



CHALMERS

Chalmers Publication Library

On the Performance of MIMO-ARQ Systems with Channel State Information at the Receiver

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

IEEE Transactions on Communications (ISSN: 0090-6778)

Citation for the published paper:

Makki, B. ; Eriksson, T. (2014) "On the Performance of MIMO-ARQ Systems with Channel State Information at the Receiver". IEEE Transactions on Communications, vol. 62(5), pp. 1588-1603.

<http://dx.doi.org/10.1109/tcomm.2014.033014.1302>

23

Downloaded from: <http://publications.lib.chalmers.se/publication/200707>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

On the Performance of MIMO-ARQ Systems with Channel State Information at the Receiver

Behrooz Makki, Thomas Eriksson

Abstract—This paper investigates the performance of multiple-input-multiple-output (MIMO) systems in the presence of automatic repeat request (ARQ) feedback. We show that, for a large range of performance metrics, the data transmission efficiency of the ARQ schemes is determined by a set of parameters which are scheme-dependent and not metric-dependent. Then, the results are used to study different aspects of MIMO-ARQ such as the effect of nonlinear power amplifiers, large-scale MIMO-ARQ, adaptive power allocation and different data communication models. The results, which are valid for various forward and feedback channel models, show the efficiency of the MIMO-ARQ techniques in different conditions.

I. INTRODUCTION

Multiple-input-multiple-output (MIMO) transmission [1]–[19], combined with the space-time coding (STC) [20]–[23], is among the best approaches for exploiting the spatial diversity, particularly over rich scattered environments. MIMO has revolutionized the modern wireless communications, is a key part of most current standards such as WiFi (IEEE 802.11) and WiMax (IEEE 802.16) [24], [25], and is expected to be the core technology for the next generation broadband wireless communication systems. The performance of the MIMO channels, however, depends strongly on the amount of channel state information (CSI) at the transmission endpoints. Channel estimation at the receiver is relatively simple and incurs negligible loss in the transmission rate, particularly when the channel experiences slow variations. On the other hand, due to the signaling load caused by reporting the channel information, assuming perfect channel knowledge at the transmitter is an overly optimistic assumption. Therefore, it is important to study the MIMO channels in the presence of limited-feedback schemes.

Hybrid automatic repeat request (ARQ) is a well-established approach for wireless networks [1]–[19], [26]–[33]. From an information-theoretic point of view, the ARQ systems can be viewed as channels with *sequential* feedback where, utilizing both forward error correction and error detection, the system performance is improved by retransmitting the data which has experienced *bad* channel conditions. Therefore, the combination of MIMO and ARQ improves the performance of wireless systems. This is the main motivation for the current MIMO-ARQ techniques such as [1]–[19] and this paper as well.

The performance of ARQ schemes in single-input-single-output (SISO) setups has been studied by, e.g., [27]–[34]. The MIMO-ARQ related works can be divided into three

(partly overlapping) categories. The first group are papers focusing on precoder design [35]–[39]. Here, considering different modulations, the optimal linear precoders minimizing the distortion or the bit error rate have been obtained in the presence of (im)perfect CSI at the transmitter (CSIT).

The second group are the papers that, following the outstanding work by Zheng and Tse [40], have studied the diversity-multiplexing-tradeoff (DMT) or diversity-multiplexing-delay-tradeoff (DMDT) of the MIMO-ARQ setups, e.g., [1]–[6]. These metrics establish the necessary trade-off between reliability and throughput in outage-limited fading channels. However, DMDT and DMT are metrics mostly for the high signal-to-noise ratio (SNR) regime and do not provide a complete picture for evaluating the performance of ARQ schemes operating at finite transmission rates and powers.

The last category of the MIMO-ARQ papers are the ones that have studied the theoretical and practical aspects of the channel at finite SNR. Here, the effect of non-orthogonal space-time block codes (STBCs) [7], Bell-Labs layered space-time (BLAST) systems [8], bit-interleaved coded modulation [9], quadrature amplitude modulation (QAM) [10]–[12] or using adaptive modulation [13] on the throughput [8], the bit error rate [9]–[11], the packet error rate [7], [12] or the spectral efficiency [13] of the MIMO-ARQ channels have been investigated. Studying the high but finite SNR DMDT of multihop MIMO-ARQ relay networks, low-complexity ARQ techniques in a two-user MIMO broadcast channel, the average rate of MIMO-ARQ schemes utilizing linear dispersion codes (LDCs) and adaptive coding by different STCs have been considered by [14]–[17], respectively. Finally, [18] developed a parallel ARQ model in correlated MIMO channels and [19] studied the outage minimization problem in MIMO channels when at each ARQ round the data is sent by only a single transmit antenna.

Reviewing [1]–[19], [26]–[33], it is clear that there is large diversity in selection of the performance metrics and fading models. This has led to repeating conceptually similar analysis for each performance yardstick individually. Also, a large effort has been undertaken to compare the ARQ schemes from different points of view and for various fading models. Thus, a natural question is whether it is necessary to study the ARQ schemes for each fading model/metric individually. The answer this paper establishes is negative.

Here, we show that, for a large range of performance metrics, the data transmission efficiency of the ARQ schemes is determined by a set of intermediate parameters which are scheme-dependent and not metric-dependent. This point, which is illustrated in Fig. 1, indicates that there is a specific set of parameters, referred to as accumulated mutual informa-

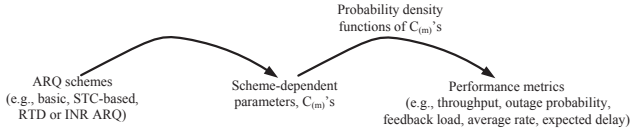


Figure 1. Illustration of the scheme-dependent parameters in ARQ approaches.

tion (AMI), from which the data transmission efficiency of ARQ schemes can be analyzed. Then, we use this result to study different aspects of the MIMO-ARQ techniques such as 1) the effect of nonlinear power amplifiers, 2) large-scale MIMO-ARQ, 3) temporal and spatial power allocation, 4) unreliable feedback channel, 5) fading channel variability and 6) different data transmission models. Also, as side-results, we present comparisons between different MIMO-ARQ schemes and show mappings between the MIMO- and SISO-ARQ setups. The mapping simplifies the performance analysis of MIMO-ARQ and makes it possible to extend many theoretical results of SISO-ARQ setups to MIMO-ARQ. Finally, we obtain the empirical probability density function (pdf) of the received SNR for different data communication models and investigate the effect of variable-length coding and different STCs on the data transmission of the MIMO-ARQ systems.

It is worth noting that the idea of utilizing intermediate parameters for performance analysis of ARQ schemes have been implicitly used by, e.g., [1] and [27], [29] where the DMDT and the throughput have been studied, respectively. Thus, this paper is the generalization of [1], [27], [29] to a wider range of performance metrics and for different forward and feedback channel conditions. Also, compared to the literature, we present new analysis on different aspects of MIMO-ARQ transmission which, to the best of authors' knowledge, have not been studied before. The results show the efficiency of the MIMO-ARQ techniques in different conditions.

