_e^{(r)} \dot{\leq} P_\ell^{-\underline{d}(I)}$. Consider encoding and transmitting a message m as a random codeword $\boldsymbol{X}$. Assuming that the channel realization is $\boldsymbol{H}$, the pairwise error probability between $\boldsymbol{X}$ and $\boldsymbol{X}'$ is bounded by [29]

$$P_{\rm PEP}\left(\boldsymbol{X} \to \boldsymbol{X}' | \boldsymbol{H}\right) \leq \exp\left(-\frac{1}{4} g^2(\boldsymbol{X}, \boldsymbol{X}', \boldsymbol{H})\right), \tag{38}$$

where, by letting $\hat{P} = \frac{P_\ell}{N_t}$,

$$g^2(\boldsymbol{X}, \boldsymbol{X}', \boldsymbol{H}) = \sum_{b=1}^{B} \sum_{j=1}^{J} \sum_{r=1}^{N_r} \left| \sum_{t=1}^{N_t} \sqrt{\hat{P}} h_{b,t,r} (\boldsymbol{X}_{b,t,j} - \boldsymbol{X}'_{b,t,j}) \right|^2. \tag{39}$$

Here, $h_{b,t,r}$ is the channel gain from transmit antenna t to receive antenna r in block b, and $X_{b,t,j}$ is the coded symbol transmitted by antenna t at time instant j of block b. Let us write $h_{b,t,r} = |h_{b,t,r}|e^{i\theta_{b,t,r}}$, where $i = \sqrt{-1}$. Further define a matrix of normalized fading gains $\boldsymbol{\alpha} \in \mathbb{R}^{B\times N_t\times N_r}$ where $\alpha_{b,t,r} \triangleq -\frac{\log(|h_{b,t,r}|^2)}{\log(\hat{P})}$, then

$$g^2(\boldsymbol{X},\boldsymbol{X}',\boldsymbol{H}) = \sum_{b=1}^{B}\sum_{j=1}^{J}\sum_{r=1}^{N_r}\left|\sum_{t=1}^{N_t}\hat{P}^{\frac{1-\alpha_{b,t,r}}{2}}e^{i\theta_{b,t,r}}(\boldsymbol{X}_{b,t,j}-\boldsymbol{X}'_{b,t,j})\right|^2. \tag{40}$$

By averaging (38) over the random coding ensemble, the pairwise error probability of random codes is

$$P^{(r)}_{\text{PEP}}(\boldsymbol{X}\to\boldsymbol{X}'|\boldsymbol{H}) \le \prod_{b=1}^{B}\left\{\frac{1}{2^{2MN_t}}\sum_{\mathbf{x}\in\mathcal{X}^{N_t}}\sum_{\mathbf{x}'\in\mathcal{X}^{N_t}}\exp\left(-\frac{1}{4}\sum_{r=1}^{N_r}\left|\sum_{t=1}^{N_t}\hat{P}^{\frac{1-\alpha_{b,t,r}}{2}}e^{i\theta_{b,t,r}}(x_t-x'_t)\right|^2\right)\right\}^J \tag{41}$$

$$\le \exp\left(BMJ\log(2)\left(-2N_t+\frac{1}{BM}T(\hat{P},\boldsymbol{\alpha})\right)\right), \tag{42}$$

where x_t is the t^{th} entry of vector $\mathbf{x}$ and

$$T(\hat{P},\boldsymbol{\alpha}) \triangleq \sum_{b=1}^{B}\log_2\left(\sum_{\mathbf{x}\in\mathcal{X}^{N_t}}\sum_{\mathbf{x}'\in\mathcal{X}^{N_t}}\exp\left(-\frac{1}{4}\sum_{r=1}^{N_r}\left|\sum_{t=1}^{N_t}\hat{P}^{\frac{1-\alpha_{b,t,r}}{2}}e^{i\theta_{b,t,r}}(x_t-x'_t)\right|^2\right)\right). \tag{43}$$

By summing over the $2^{BRJ}-1$ possible error events, the union bound on the word error probability is given by

$$P_e^{(r)}(\boldsymbol{H}) \le \min\left\{1,\exp\left(BMJ\log(2)\left(-2N_t+\frac{R}{M}+\frac{1}{BM}T(\hat{P},\boldsymbol{\alpha})\right)\right)\right\}. \tag{44}$$

For any $\epsilon > 0$, denote $\mathcal{S}_b^{(\epsilon)} \triangleq \bigcup_{r=1}^{N_r}\mathcal{S}_{b,r}^{(\epsilon)}$, and $\kappa_b \triangleq |\mathcal{S}_b^{(\epsilon)}|$, where

$$\mathcal{S}_{b,r}^{(\epsilon)} \triangleq \{t : \alpha_{b,t,r} \le 1-\epsilon, t = 1,\dots,N_t\}. \tag{45}$$

Then, for any given $r \in \{1,\dots,N_r\}$, and letting $\alpha_{b,r} = \max\{\alpha_{b,t,r}, t \in \mathcal{S}_{b,r}^{(\epsilon)}\}$, we can write

$$\lim_{\hat{P}\to\infty}\sum_{t=1}^{N_t}\hat{P}^{\frac{1-\alpha_{b,t,r}}{2}}e^{i\theta_{b,t,r}}(x_t-x'_t) \ge \lim_{\hat{P}\to\infty}\sum_{\substack{t\in\mathcal{S}_{b,r}^{(\epsilon)}\\ x_t\neq x'_t}}\hat{P}^{\frac{1-\alpha_{b,t,r}}{2}}e^{i\theta_{b,t,r}}(x_t-x'_t) \tag{46}$$

$$\ge \lim_{\hat{P}\to\infty}\hat{P}^{\frac{1-\alpha_{b,r}}{2}}\sum_{\substack{t\in\mathcal{S}_{b,r}^{(\epsilon)}\\ x_t\neq x'_t}}e^{i\theta_{b,t,r}}(x_t-x'_t). \tag{47}$$