II. SYSTEM MODEL

In this section, we present the definitions and the system model on which the analytical/numerical analysis is based. The following notation is used throughout the paper:

- A packet is defined as the transmission of a sub-codeword along with all its possible retransmission rounds. A maximum of M ARQ retransmission rounds is considered, i.e., the data is (re)transmitted a maximum of $M + 1$ rounds.
- l_m (in channel uses) is the length of the sub-codeword (re)transmitted in the m -th (re)transmission round and $l_{(m)} = \sum_{n=1}^m l_n$ represents the total channel uses up to the end of the m -th (re)transmission round.
- $R_{(m)}$ (in nats-per-channel-use (npcu)) denotes the *equivalent* data rate at the end of the m -th round, i.e., the ratio of the transmitted information nats and the total channel uses up to the end of round m . Thus, denoting the number of information nats considered for a packet by Q , we have $R_{(m)} = \frac{Q}{l_{(m)}}$. Also, to simplify the equations, we define $R_{(0)} \doteq \infty$ and $R_{(M+2)} \doteq 0$.
- We define $C_{(m)}$ as the AMI of the MIMO-ARQ model at the end of the m -th round. That is, $C_{(m)}$ is the maximum

equivalent rate that can be decoded at the m -th round of the considered ARQ scheme. As seen in the sequel, the AMIs are determined based on the encoding/decoding procedure of the ARQ schemes and are random variables depending on the channel realization(s).

- A_m represents the event that the codeword is decodable at the end of the m -th round. Also, $A_0 = \emptyset$ denotes the empty set.
- The matrices are represented by bold capital letters. Finally, $\text{Tr}(\mathbf{X})$, $|\mathbf{X}|$, \mathbf{X}^T and \mathbf{X}^H represent the trace, the determinant, the transpose and the Hermitian of the matrix \mathbf{X} , respectively.

General system model. We study the system performance for three different fading conditions:

- **Fast-fading.** Here, it is assumed that a finite number of channel realizations are experienced within each ARQ retransmission round.
- **Slow-fading.** In this model, the channel is supposed to change between two successive retransmission rounds, while it is fixed for the duration of each sub-codeword.
- **Quasi-static.** The channel is assumed to remain fixed within a packet period.

Fast-fading is an appropriate assumption for fast-moving users or users with long codewords compared to the channel coherence time [41]. On the other hand, slow-fading is a good model for the users with moderate speeds [3]–[6], [36]–[39]. Finally, the quasi-static case can properly model the channel characteristics in slow-moving or stationary users, e.g., [1], [14], [16], [17], [19], [28], [30]–[32], [39].

In Sections III–VI, we consider the quasi-static fading condition, unless otherwise stated. As stated in [1], the quasi-static model makes it possible to decouple the ARQ gain from the temporal interleaving gain. That is, from the diversity point of view, the quasi-static condition corresponds to the worst case scenario since no time diversity can be exploited by the ARQ. Also, we note that the quasi-static model is justified in practice by considering, e.g., time-division multiple access (TDMA) systems where the channel is allocated to a transmitter-receiver pair sporadically and (at least) one packet is sent in each slot [1]; if the channel coherence time is larger than the maximum length of the packet, but it is smaller than the idle time between two consecutive active times, the quasi-static model accurately describes the channel. Also, compared to the slow- and fast-fading models, the quasi-static assumption matches better with the assumption of perfect CSI at the receiver, which is considered in the paper. Later, in Section VII, we show that 1) many conclusions of the paper are valid independent of the fading model and 2) there are cases where the quasi-static, slow- and fast-fading conditions can be mapped to each other when studying the performance of ARQ schemes. Finally, the simulation results are presented for independent and identically distributed (iid) fading conditions. However, as illustrated in Section IV, the qualitative discussions also hold for the temporally-correlated fading models, unless otherwise stated.

For different fading conditions, the channel coefficients are assumed to be known at the receiver, in harmony with [1]–

[6], [30]–[33], [41]. The ARQ feedback bits are sequential signals increasing the CSI at the transmitter; using ARQ, the transmitter considers some initial transmission rate and power with no pre-knowledge about the channel quality. Then, with the help of ARQ, the transmitter CSI is refined based on the message decoding status, and the transmission rates/powers are adapted. Finally, although not considered in the paper, the same procedure as in [27], [42] (with optimistic assumptions on the feedback delay) can be used for sending quantized CSI at the beginning of each ARQ packet, and all theory developed here is still valid since the channel with quantized CSIT can be modeled as a channel with no CSIT and a modified fading pdf, e.g., [30]. Moreover, with a slow-fading condition, [34] studies the combination of ARQ and delayed CSI feedback schemes where, as it also deals with the maximum achievable rates in different conditions, our analysis can be useful in that case as well.

The ARQ feedback bits are initially supposed to be received delay- and error-free, while we later investigate the system performance for erroneous feedback signals as well (Section VII). Also, we initially consider the *bursty* communication model [28], [31], [32] where there is an idle period between two successive packet transmissions. However, in Section IV the results are compared with the ones achieved with continuous communication, where the data is sent continuously.

MIMO-ARQ model. The MIMO-ARQ setup is represented as follows. Let $\mathbf{H} \in \mathcal{C}^{L_r \times L_t}$ denote the channel matrix where L_t and L_r are the number of transmit and receive antennas, respectively. Selecting a message w from a set of uniformly distributed messages \mathcal{W} , the encoder of a MIMO-ARQ scheme produces a *mother* codeword $\mathbf{X}(w) \in \mathcal{C}^{L_t \times l_{(M+1)}}$. Then, the codeword is divided into $M + 1$ sub-codewords $\mathbf{X}_m \in \mathcal{C}^{L_t \times l_m}$, with length of l_m channel uses, such that $\mathbf{X}(w) = [\mathbf{X}_1 \dots \mathbf{X}_{M+1}]$. At the m -th (re)transmission round, the sub-codeword \mathbf{X}_m is sent by the transmitter. Thus, defining $\mathbf{X}_{(m)} = [\mathbf{X}_1 \dots \mathbf{X}_m]$ as the total signal transmitted up to the end of the m -th round, we have

$$\mathbf{Y}_{(m)} = \mathbf{H}\mathbf{X}_{(m)} + \mathbf{Z}_{(m)}, \mathbf{Z}_{(m)} \in \mathcal{C}^{L_r \times l_{(m)}}, \quad (1)$$

where $\mathbf{Y}_{(m)} = [\mathbf{Y}_1 \dots \mathbf{Y}_m]$ is the total signal received up to the end of the m -th round and $\mathbf{Z}_{(m)}$ denotes the iid complex Gaussian noise matrix whose elements follow $\mathcal{CN}(0, 1)$.

The ARQ schemes are normally divided into three categories, Types I–III [43]. Type I is the simplest version of hybrid ARQ, where both the error-detecting and the forward error correction information are added to each message and the receiver disregards the previous messages, if received in error. In Types II and III, however, the successive retransmissions are combined to improve the system performance. The difference between Types II and III protocols is in what is retransmitted in each retransmission. In this work, we consider the following examples of hybrid ARQ:

- **Incremental redundancy (INR) protocols.** Here, new sub-codewords are transmitted in the (re)transmissions and in each round the receiver combines all signals received up to the end of that round (Type II hybrid ARQ).

- **Repetition time diversity (RTD) protocols.** Using the RTD, the same data is repeated in the (re)transmission rounds and, in each round, the receiver performs maximum ratio combining (MRC) of all received signals (Type III hybrid ARQ).
- **STC-based protocols.** With STC, a permuted version of the initial sub-codeword is sent in the retransmissions. That is, the original STC, which is produced by combining the permutations of an initial sub-codeword, is divided into a number of sub-codewords. The sub-codewords are then sent in the retransmissions, until the message is successfully decoded or the maximum number of retransmissions is reached. Here, we study the system performance for the Alamouti, the cyclic delay diversity (CDD), the spatial multiplexing with repetition (SMR) or the antenna switching (AS) kinds of STCs.

In the literature, it is common to use an isotropic Gaussian input distribution over all transmit antennas, i.e.,

$$\mathbf{K}_m = \frac{\phi}{L_t} \mathbf{I}_{L_t}, \forall m, \quad (2)$$

where \mathbf{K}_m denotes the input covariance matrix in the m -th round, ϕ is the total transmission power, $\frac{\phi}{L_t}$ is the transmission power per transmit antenna and \mathbf{I}_{L_t} represents the $L_t \times L_t$ identity matrix. This is because the problem of finding the optimal covariance matrix that maximizes the capacity with no CSIT is still open. In all simulations, except Figs. 8 and 9, we use (2) as the input covariance matrix. Section VI, however, investigates the effect of power allocation on the system performance, and the corresponding simulation results are presented in Figs. 8 and 9.

III. PERFORMANCE ANALYSIS AS A FUNCTION OF AMIs

This section is devoted to clarifying the relationship in Fig. 1. The final conclusion of the section is that to analyze many performance metrics in ARQ protocols it is enough to study their AMIs, which are scheme-dependent and not metric-dependent. Along with simplifying the mathematical analysis, this point leads to a number of outcomes such as:

- It allows us to compare the ARQ schemes based on their AMIs and not for a specific metric.
- As the system performance depends only on the AMIs, many closed-form/approximate expressions previously derived for MIMO and SISO setups can be applied in the MIMO-ARQ channels directly. As some examples, we study the effect of nonlinear power amplifiers and large-scale MIMO on the performance of MIMO-ARQ protocols (For more details, see Examples 1 and 2).
- The performance of ARQ protocols with different temporal fading variations, such as quasi-static, slow- and fast-fading, can be mapped to each other (See Section VII).
- Finally, we can elaborate on the properties of an optimal ARQ protocol which outperforms every other ARQ scheme, for all studied performance metrics. Then, as a side-result, we show that there is no ARQ protocol that can reach the channel ergodic capacity with bursty data

communication in quasi-static channels (See Section IV for more details).

A. On the definition of performance metrics

Here, we study a number of performance metrics that are commonly used for evaluating ARQ protocols and show that they can be expressed as functions of the AMIs. Note that the considered metrics are only examples for clarifying the main conclusion, and we concentrate on no specific metric throughout the paper. Also, the metrics are obtained independently of the ARQ protocol and the fading channel.

Let us first find the long-term (LT) throughput (in npcu) which, based on the renewal-reward theorem, is defined as [27]–[32], [44]

$$\eta = \frac{E\{\tilde{Q}\}}{E\{\tilde{L}\}}. \quad (3)$$

Here, \tilde{Q} and \tilde{L} represent the number of successfully decoded information nats and the number of channel uses during a packet period, respectively, and $E\{\cdot\}$ denotes the expectation operator. The LT throughput is found as follows.

Outage occurs if the data is decodable in none of the (re)transmission rounds. Thus, according to the definition of \bar{A}_m (see Section II), we have $\Pr(\text{Outage}) = \Pr(\bar{A}_1, \dots, \bar{A}_{M+1})$. If the data is decoded at any (re)transmission round, all information nats are received by the receiver. Hence, the expected number of received information nats in each packet is

$$E\{\tilde{Q}\} = Q(1 - \Pr(\text{Outage})) = Q(1 - \Pr(\bar{A}_1, \dots, \bar{A}_{M+1})). \quad (4)$$

If the data is decoded at the end of the m -th round, the total number of channel uses is $l_{(m)}$. Also, the total number of channel uses would be $l_{(M+1)}$ if an outage occurs, where all possible retransmission rounds are used. Therefore, the expected number of channel uses is obtained by

$$E\{\tilde{L}\} = \sum_{m=1}^{M+1} l_{(m)} \Pr(\bar{A}_0, \dots, \bar{A}_{m-1}, A_m) + l_{(M+1)} \Pr(\bar{A}_0, \dots, \bar{A}_{M+1}). \quad (5)$$

Here, $\Pr(\bar{A}_1, \dots, \bar{A}_{m-1}, A_m)$ is the probability that the data is decoded at the end of the m -th (re)transmission round while it was not decoded before.

The data is decoded at the end of the m -th round (and not before) iff 1) all previous equivalent data rates are higher than the maximum decodable rates, i.e., $C_{(n)} < R_{(n)}, \forall n < m$, and 2) using the m -th retransmission round, the data can be decoded, i.e., $C_{(m)} \geq R_{(m)}$. Thus, with some manipulations on the probabilities, we have $\Pr(\bar{A}_0, \dots, \bar{A}_m) = \Pr(C_{(m)} < R_{(m)})$ and (5) can be rewritten as

$$E\{\tilde{L}\} = \sum_{m=1}^{M+1} l_m \Pr(\bar{A}_0, \dots, \bar{A}_{m-1}) + \sum_{m=1}^{M+1} l_m \Pr(C_{(m-1)} < R_{(m-1)}). \quad (6)$$

Intuitively, (6) is based on the fact that the m -th retransmission round, with length l_m , is used if the data has not been decoded before. In this way, using $l_m = \frac{Q}{R_{(m)}} - \frac{Q}{R_{(m-1)}}$, (3)-(6) and $\Pr(\bar{A}_1, \dots, \bar{A}_{M+1}) = \Pr(C_{(M+1)} < R_{(M+1)})$, the LT throughput and the outage probability are rephrased as

$$\eta = \frac{1 - \Pr(C_{(M+1)} < R_{(M+1)})}{\sum_{m=1}^{M+1} (\frac{1}{R_{(m)}} - \frac{1}{R_{(m-1)}}) \Pr(C_{(m-1)} < R_{(m-1)})} \quad (7)$$

and

$$\Pr(\text{Outage}) = \Pr(C_{(M+1)} < R_{(M+1)}), \quad (8)$$

respectively. Also, for ARQ protocols with fixed-length coding we have $l_m = l, \forall m$, which leads to $R_{(m)} = \frac{Q}{ml} = \frac{R}{m}$, $\Pr(\text{Outage}) = \Pr(C_{(M+1)} < \frac{R}{M+1})$ and

$$\eta = R \frac{1 - \Pr(C_{(M+1)} < \frac{R}{M+1})}{1 + \sum_{m=1}^M \Pr(C_{(m)} < \frac{R}{m})}, \quad (9)$$

where $R = \frac{Q}{l}$ is the initial sub-codeword rate and l is the length of the sub-codewords.