Since the $\theta_{b,t,r}$'s are uniformly drawn from $[-\pi,\pi]$, we have that

$$\lim_{\hat{P}\to\infty}\sum_{t=1}^{N_t}\hat{P}^{\frac{1-\alpha_{b,t,r}}{2}}e^{i\theta_{b,t,r}}(x_t-x'_t) = \infty \tag{48}$$

with probability 1 if there exists $t \in \mathcal{S}_{b,r}^{(\epsilon)}$ such that $x_t \neq x_t'$. Noting that $\kappa_b = |\mathcal{S}_b^{(\epsilon)}|$, it follows from (43) that

$$\begin{aligned}\lim_{\hat{P}\to\infty} T(\hat{P},\boldsymbol{\alpha}) &= \sum_{b=1}^{B} \lim_{\hat{P}\to\infty} \log_2 \left(\sum_{\mathbf{x}\in\mathcal{X}^{N_t}} \sum_{\substack{\mathbf{x}'\in\mathcal{X}^{N_t} \\ x_t'=x_t,\forall t\in\mathcal{S}_b^{(\epsilon)}}} \exp\left(-\frac{1}{4} \sum_{r=1}^{N_r} \left| \sum_{t=1}^{N_t} \hat{P}^{\frac{1-\alpha_{b,t,r}}{2}} e^{i\theta_{b,t,r}} (x_t - x_t') \right|^2 \right) \right) \\ &\leq \sum_{b=1}^{B} \log_2 \left(2^{MN_t} 2^{M(N_t-\kappa_b)} \right) \\ &= \sum_{b=1}^{B} M(2N_t - \kappa_b). \end{aligned} \tag{49}$$

Thus, the error probability in (44) is asymptotically upper-bounded by

$$\lim_{\hat{P}\to\infty} P_e^{(r)}(\boldsymbol{H}) \leq \min\left\{1, \exp\left(-BMJ\log(2)\left(\frac{1}{B}\sum_{b=1}^{B}\kappa_b - \frac{R}{M}\right)\right)\right\}. \tag{50}$$

Let $\mathcal{B}^{(\epsilon)} \triangleq \left\{\boldsymbol{\alpha} \in \mathbb{R}^{B\times N_t\times N_r} : \sum_{b=1}^{B} \kappa_b \leq \frac{BR}{M}\right\}$ be the outage set. By averaging over the fading matrix and letting $J \to \infty$, the error probability is bounded by

$$P_e^{(r)} \leq \int_{\boldsymbol{\alpha}\in\mathcal{B}^{(\epsilon)}} f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) d\boldsymbol{\alpha}, \tag{51}$$

where $f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$ is the joint pdf of the random vector $\boldsymbol{\alpha}$. Following the analysis in [28], and letting $J \to \infty$, the SNR-exponent for the case of using random codes is lower bounded by

$$\begin{aligned}\inf_{\boldsymbol{\alpha}\in\mathcal{B}^{(\epsilon)}\cap\mathbb{R}_+^{BN_r\times BN_t}} \left\{\sum_{b=1}^{B}\sum_{t=1}^{N_t}\sum_{r=1}^{N_r} \alpha_{b,t,r}\right\} &= N_r\left(BN_t - \left\lfloor \frac{BR}{M}\right\rfloor\right)(1-\epsilon) &(52)\\ &= N_r\left\lceil B\left(N_t - \frac{R}{M}\right)\right\rceil(1-\epsilon). &(53)\end{aligned}$$

Thus, by letting $\epsilon \downarrow 0$, the outage diversity $\underline{d}(R)$ is achievable using random codes. Therefore we have from (37) that

$$\Pr\{I \leq R\} \dot{\leq} \hat{P}^{-\underline{d}(R)} \doteq P^{-\underline{d}(R)}. \tag{54}$$

Thus, (33) is obtained from (35).

Appendix B
Proof of Theorem 2

A sketch of the proof is given as follows. We first lower-bound the outage diversity by considering a sub-optimal ARQ system with $\underline{K} = \left\lceil \frac{BR}{M} \right\rceil + 1$ feedback levels, where the quantization thresholds are placed at $\overline{I}([\boldsymbol{k}_{\ell-1}, k_\ell]) = \frac{k_\ell M}{B}, k_\ell = 0, \ldots, \left\lfloor \frac{BR}{M} \right\rfloor$. Using Theorem

1, we prove by induction that the outage diversity of the sub-optimal ARQ system at round ℓ is $d_\ell(R)$.

Conversely, consider an optimal INR-ARQ system with $K \geq \left\lceil \frac{BR}{M} \right\rceil + 1$ feedback levels. The outage performance of the system can be improved by adding $\left\lfloor \frac{BR}{M} \right\rfloor + 1$ extra quantization thresholds (and corresponding feedback indices) at $\frac{tM}{B}, t = 0, \ldots, \left\lfloor \frac{BR}{M} \right\rfloor$. Using Theorem 1, we prove by induction that the outage diversity at round ℓ of the improved systems (with $K + \left\lfloor \frac{BR}{M} \right\rfloor + 1$ feedback levels) is also given by $d_\ell(R)$. Therefore, $d_\ell(R)$ is the optimal outage diversity at round ℓ for an ARQ system with $K \geq \left\lceil \frac{BR}{M} \right\rceil + 1$ feedback levels.

A. Lower bound on the optimal outage diversity

To get a lower bound to the outage diversity, consider an ARQ system with $K = \left\lceil \frac{BR}{M} \right\rceil + 1$ feedback levels, where the following (sub-optimal) set of feedback thresholds is employed,

$$\overline{I}(\boldsymbol{k}_\ell) = \begin{cases} \hat{I}_{k_\ell}, & 0 \leq k_\ell < K - 1 \\ R, & k_\ell = K - 1, \end{cases} \tag{55}$$

with $\hat{I}_t = \frac{tM}{B}$. In this case, feedback index $k_\ell = t$ is delivered at round ℓ if $I_{\overline{1,\ell}} \in \left[\hat{I}_t, \hat{I}_{t+1}\right)$, regardless of the realized feedback indices of the previous rounds. At round ℓ, the transmit power is sub-optimally adapted to the feedback index $k_{\ell-1}$ as $P_\ell = P_\ell(k_{\ell-1})$, where

$$P_\ell(k_{\ell-1}) = \begin{cases} \frac{P}{KL \Pr\left\{I_{\overline{1,\ell-1}} \in \left[\hat{I}_{k_{\ell-1}}, \hat{I}_{k_{\ell-1}+1}\right)\right\}}, & k_{\ell-1} < K - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{56}$$

The power adaptation rule in (56) satisfies the power constraint in (10). We now derive the outage diversity achieved by the aforementioned system.

For $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$, we have from Theorem 1 that

$$\Pr\left\{I_1 < I\right\} \doteq \Pr\left\{I_1 \in \left[\hat{I}_t, \hat{I}_{t+1}\right)\right\} \doteq P^{-\delta_1(t)}, \tag{57}$$

where $\delta_1(t) \triangleq d(\hat{I}_{t+1}) = N_r(BN_t - t)$.

For $t = 0, \ldots, BN_t - 1$ and a given $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$, we now prove by induction that for $\ell = 1, \ldots, L$,

$$\Pr\left\{I_{\overline{1,\ell}} < I\right\} \doteq \Pr\left\{I_{\overline{1,\ell}} \in \left[\hat{I}_t, \hat{I}_{t+1}\right)\right\} \doteq P^{-\delta_\ell(t)}, \tag{58}$$

where $\delta_\ell(t) = d_\ell\left(\hat{I}_{t+1}\right)$ is given in (20).

Equation (57) shows that (58) is correct at round 1. Assume now that (58) is correct at round ℓ. From (56) we have that

$$P_{\ell+1}(t) = \frac{P}{KL \Pr\left\{ I_{\overline{1,\ell}} \in \left[\hat{I}_t, \hat{I}_{t+1}\right) \right\}} \doteq P^{1+\delta_\ell(t)}. \tag{59}$$

Therefore, for $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$,

$$\Pr\left\{ I_{\overline{1,\ell+1}} < I \right\} = \sum_{j=0}^{t} \Pr\left\{ I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_j + I - \hat{I}_t\right) \right\} \Pr\left\{ I_{\ell+1} < I - I_{\overline{1,\ell}} \middle| I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_j + I - \hat{I}_t\right) \right\} + \sum_{j=0}^{t} \Pr\left\{ I_{\overline{1,\ell}} \in \left[\hat{I}_j + I - \hat{I}_t, \hat{I}_{j+1}\right) \right\} \Pr\left\{ I_{\ell+1} < I - I_{\overline{1,\ell}} \middle| I_{\overline{1,\ell}} \in \left[\hat{I}_j + I - \hat{I}_t, \hat{I}_{j+1}\right) \right\}. \tag{60}$$

Given $I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_j + I - \hat{I}_t\right)$ and $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$, we have that $I - I_{\overline{1,\ell}} \in \left(\hat{I}_{t-j}, \hat{I}_{t-j+1}\right)$. Therefore, by applying Theorem 1, and noting the transmit power in (59), we have that

$$\Pr\left\{ I_{\ell+1} < I - I_{\overline{1,\ell}} \middle| I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_j + I - \hat{I}_t\right) \right\} \doteq P^{-(1+\delta_\ell(j))N_r(BN_t - t + j)}. \tag{61}$$

Since (58) is assumed at round ℓ, the first summation dominates in (60). Thus from (61), we have that

$$\Pr\left\{ I_{\overline{1,\ell+1}} < I \right\} \doteq \sum_{j=0}^{t} P^{-\delta_\ell(j) - [1+\delta_\ell(j)]N_r(BN_t - t + j)}. \tag{62}$$

The asymptotic exponent in (62) is given by

$$\min_{j=0,\ldots,t} \delta_\ell(j) + \left[1 + \delta_\ell(j)\right] N_r(BN_t - t + j) \tag{63}$$

$$= \min_{j=0,\ldots,t} -1 + (1 + BN_rN_t)^{\ell-1} \left[1 + N_r(BN_t - j)\right] \left[1 + N_r(BN_t - t + j)\right] \tag{64}$$

$$= -1 + (1 + BN_tN_r)^{\ell} \left[1 + N_r(BN_t - t)\right] \tag{65}$$

$$= \delta_{\ell+1}(t), \tag{66}$$

where (64) follows from assumption $\delta_\ell(j) = d_\ell(\hat{I}_{j+1})$ in (58), and (65) follows since the minimum in (64) is achieved with either $j = 0$ or $j = t$. Therefore, from (62),

$$\Pr\left\{ I_{\overline{1,\ell+1}} < I \right\} \doteq P^{-\delta_{\ell+1}(t)}, \tag{67}$$

where $\delta_{\ell+1}(t) = d_{\ell+1}(\hat{I}_{t+1})$ in (20). Thus, (58) is correct for $\ell = 1, \ldots, L$ by induction. Consequently, for any $R \in \left(\hat{I}_\tau, \hat{I}_{\tau+1}\right)$, we have that

$$\Pr\left\{ I_{\overline{1,\ell}} < R \right\} \doteq P^{-\delta_\ell(\tau)} = P^{-d_\ell(\hat{I}_{\tau+1})}, \tag{68}$$

and thus, the diversity in (20) is achieved by the ARQ system with $\tau+2 = \left\lceil \frac{BR}{M} \right\rceil + 1$ feedback levels.

Noting when $\Pr\left\{I_{\overline{1,\ell+1}} < R\right\} \doteq P^{-\delta_\ell(\tau)}$, the outage probability at round ℓ is dominated by the events with $j = 0$ and $j = \tau$ in (62), which correspond to the events with $I_{\overline{1,\ell}} \in \left[0, \hat{I}_1\right) \cup \left[\hat{I}_\tau, R\right)$. The observation is useful for designing the feedback thresholds for the system, as summarized in Remark 1.

B. Upper bound on the optimal outage diversity

Conversely, we derive an upper bound to the outage diversity achieved by a system with optimal feedback threshold $\overline{I}(\boldsymbol{k}_\ell)$ with K levels per transmission round. We first assume that $R \in \left(\hat{I}_\tau, \hat{I}_{\tau+1}\right)$ for some $\tau \in \{0, \ldots, BN_t - 1\}$. Consider improving the performance of the system by employing a feedback threshold set $I^\dagger(\boldsymbol{k}_\ell)$ with $\overline{K} = K + \tau + 1$ feedback levels per ARQ round by adding $\tau + 1$ levels to the optimal feedback threshold set $\{\overline{I}(\boldsymbol{k}_\ell)\}$. The extra $\tau + 1$ levels are located at $\hat{I}_t = \frac{tM}{B}, t = 0, \ldots, \tau$.

Let $\mathcal{A}_{\boldsymbol{k}_{\ell-1}}(k) \triangleq \left[I^\dagger([\boldsymbol{k}_{\ell-1}, k]), I^\dagger([\boldsymbol{k}_{\ell-1}, k+1])\right), \ell = 1, \ldots, L, k = 0, \ldots, \overline{K} - 2$, and further let $\overline{\mathcal{A}}_{\boldsymbol{k}_{\ell-1}}(k) \triangleq \left(I^\dagger([\boldsymbol{k}_{\ell-1}, k]), I^\dagger([\boldsymbol{k}_{\ell-1}, k+1])\right)$. Then, given that the feedback vector at round $\ell-1$ is $\boldsymbol{k}_{\ell-1}$, the receiver delivers feedback index $\overline{K}-1$ if $I_{\overline{1,\ell}} \geq I^\dagger\left([\boldsymbol{k}_{\ell-1}, \overline{K} - 1]\right) = R$; otherwise, it delivers index k_ℓ, where k_ℓ is chosen such that $I_{\overline{1,\ell}} \in \mathcal{A}_{\boldsymbol{k}_{\ell-1}}(k_\ell)$.