With the same procedure as in (6), the expected number of ARQ-based retransmission rounds, which is of interest in terms of complexity [30], is found as $s = 1 + \sum_{m=1}^M \Pr(C_{(m)} < R_{(m)})$. Moreover, the feedback load defined as the expected number of feedback bits transmitted in a packet period is obtained by

$$b = 1 + \sum_{m=1}^{M-1} \Pr(C_{(m)} < R_{(m)}) \text{ (in bits per packet)} \quad (10)$$

which is based on (6) and the fact that no ARQ feedback is sent at the end of the packet. Finally, as another metric of interest, the average rate [16], [31] of an ARQ approach is obtained by

$$\begin{aligned} r &= \sum_{m=1}^{M+1} R_{(m)} \Pr(\bar{A}_1, \dots, \bar{A}_{m-1}, A_m) \\ &= \sum_{m=1}^{M+1} R_{(m)} \Pr(C_{(m-1)} < R_{(m-1)}, C_{(m)} \geq R_{(m)}) \\ &= \sum_{m=1}^{M+1} (R_{(m)} - R_{(m+1)}) \Pr(C_{(m)} \geq R_{(m)}) \text{ (in npcu)}, \end{aligned} \quad (11)$$

which follows from the fact that the achievable rate is $R_{(m)}$ if the data is decoded at the m -th round (and not before). Also, the last equality in (11) is obtained with the same procedure as in (6).

Note that in (6)-(11) we have considered no specific ARQ protocol. For the performance analysis of different MIMO-ARQ protocols, the readers are referred to Section V.

B. On the effect of AMIs

Let us start the discussions on the AMIs with the following lemma which has been implicitly used by, e.g., [1], [27], [29].

Lemma 1: If a performance metric of an ARQ protocol is a monotonically increasing function of the probabilities

$\Pr(C_{(m)} < R_{(m)}), \forall m$, the dominance of the cumulative distribution functions (cdf)¹ of the AMIs is a sufficient (but not necessary) condition for increasing the metric.

Proof. Let $C_{(m)}$ and $C'_{(m)}$, $m = 1, \dots, M+1$, be two sets of AMIs and define $\theta(\Pr(C_{(m)} < x), \forall m)$ as a metric which is a monotonically increasing function of the probabilities $\Pr(C_{(m)} \leq x), \forall x, m$. With $F_{C_{(m)}}(x) \leq F_{C'_{(m)}}(x), \forall x, m$, we have $\Pr(C_{(m)} \leq x) \leq \Pr(C'_{(m)} \leq x)$ and, as stated in the lemma, $\theta(\Pr(C_{(m)} < x), \forall m) \leq \theta(\Pr(C'_{(m)} < x), \forall m)$. \square

Considering (6)-(11), it is found that many performance metrics of ARQ protocols are monotonic functions of the probabilities $\Pr(C_{(m)} < R_{(m)}), \forall m$. This is because the system performance² depends on the retransmission round in which the codewords are correctly decoded. Moreover, the interesting point is that the probability $\Pr(C_{(m)} < x)$ is directly linked to the AMI $C_{(m)}$ which is a random variable and function of the fading realizations experienced in rounds $n = 1, \dots, m$. That is, as soon as the AMIs $C_{(m)}, m = 1, \dots, M+1$, are determined as functions of channel realization(s), their corresponding pdfs $f_{C_{(m)}}, m = 1, \dots, M+1$, are found, from which the probabilities $\Pr(C_{(m)} < R_{(m)})$ and, consequently, the considered performance metrics are obtained. Therefore, instead of concentrating on different performance metrics, we can study the data transmission efficiency of the ARQ protocols via analyzing the AMIs. In other words, as illustrated in Fig. 1, the performance of the ARQ protocols can be determined by a set of intermediate parameters which are scheme-dependent and not metric-dependent.

Note that the relations of the AMIs to the fading channel are determined based on the encoding/decoding procedure of the considered ARQ protocol. Also, the AMIs are not known by the transmitter, because there is no CSI at the transmitter except the ARQ feedback bits. Moreover, since we are always dealing with the probabilities $\Pr(C_{(m)} < R_{(m)})$, the distributions of $C_{(m)}$'s are of interest and not their instantaneous realizations. However, determining the AMIs as functions of the channel realization(s) we can find their corresponding pdfs and the probabilities $\Pr(C_{(m)} < R_{(m)})$. This conclusion which is in harmony with intuition, indicates that 1) the only difference in performance between the different ARQ protocols is in their AMIs and 2) for every given set of data rates, the system performance is independent of the selected ARQ protocol if the distributions of AMIs are given. Finally note that 1) the conclusion is derived for the set of metrics (6)-(11) and it is not proved for every possible utility function. However, the arguments on the effect of the AMIs hold for every other metric that we have checked, e.g., the expected transmission energy per packet [46, eq. 4], the transmission throughput [1, eq. 11] and the diversity gain [1, eq. 14]. Also, 2) up to now, the results are general in the sense that they are independent of the ARQ protocol and the fading model (For more details see Sections V and VII where different ARQ

schemes and temporal variations of the fading channel are discussed). In the following, we provide various examples that demonstrate the importance of the arguments.

Remark 1: Let $C_{(m)}^p, \forall m$, and θ^p be, respectively, the AMIs and a utility function, e.g., (6)-(11), for the ARQ protocol "p". Also, define $\theta^q \prec \theta^p$ as *scheme p outperforms the scheme q, in terms of metric θ* . Then, $\theta^q \prec \theta^p$ if $C_{(m)}^q \leq C_{(m)}^p, \forall m, \mathbf{H}$, which is because $\Pr(C_{(m)}^p < R_{(m)}) \leq \Pr(C_{(m)}^q < R_{(m)})$ if $C_{(m)}^q \leq C_{(m)}^p, \forall m, \mathbf{H}$, i.e., with the same transmission parameters there is higher chance for scheme p, compared to scheme q, to decode the message in the earlier rounds. Thus, to compare two ARQ protocols we can verify if the AMIs of one of them exceeds the AMIs of the other one for every possible channel realization (sufficient condition), instead of deriving the cdfs; this is usually much simpler.

Examples of sorting the ARQ schemes based on their AMIs are given in Section V. Note that 1) the remark introduces a sufficient (but not necessary) condition for comparing ARQ protocols and 2) it is useful when all AMIs of a protocol exceed the corresponding AMIs of the other scheme for every channel realization(s), which is the case for all ARQ comparisons that we tested (see Section V). Also, 3) the remark is of interest as it frees us from the necessity of calculating the pdfs $f_{C_{(m)}}, \forall m$.