From the power constraint (10), the optimal power allocation rule is upper-bounded by

$$\overline{P}_\ell(\boldsymbol{k}_{\ell-1}) = \begin{cases} P, & \ell = 1 \\ \frac{P}{\Pr\left\{I_{\overline{1,\ell-1}} \in \mathcal{A}_{\boldsymbol{k}_{\ell-2}}(k_{\ell-1})\right\}}, & k_{\ell-1} < K - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{69}$$

Meanwhile, the power adaptation rule

$$\underline{P}_\ell(\boldsymbol{k}_{\ell-1}) = \begin{cases} \frac{P}{L}, & \ell = 1 \\ \frac{P}{\overline{K} L \Pr\left\{I_{\overline{1,\ell-1}} \in \mathcal{A}_{\boldsymbol{k}_{\ell-2}}(k_{\ell-1})\right\}}, & k_{\ell-1} < K - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{70}$$

satisfies the power constraint in (10). Therefore, the optimal power allocation rule asymptotically satisfies $P_\ell(\boldsymbol{k}_{\ell-1}) \doteq \overline{P}_\ell(\boldsymbol{k}_{\ell-1})$ given in (69).

For $t = 0, \ldots, \tau$, let $\mathcal{S}_{\boldsymbol{k}_{\ell-1}}(t) = \left\{k \in \{1, \ldots, \overline{K} - 2\} : \mathcal{A}_{\boldsymbol{k}_{\ell-1}}(k) \subseteq \left[\hat{I}_t, \hat{I}_{t+1}\right)\right\}$. Since $\hat{I}_t$, for $t = 1, \ldots, \tau$, belongs to the set of thresholds $\{I^\dagger([\boldsymbol{k}_{\ell-1}, k_\ell]), k_\ell = 0, \ldots, \overline{K} - 1\}$,

$\overline{\mathcal{A}}_{\boldsymbol{k}_{\ell-1}}(k) \subseteq \left(\hat{I}_t, \hat{I}_t + 1\right)$ for some $t \in \{1, \ldots, \tau\}$. Applying Theorem 1, for any $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$ and $k \in \mathcal{S}_{\boldsymbol{k}_0}(t)$, we have that

$$\Pr\left\{I_1 < I\right\} \doteq P^{-N_r(BN_t - t)} \doteq P^{-\delta_1(t)} \tag{71}$$

$$\Pr\left\{I_1 \in \mathcal{A}_{\boldsymbol{k}_0}(k)\right\} \doteq \Pr\left\{I < I^{\dagger}\left([k+1]\right)\right\} \doteq P^{-\delta_1(t)}, \tag{72}$$

where $\delta_1(t) = d_1(\hat{I}_{t+1})$ given in (20).

For the induction proof, assume that when $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$ and $k \in \mathcal{S}_{\boldsymbol{k}_{\ell-1}}(t)$, we have

$$\Pr\left\{I_{\overline{1,\ell}} < I\right\} \doteq \Pr\left\{I_{\overline{1,\ell}} \in \mathcal{A}_{\boldsymbol{k}_{\ell-1}}(k)\right\} \doteq P^{-\delta_\ell(t)}, \tag{73}$$

where $\delta_\ell(t) = d_\ell(\hat{I}_{t+1})$ given in (20). The assumption is correct for $\ell = 1$. We prove that (73) is also valid at round $\ell + 1$. In fact, considering $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$, we have

$$\Pr\left\{I_{\overline{1,\ell+1}} < I\right\} = \sum_{j=0}^{t} \Pr\left\{I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_j + I - \hat{I}_t\right)\right\} \Pr\left\{I_{\ell+1} < I - I_{\overline{1,\ell}} \middle| I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_j + I - \hat{I}_t\right)\right\}$$
$$+ \sum_{j=0}^{t} \Pr\left\{I_{\overline{1,\ell}} \in \left[\hat{I}_j + I - \hat{I}_t, \hat{I}_{j+1}\right)\right\} \Pr\left\{I_{\ell+1} < I - I_{\overline{1,\ell}} \middle| I_{\overline{1,\ell}} \in \left[\hat{I}_j + I - \hat{I}_t, \hat{I}_{j+1}\right)\right\}.$$

From assumption (73) and power allocation rule (69), when $I_{\overline{1,\ell}} \in \mathcal{A}_{\boldsymbol{k}_{\ell-1}}(k_\ell)$, the transmit power in round $\ell + 1$ is $P_{\ell+1} \doteq \frac{P}{\Pr\left\{I_{\overline{1,\ell}} \in \mathcal{A}_{\boldsymbol{k}_{\ell-1}}(k_\ell)\right\}} \doteq P^{1+\delta_\ell(j)}$ for all $k_\ell \in \mathcal{S}_{\boldsymbol{k}_\ell}(j)$. Therefore, when $I_{\overline{1,\ell}} \in \left[\hat{I}_j, \hat{I}_{j+1}\right)$, $P_{\ell+1} \doteq P^{1+\delta_\ell(j)}$. Thus, with similar arguments that are used to derive (61), we have that

$$\Pr\left\{I_{\overline{1,\ell+1}} < I\right\} \doteq \sum_{j=0}^{t} P^{-(\delta_\ell(j) + (1+\delta_\ell(j))N_r(BN_t - t + j))} \tag{74}$$

as given in (62). Therefore, following the steps used to derive (67), we have that

$$\Pr\left\{I_{\overline{1,\ell+1}} < I\right\} \doteq P^{-\delta_{\ell+1}(t)} \tag{75}$$

for $I \in \left(\hat{I}_t, \hat{I}_{t+1}\right)$. It follows that

$$\Pr\left\{I_{\overline{1,\ell+1}} \in \mathcal{A}_{\boldsymbol{k}_\ell}(k)\right\} \doteq \Pr\left\{I_{\overline{1,\ell+1}} < I^{\dagger}\left([\boldsymbol{k}_\ell, k]\right)\right\} \doteq P^{-\delta_{\ell+1}(t)} \tag{76}$$

for all $k \in \mathcal{S}_{\boldsymbol{k}_\ell}(t)$. The results in (75) and (76) prove that assumption (73) is valid at round $\ell + 1$, and thus by mathematical induction, (73) is valid for $\ell = 1, \ldots, L$.

Since $R \in \left(\hat{I}_\tau, \hat{I}_{\tau+1}\right)$,

$$\Pr\left\{I_{\overline{1,\ell}} < R\right\} \doteq P^{-\delta_\ell(\tau)} \doteq P^{-d_\ell(\hat{I}_{\tau+1})}, \tag{77}$$

which proves that the outage diversity of the system with $\overline{K}-$level feedback is the same as that given in (20).

REFERENCES

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

[2] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Informatic-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[3] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. IT-19, pp. 357–359, 1973.

[4] E. Malkamäki and H. Leib, "Coded diversity on block-fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 771–781, Mar. 1999.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley and Sons, 2006.

[6] 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.

[7] G. Caire, G. Taricco, and E. Biglieri, "Optimal power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.

[8] G. Caire, D. Tuninetti, and S. Verdú, "Variable-rate coding for slowly fading Gaussian multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2271–2292, Oct. 2004.

[9] S. Nanda, R. Walton, J. Ketchum, M. Wallace, and S. Howard, "A high-performance MIMO OFDM wireless LAN," *IEEE Commun. Mag.*, vol. 43, pp. 101–109, Feb. 2005.

[10] A. Ghosh, D. Wolters, J. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: Current performance benchmarks and future potential," *IEEE Commun. Mag.*, vol. 43, pp. 129–136, Feb. 2005.

[11] D. J. Costello, J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error-control coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2531–2560, Oct. 1998.

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

[13] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–764, Mar. 1998.

[14] H. F. Lu and P. V. Kumar, "Rate-diversity tradeoff of space-time codes with fixed alphabet and optimal constructions for PSK modulation," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2747–2751, Oct. 2003.

[15] ——, "A unified construction of space-time codes with optimal rate-diversity tradeoff," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1709–1730, May 2005.

[16] J. Perret and D. Tuninetti, "Repetition protocols for block fading channels that combine transmission requests and state information," in *Proc. IEEE Int. Conf. on Communications*, Beijing, China, May 2008, pp. 1297–1301.

[17] E. Visotsky, Y. Sun, and V. Tripathi, "Reliability-based incremental redundancy with convolutional codes," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 987–997, Jun. 2005.

[18] Z. Yiqing and W. Jiangzhou, "Optimal subpacket transmission for hybrid ARQ systems," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 934–942, May 2006.

[19] A. Steiner and S. Shamai (Shitz), "Multi-layer broadcasting Hybrid-ARQ strategies for block-fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2640–2650, Jul. 2008.

[20] 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.

[21] A. Chuang, A. Guillén i Fàbregas, L. K. Rasmussen, and I. B. Collings, "Optimal throughput-diversity-delay tradeoff in MIMO ARQ block-fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3968–3986, Sep. 2008.

[22] K. D. Nguyen, L. K. Rasmussen, A. Guillén i Fàbregas, and N. Letzepis, “Diversity-rate-delay tradeoff for ARQ systems over the MIMO block-fading channels,” in *Proc. Aus. Comm. Theory Workshop*, Sydney, Australia, Feb. 2009, pp. 116–121.

[23] H. Liu, L. Razoumov, N. Mandayam, and Spasojević, “An optimal power allocation scheme for the STC Hybrid-ARQ over energy limited networks,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5718–5722, Dec. 2009.

[24] T. Kim, K. Nguyen, and Guillén i Fàbregas, “Coded modulation with mismatched CSIT over MIMO block-fading channels,” *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5631–5640, Dec 2010.

[25] R. Knopp and P. A. Humblet, “On coding for block fading channels,” *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 189–205, Jan. 2000.

[26] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: Theory and algorithms*, 3rd ed. J. Wiley & Sons, 2006.

[27] K. D. Nguyen, A. Guillén i Fàbregas, and L. K. Rasmussen, “A tight lower bound to the outage probability of block-fading channels,” *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4314–4322, Nov. 2007.

[28] 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.

[29] A. J. Viterbi and J. K. Omura, *Principles of Digital Communications*. McGraw-Hill, 1979.

TABLE I
SUMMARY OF NOTATIONS

B	No. of blocks per round	L	Maximum No. rounds
N_t	No. transmit antennas	N_r	No. receive antennas
R	Rate (bits per channel uses)	J	No. channel uses per block
P	Power constraint	P_ℓ	Transmit power
$\mathcal{X}$	input constellation	M	Constellation size
K	No. feedback levels	$k, \boldsymbol{k}$	Feedback index, vector
I	Mutual information	$\hat{I}, \overline{I}$	Quantization thresholds of I
$\boldsymbol{X}, \boldsymbol{x}, x$	Transmit signal	$\boldsymbol{Y}, \boldsymbol{y}, y$	Receive signal
$\boldsymbol{H}$	Channel gain matrix	$\boldsymbol{W}$	AWGN noise
P_e, p	Error, outage probability	d	Outage diversity, SNR exponent

TABLE II
FEEDBACK THRESHOLDS FOR $N_t = N_r = 1, B = 2, L = 2, R = 3.5$.

	$k_2 = 0$	$k_2 = 1$	$k_2 = 2$
$k_1 = 0$	0	2	2.75
$k_1 = 1$	2	2.5	3.0
$k_1 = 2$	2.75	3.0	3.25

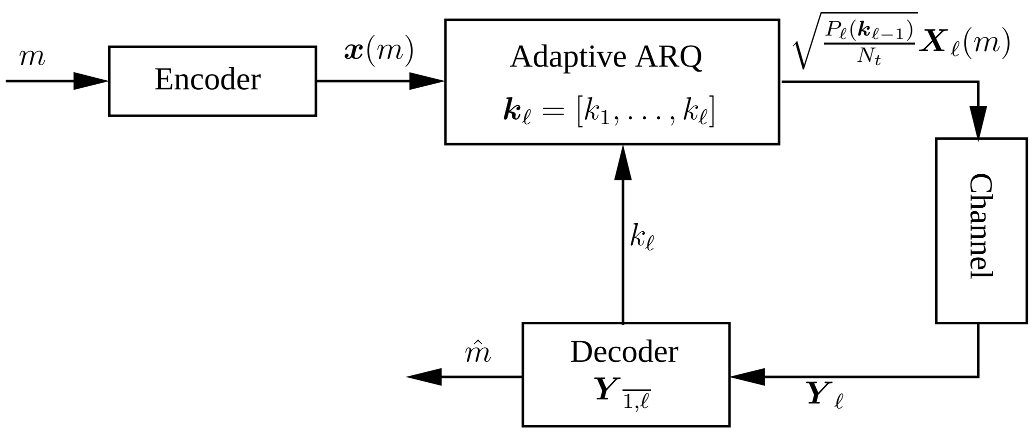


Fig. 1. The INR-ARQ system with multi-bit feedback.