Definition 1: (Optimal ARQ protocol) As we know, the decodable rate of any communication channel is limited to the maximum achievable mutual information accumulated during data transmission. Thus, for a given set of (re)transmission rates and powers, we have $C_{(m)} \leq C_{(m),\max}, \forall m, \mathbf{H}$, where $C_{(m),\max}$ is the maximum achievable AMI up to the end of round m , maximized over all possible encoding/decoding procedures and input pdfs. Therefore, following Remark 1, we refer to an ARQ protocol as *optimal*, if $C_{(m)} = C_{(m),\max}, \forall m, \mathbf{H}$. For instance, with an input covariance matrix \mathbf{K} , the maximum achievable mutual information of a MIMO setup experiencing fading realization \mathbf{H} is $\log |\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h|$, where \mathbf{I}_{L_r} is the $L_r \times L_r$ identity matrix [47]. Thus, a MIMO-ARQ protocol is optimal if $C_{(m)} = \log |\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h|, \forall m, \mathbf{H}$. As seen in the sequel, the INR and some STC-based MIMO-ARQ protocols are optimal. Also, examples of non-optimal ARQ schemes are seen in Section V. Finally, note that in the definition the optimality condition is on the $C_{(m)}$'s via proper selection of the encoding/decoding procedure and the input pdf. Meanwhile, as seen in the following, we can still improve the system performance by optimizing the transmission rates/powers, in terms of the considered metric.

Example 1: (Nonlinear power amplifiers in MIMO-ARQ protocols) In [48], the effect of nonlinear power amplifiers on the performance of the open-loop systems is studied where, for a given channel realization \mathbf{H} , the maximum decodable rate of a multiple-input-single-output (MISO) channel is upper bounded by $\log \left(1 + \|\mathbf{H}\|_F^2 \zeta^2 \left(\sqrt{\frac{\phi}{L_r}} \right) \right)$. Here, $\|\cdot\|_F$ is the Frobenius norm operator and $\zeta(\cdot)$ is a function representing the amplitude-to-amplitude mapping of a nonlinear power amplifier.

Now, according to the above discussions, the results of [48] can be applied to the MIMO-ARQ schemes. Specifically,

¹The pdf and the cdf of a random variable Λ are represented by $f_\Lambda(\cdot)$ and $F_\Lambda(\cdot)$, respectively. The random variable Δ dominates the random variable Λ if $F_\Lambda(x) \leq F_\Delta(x), \forall x$ [45].

²Throughout the paper, by the performance we mean any of the metrics obtained in (6)-(11), unless otherwise specified.

we can use the same procedure as in [48, Section III.D] to show that the AMIs of a MISO setup using an optimal ARQ protocol, such as the INR, are bounded by

$$C_{(m)} \leq \log \left(1 + \|\mathbf{H}\|_F^2 \zeta^2 \left(\sqrt{\frac{\phi}{L_t}} \right) \right), \forall m. \quad (12)$$

In this way, from (9), the LT throughput of the fixed-length MISO-INR is bounded by

$$\eta \leq R \frac{1 - F_U \left(\frac{e^{\frac{R}{M+1}} - 1}{\zeta^2 \left(\sqrt{\frac{\phi}{L_t}} \right)} \right)}{1 + \sum_{m=1}^M F_U \left(\frac{e^{\frac{R}{m}} - 1}{\zeta^2 \left(\sqrt{\frac{\phi}{L_t}} \right)} \right)} \quad (13)$$

where F_U is the cdf of the random variable $U = \|\mathbf{H}\|_F^2$. Also, (13) is based on (9) and the fact that the LT throughput is a decreasing function of $\Pr(C_m < R_{(m)})$, $\forall m$. Hence, an upper bound (resp. a lower bound) of the LT throughput is obtained if $C_{(m)}$ is replaced by its upper bound (resp. lower bound). Finally, note that with a linear power amplifier $\zeta(x) = x$ we have $C_{(m)} = \log \left(1 + \frac{\phi}{L_t} \|\mathbf{H}\|_F^2 \right)$, $\forall m$, as in Definition 1.

Considering a solid-state power amplifier (SSPA) with $\zeta(x) = \frac{x}{(1 + (\frac{x}{A_{os}})^{2\tau})^{\frac{1}{2\tau}}}$ [48] and typical constants $A_{os} = 5$ and $\tau = 1$, Fig. 2 demonstrates the effect of nonlinear power amplifiers on the LT throughput of 4×1 MISO-INR setup with Rayleigh-fading channels³. Here, the SNR is defined as $10 \log_{10}(\phi)$ in dB. Note that, according to [48, Fig. 6] and the considered parameter setting, the bound is very tight for the AMIs and, consequently, the LT throughput. As seen in the figure, the performance of the MIMO-ARQ protocols is substantially affected by the efficiency of the amplifiers when the input SNR increases. However, the numerical results of the example should not be overemphasized because we have considered no compensation for the nonlinear power amplifier (of course, we can apply the same mathematical procedure for the cases with compensation of the nonlinear power amplifier). Finally, as a side-result, the figure indicates that, compared to the open-loop communication setup, considerable throughput increment is achieved by a limited number of ARQ-based retransmissions.

Example 2: (Large-scale MIMO-ARQ) Recently, the implementation of very large antenna arrays at the transmitters and/or receivers, referred to as large-scale MIMO, has become a hot research topic, because it offers large degrees of freedom, the possibility for simultaneous transmission to several users, etc. [49], [50]. One of the main bottlenecks of the large-scale MIMO is the CSI acquisition, specially at the transmitter, which has limited the large-scale MIMO to an academic research topic [49]. This is the place where ARQ is of great value, leading to substantial feedback load reduction. Thus, it is expected to see many works dealing with large-scale MIMO-ARQ in the future. This example is devoted to provide some hints for studying the large-scale MIMO-ARQ. Specifically, it is seen that the properties of the AMIs let us map a large-

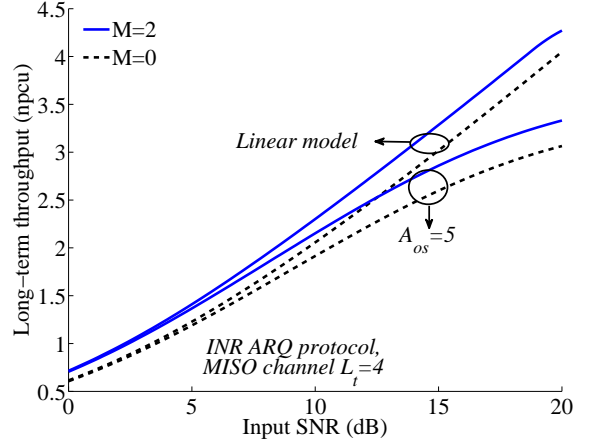


Figure 2. LT throughput in different power amplifier models, MISO channel $L_t = 4$, optimal (INR) ARQ protocol, isotropic input distribution $\mathbf{K}_m = \frac{\phi}{L_t} \mathbf{I}_{L_t}$, $\forall m$. The input SNR is defined as $10 \log_{10}(\phi)$ dB. For each SNR, the initial rate R is optimized, to maximize the LT throughput.

scale MIMO-ARQ into an equivalent SISO-ARQ setup, which simplifies the performance analysis considerably.