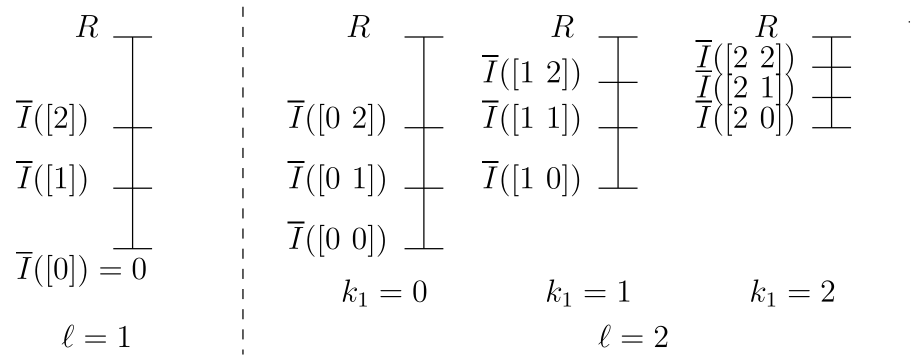
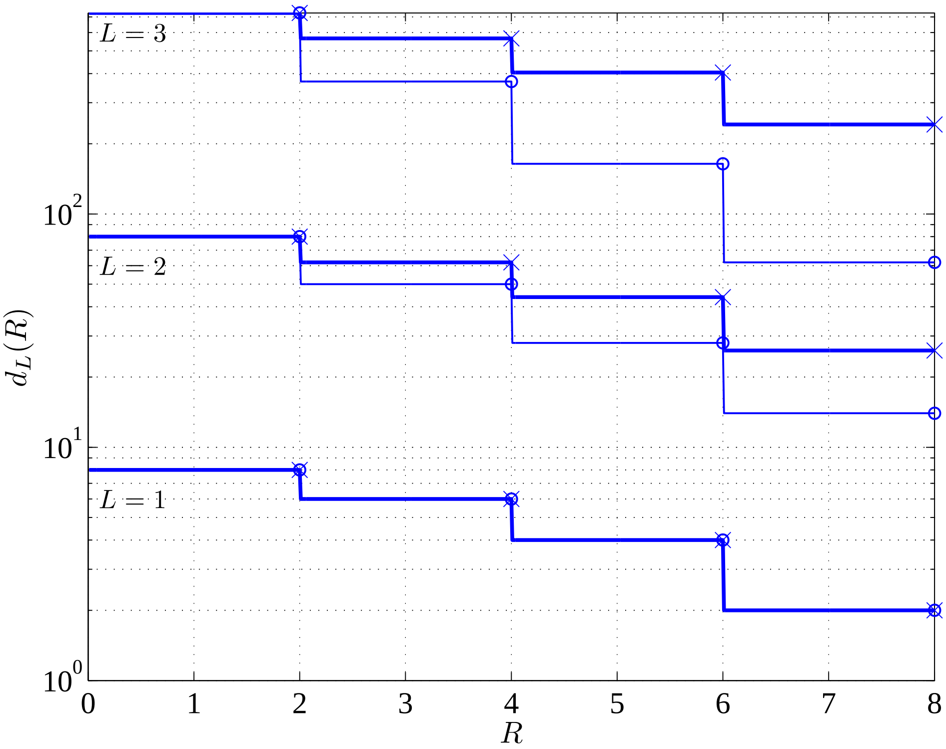
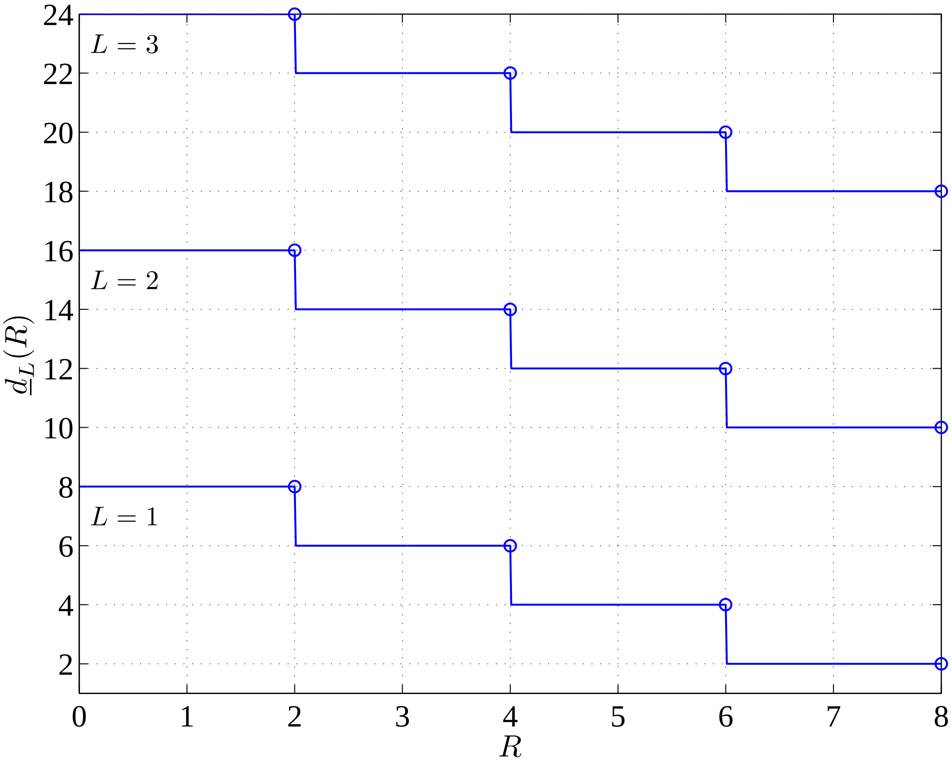


Fig. 2. An example of feedback thresholds.



(a) Long-term power constraint tradeoff.



(b) Constant transmit power tradeoff.

Fig. 3. Optimal rate-diversity-delay tradeoff of ARQ transmission with long-term power constraint (a) and constant power (b). 16-QAM is used over a MIMO block-fading channel with $N_t = N_r = 2, B = 2, L = 1, 2, 3$. Thick and thin lines in (a) represent the optimal tradeoffs $d_L(R)$ achieved by multi-bit feedback ($K \geq \lceil BR/M \rceil + 1$) and $\hat{d}_L(R)$ achieved by one-bit feedback ($K = 2$), respectively. Crosses and circles correspond to the rate points where the SNR-exponent of random codes does not achieve the optimal diversity.

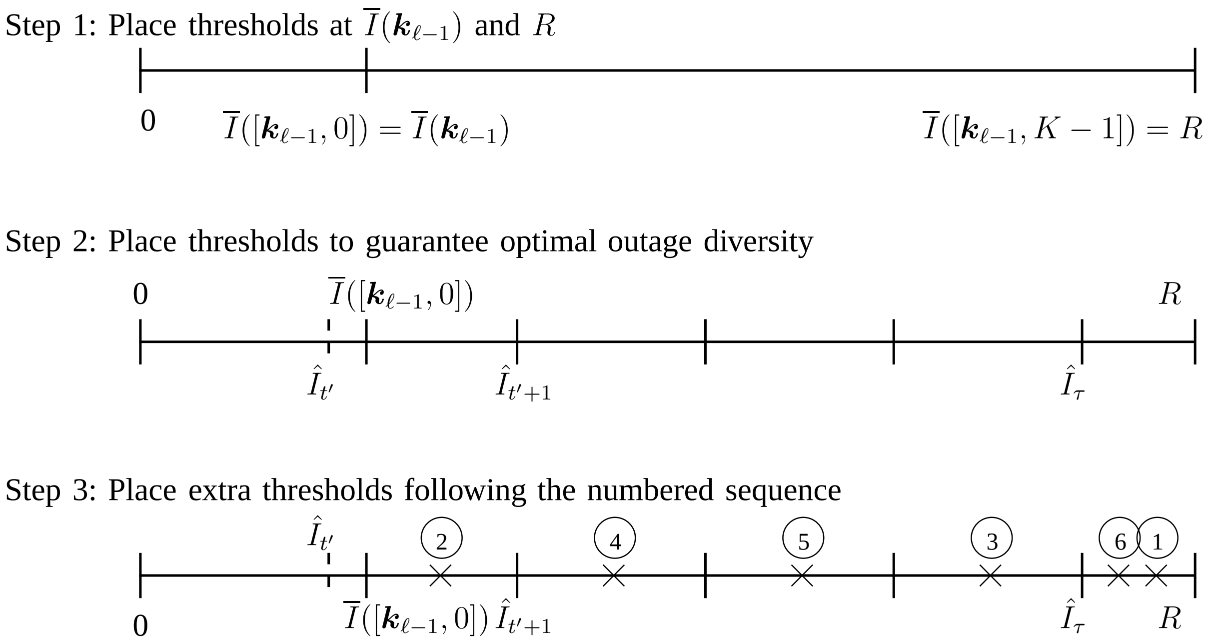


Fig. 4. An example of feedback threshold design ($K = 12$).

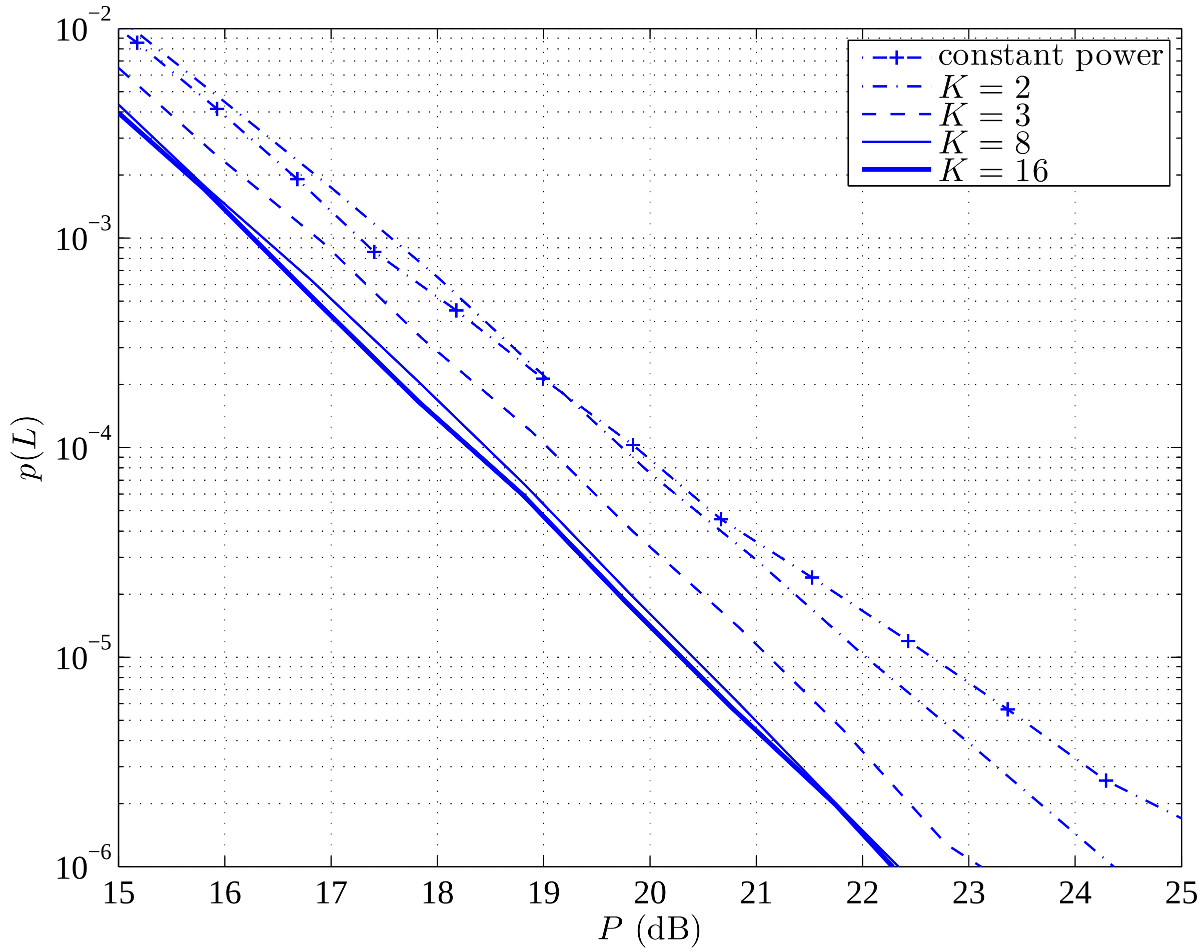


Fig. 5. Outage performance of ARQ transmission schemes for a 16-QAM input block-fading channel with $L = 2, N_t = N_r = 1, B = 2, R = 3.5$.