Consider a MIMO transmission scheme utilizing an optimal, e.g., INR, ARQ protocol. We can use the results of [51, Section II.A] to show that the random variable $C_{(m)} = \log |\mathbf{I}_{L_r} + \frac{\phi}{L_t} \mathbf{H} \mathbf{H}^H|$, $\forall m$, i.e., the AMIs of the ARQ protocol, converges in distribution to a Gaussian random variable \mathcal{Y} which, depending on the number of transmit/receive antennas, has the following characteristics

$$\mathcal{Y} \sim \begin{cases} \mathcal{N}(L_t \log(1 + \frac{L_r \phi}{L_t}), \frac{L_r}{L_t}) & \text{if large } L_r, \text{ fixed } L_t \\ \mathcal{N}(L_r \log(1 + \phi), \frac{L_r \phi^2}{L_t(1+\phi)^2}) & \text{if large } L_t, \text{ fixed } L_r \\ \mathcal{N}(L_r \phi, \frac{L_r}{L_t} \phi^2) & \text{if large } L_r \& L_t, \text{ small } \phi \\ \mathcal{N}(\check{\mu}, \check{\sigma}^2) & \text{if large } L_r \& L_t, \text{ large } \phi \end{cases}$$

$$\begin{aligned} \check{\mu} &= L_{\min} \log(\frac{\phi}{L_t}) + L_{\min} (\sum_{i=1}^{L_{\max}-L_{\min}} \frac{1}{i} - \gamma) \\ &\quad + \sum_{i=1}^{L_{\min}-1} \frac{i}{L_{\max}-i}, \gamma = 0.5772 \dots \\ \check{\sigma}^2 &= \sum_{i=1}^{L_{\min}-1} \frac{i}{(L_{\max}-L_{\min}+i)^2} + L_{\min} (\frac{\pi^2}{6} - \sum_{i=1}^{L_{\max}-1} \frac{1}{i^2}), \\ L_{\max} &= \max(L_t, L_r), L_{\min} = \min(L_t, L_r). \end{aligned} \quad (14)$$

Using (14), we have $\Pr(C_{(m)} \leq R_{(m)}) \rightarrow F_{\mathcal{Y}}(R_{(m)})$ which let us to investigate the performance of large-scale MIMO-ARQ protocols with very high accuracy⁴. For instance, setting $L_r = 1$, Fig. 3 shows the LT throughput and the average rate of an optimal ARQ approach which is obtained via (14) and compares the results with exact ones obtained by $C_{(m)} = \log |\mathbf{I}_{L_r} + \mathbf{H} \mathbf{K} \mathbf{H}^H|$, $\forall m$. The figure indicates the accuracy of the approximation in (14). Therefore, assuming iid Rayleigh-fading channels, the data transmission efficiency of large-scale MIMO-ARQ can be studied via a SISO-ARQ setup following (14), for which many results have been previously derived [27]–[32].

³All simulation results are given for iid Rayleigh-fading channels where each element of the channel matrix \mathbf{H} follows $\mathcal{CN}(0, 1)$.

⁴According to [51], the approximations in (14) are very accurate even for moderate values of L_t and L_r .

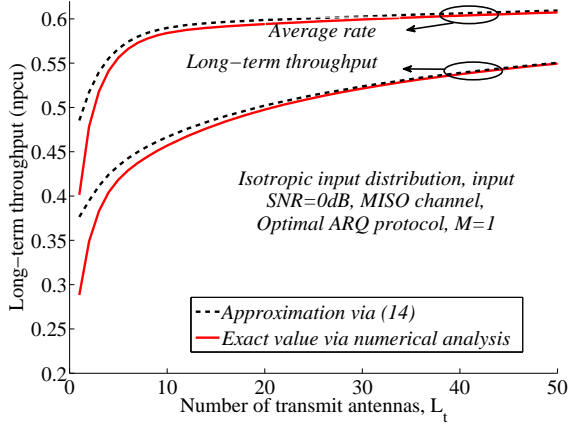


Figure 3. LT Throughput and the average rate versus the number of transmit antennas, MISO channel with an isotropic input distribution $\mathbf{K} = \frac{\phi}{L_t} \mathbf{I}_{L_t}$, transmission SNR $10 \log_{10} \phi = 0$ dB, an optimal ARQ protocol with a maximum of $M = 1$ retransmission. For each number of transmit antennas, the initial rate R is optimized, to maximize the LT throughput/average rate.

IV. ON THE EFFECT OF BURSTY COMMUNICATION

ARQ protocols are normally studied in two different data communication models, namely, bursty [27], [28], [31] and continuous [30], [31]. In contrast to the bursty model where there is an idle period between two successive packet transmissions, the continuous model is based on the assumption that a new packet transmission starts as soon as the previous packet transmission ends. Continuous communication is an appropriate model for the cases where there is a large pool of information nats to be sent to the receivers. On the other hand, the bursty communication model is more accurate for cases where the spectrum is used sporadically, for instance in random access, cognitive radio, spectrum sharing or TDMA-based networks [27], [28], [31], [52], [53]. Particularly, the bursty model is of interest when we remember that, according to Federal Communications Commission (FCC), temporal and geographical variations in the utilization of the assigned spectrum range from 15% to 85% with a high variance in time [53].

From the discussions of the previous section, the goal of this part is to evaluate the effect of bursty communication on the performance of the MIMO-ARQ setup. Along with deriving the empirical fading pdf, two interesting conclusions of the section are 1) a MIMO-ARQ setup can be mapped to an equivalent SISO-ARQ model, in the sense that the same performance is achieved in the two schemes, and 2) depending on the fading condition, there may be no ARQ protocol that can reach the channel ergodic capacity in a bursty communication model. Also, as a side-result, we show that to maximize the LT throughput and for large M , the equivalent data rates of an optimal ARQ protocol should follow a geometric sequence.

Let us first illustrate the following example.

Example 3: (Equivalent ARQ models) As explained before, the performance of an ARQ protocol is independent of the fading model as soon as the pdfs of the AMIs are given.

Also, for an optimal MIMO-ARQ with non-adaptive input covariance matrix \mathbf{K} we have $C_{(m)}^{\text{MIMO}} = \log |\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h|$, $\forall m$. On the other hand, the AMIs of the optimal ARQ scheme in a SISO channel is $C_{(m)}^{\text{SISO}} = \log(1 + \Upsilon)$ where Υ is the channel SNR random variable. Therefore, the performance of the MIMO and the SISO setups are the same if $F_{\Upsilon}(x) = F_{|\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h| - 1}(x)$, $\forall x$, which leads to $\Pr(C_{(m)}^{\text{MIMO}} \leq x) = \Pr(C_{(m)}^{\text{SISO}} \leq x)$, $\forall x$ (Note that $|\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h| \geq 1$). In simple words, using optimal ARQ it is possible to map a MIMO channel to a SISO setup, in the sense that the same performance is obtained in these schemes. Finally, the discussions were presented for the optimal ARQ schemes. However, as also seen in the following, the same arguments are valid for the non-optimal ARQ protocols, e.g., the RTD.

In contrast to Example 2 on large-scale MIMO-ARQ, there is no approximation in Example 3. Also, it should be mentioned that the concept of mapping the open-loop MIMO (without feedback) to a SISO channel has been previously indicated by, e.g., [54]. In Example 3, however, the main point is that the performance of the ARQ-based approach is fully characterized by $C_{(m)}$'s which lets us derive the conclusion for the MIMO-ARQ protocols. Moreover, the example has interesting outcomes such as:

- Recently, several theoretical analysis/comparisons for the ARQ protocols have been presented in SISO systems where the assertions are valid independent of the channel distribution [27], [30], [31]. In this way, Example 3 shows that all results previously proved by, e.g., [27], [30], [31] for the SISO channels also hold for the MIMO setups, as they can be mapped to an equivalent SISO model.
- The example helps us to 1) analyze different power allocation methods, 2) discuss about the optimal equivalent data rates of the ARQ protocols and also 3) find the empirical pdf of the channel quality in a MIMO-ARQ setup. The first item is covered in Section VI. The two last items are studied here as follows.

The key difference between the bursty and continuous models returns back to the way the fading channel is observed at the transmission endpoints. Clearly, the empirical channel distribution matches the true one under the continuous communication model. This is because all channel uses of a fading block are utilized for data transmission. Under the bursty communication model, however, the channel is seen as *worse* than what it is in reality. The reason is that if the channel is *good*, the packet transmission ends at the first (re)transmission round(s). However, many channel uses are utilized for sending a packet when the channel is *bad*. Hence, a large portion of the data (re)transmission is carried out when the channel experiences low quality, while the transmitter is mostly off when the channel is good.

To be more specific, we find the empirical channel quality distribution in the bursty model; Assuming non-adaptive power allocation, let $f_{\Upsilon}(v)$ be the true channel quality pdf as explained in Example 3. With the same procedure as in (5), the number of channel uses for different channel qualities is

$$l(v) = \begin{cases} l_{(m)} & \text{if } e^{R_{(m)}} - 1 \leq v < e^{R_{(m-1)}} - 1, m = 1, \dots, M \\ l_{(M+1)} & \text{if } v < e^{R_{(M)}} - 1. \end{cases} \quad (15)$$

Considering the number of fading blocks $N \rightarrow \infty$, the total number of channel uses in which the channel SNR $\Upsilon = v$ is observed is $Nl(v)f_{\Upsilon}(v)$. Moreover, the total number of channel uses in N fading blocks is $N \sum_{\forall v} l(v)f_{\Upsilon}(v)$. Hence, the empirical channel quality pdf and cdf in the bursty model are respectively found as

$$\begin{aligned} f_{\Upsilon}^{\text{emp}}(v) &= \frac{l(v)f_{\Upsilon}(v)}{\sum_{\forall v} l(v)f_{\Upsilon}(v)} \\ &= \begin{cases} \frac{l_{(m)}f_{\Upsilon}(v)}{E\{\tilde{L}\}} & \text{if } e^{R_{(m)}} - 1 \leq v < e^{R_{(m-1)}} - 1, m = 1, \dots, M \\ \frac{l_{(M+1)}f_{\Upsilon}(v)}{E\{\tilde{L}\}} & \text{if } v < e^{R_{(M)}} - 1 \end{cases} \\ F_{\Upsilon}^{\text{emp}}(v) &= \int_0^v f_{\Upsilon}^{\text{emp}}(x)dx \\ &= \begin{cases} \frac{F_{\Upsilon}(e^{R_{(M)}}-1)}{R_{(M+1)}\Omega} + \sum_{n=1}^{m-1} \frac{F_{\Upsilon}(e^{R_{(M-n)}}-1) - F_{\Upsilon}(e^{R_{(M+1-n)}}-1)}{R_{(M+1-n)}\Omega} \\ + \frac{F_{\Upsilon}(v) - F_{\Upsilon}(e^{R_{(M+1-m)}}-1)}{R_{(M+1-m)}\Omega} & \text{if } e^{R_{(M+1-m)}} - 1 \leq v < e^{R_{(M-m)}} - 1, m = 1, \dots, M \\ \frac{F_{\Upsilon}(v)}{R_{(M+1)}\Omega} & \text{if } v < e^{R_{(M)}} - 1, \end{cases} \\ \Omega = \frac{E\{\tilde{L}\}}{Q} &= \sum_{m=1}^{M+1} \frac{F_{\Upsilon}(e^{R_{(m-1)}}-1) - F_{\Upsilon}(e^{R_{(m)}}-1)}{R_{(m)}} + \frac{F_{\Upsilon}(e^{R_{(M+1)}}-1)}{R_{(M+1)}} \end{aligned} \quad (16)$$

With $R_{(m)} = \frac{R}{m}$, $R = 1$ and Rayleigh-fading condition, Fig. 4 compares the empirical channel quality cdf with the true one when implementing a fixed-length optimal ARQ protocol, e.g., fixed-length INR. Again, note that the empirical SNR cdf matches the true channel SNR cdf in the continuous communication model (black dashed curves in the figure). In harmony with the discussions, we observe that the empirical cdf is worse in the bursty model (the curve has shifted to the left). Moreover, with 0dB input SNR and $M = 1$, Fig. 5 demonstrates the throughput versus the initial transmission rate in different data communication models⁵. As it can be seen, there is a gap between the LT throughput of the two models at moderate values of R . Importantly, the maximum throughput achieved in these models is considerably different. With low (resp. high) rates, however, the two models lead to the same LT throughput, because the data is always decoded at the first round (resp. it is almost never decoded and the LT throughput converges to zero.). Finally, how much the cdf curve is shifted to the left and the LT throughput decreases due to bursty communication depends on the number of transmit/receive antennas, equivalent data rates, number of retransmissions, etc.

To further clarify the differences between the bursty and continuous models, we study their asymptotic behavior as follows.

With $M \rightarrow \infty$ retransmissions and a quasi-static fading channel, Shen, et. al, have shown that the LT throughput of the variable-length coding INR protocol converges to the channel

⁵The LT throughput is obtained by (7) in the bursty model. Under continuous communication model, however, [30], [31] showed that the LT throughput (3) degenerates to the average rate (11), i.e., $\eta = r$. That is, the solid lines (resp. dashed lines) of Fig. 5 are obtained by (7) (resp. (11)).