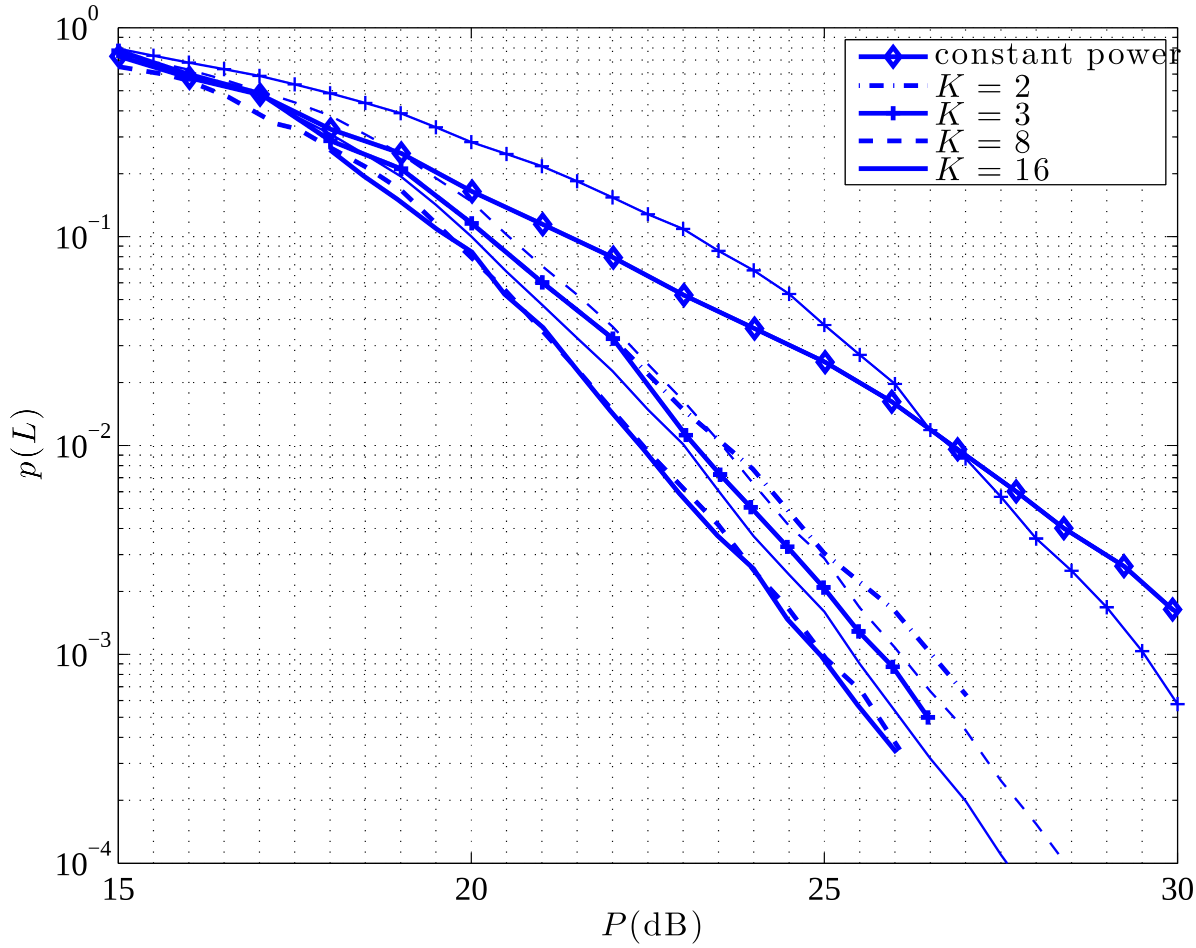


Fig. 6. Outage performance of ARQ transmission using the 16-QAM input constellation over the block-fading channel with $L = 2, N_t = 2, N_r = 1, B = 1, R = 7.5$. Systems with constant transmit power, and systems employing power adaptation with $K = 2, 3, 8, 16$ are considered.

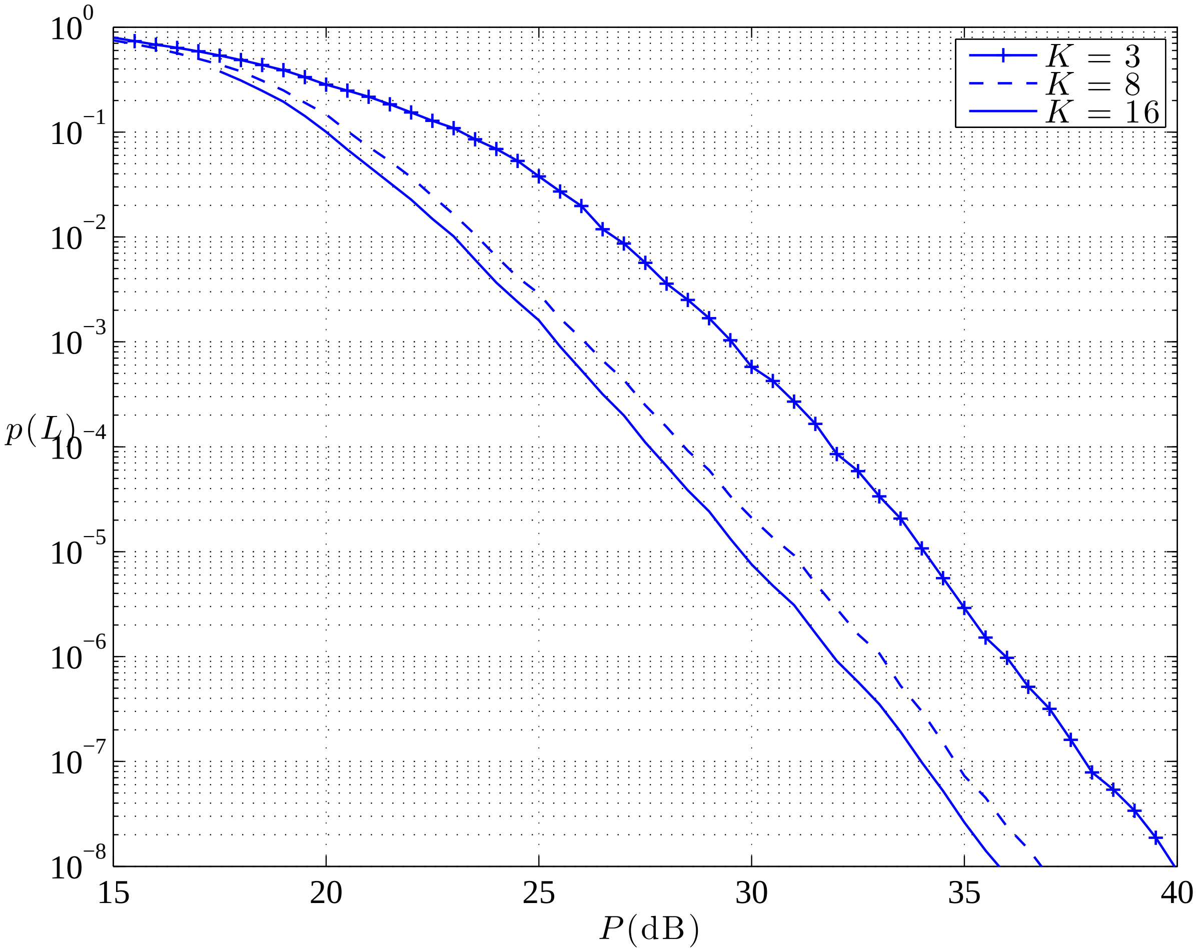


Fig. 7. Upper bound on outage performance of ARQ transmission using 16-QAM input constellations over the block-fading channel with $L = 2, N_t = 2, N_r = 1, B = 1, R = 7.5$ and $K = 3, 8, 16$.