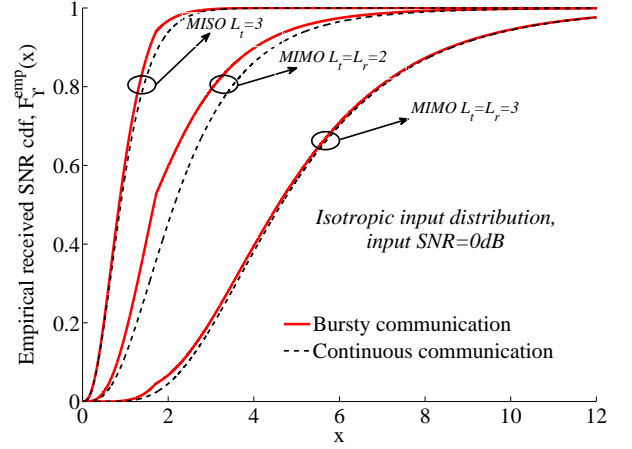


Figure 4. Empirical cdf of the received SNR, isotropic input distribution, transmission SNR=0dB, $M = 1$. For the continuous communication model the empirical cdf matches the true one.

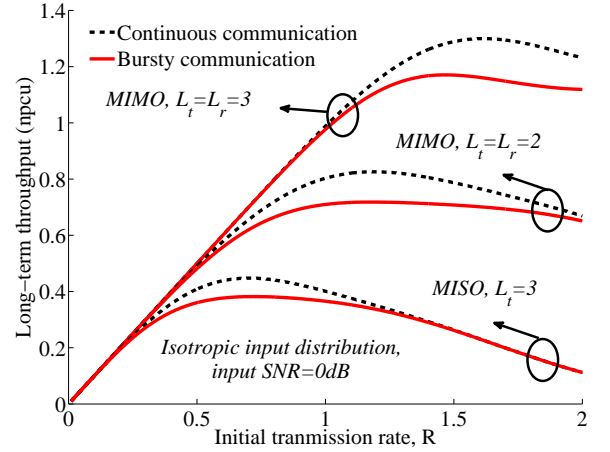


Figure 5. LT Throughput vs the initial transmission rate R in different data communication models, isotropic input distribution, SNR=0dB, $M = 1$.

ergodic capacity $E\{C\} = \int_0^\infty f_{\Upsilon}(v) \log(1+v)dv$ if the data is transmitted continuously [31, lemma 2]⁶. However, using (16) and because of the worse fading distribution, it can be easily proved that there is no ARQ scheme that can reach the channel ergodic capacity, if the data is transmitted in a bursty fashion and the channel is quasi-static. This point, which is because the good channels are not fully utilized with the bursty model, is further elaborated in Example 4.

Example 4: (Feedback load versus LT throughput) Shown in Fig. 6 is the feedback load (10) versus the LT throughput (9) for the 2×2 and 3×3 MIMO setups utilizing an optimal ARQ and bursty communication. The results are obtained for different number of retransmissions when the initial rate R goes from zero to infinity. Also, the points $\eta_{\max,2}$ and $\eta_{\max,3}$ show the maximum LT throughput with $L_t = L_r = 2$ and $L_t = L_r = 3$, respectively, and a maximum of $M = 1$ retransmission

⁶In [31], the SISO model is considered. But, the same argument as in [31, lemma 2] is valid for the MIMO setup too.

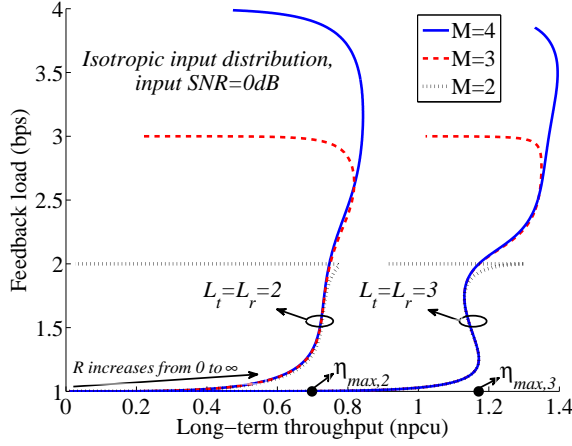


Figure 6. Feedback load (in bit per slot (bps)) vs LT throughput, isotropic input distribution $\mathbf{K} = \frac{\phi}{L_t} \mathbf{I}_{L_t}$, transmission SNR $10 \log_{10} \phi = 0\text{dB}$.

round, for which the feedback load is $b = 1, \forall R$. Considering the maximum LT throughput in each curve of the case with, e.g., $L_t = L_r = 3$, it is found that increasing the number of retransmissions leads to marginal LT throughput increment at the cost of considerable feedback load increment. This is one of consequences of the bursty communication model; with the bursty model, the system performance is more affected by the bad channel qualities if M increases. Thus, there is a tradeoff between increasing the number of retransmissions and deteriorating the empirical channel quality and, as a result, the maximum achievable LT throughput increases with M marginally. Finally, as a side-result, Fig. 6 shows that with a given feedback load increasing the number of transmit/receive antennas results in substantial LT throughput increment.

Motivated by Fig. 6, the rest of the simulation results are given with a maximum of $M = 1$ retransmission round, i.e., a maximum of two (re)transmission rounds, unless otherwise stated. Also, we do not consider the continuous communication model any further. Finally, talking about large number of retransmissions, we use the results of Example 3, on the equivalence of SISO- and MIMO-ARQ models, to determine the equivalent data rates of an optimal ARQ protocol, when M increases. The result, which is presented in Theorem 1, is interesting when we remember that depending on the fading pdf and number of retransmissions there is no general closed-form solution for the equivalent data rates maximizing the LT throughput [27], [31], [32], [55].

Theorem 1: Consider non-adaptive power allocation over the retransmissions, $\mathbf{K}_m = \mathbf{K}, \forall m$, in an optimal MIMO-ARQ model defined in Definition 1. For high values of m , the optimal (in terms of LT throughput) equivalent data rates, i.e., the set of $R_{(m)}$'s, can be approximated by the geometric sequence $R_{(m)}^2 = R_{(m-1)} R_{(m+1)}$.

Proof. See the appendix. \square

Note that the conclusion of the theorem is independent of the fading pdf and the total transmission power.

Finally, it is worth noting that, with the bursty communication model, the empirical pdf matches the true one if the

channel is temporally-independent slow-fading. The reason is that an independent channel realization is experienced in each sub-codeword transmission, independently of the ARQ process and the message decoding status. However, the empirical and the true channel quality pdfs differ in the practical cases of temporally-correlated fast- and slow-fading channels, for instance, in the indoor ultra wideband (UWB) channels [41], [56], where the channel quality varies slowly and smoothly in successive fading blocks. This is because of the ARQ random packet length which, with temporally-dependent fading condition, affects the total number of channel uses in which different fading realizations are observed.

V. EXAMPLES OF MIMO-ARQ PROTOCOLS

The goal of the section is to provide more insight into the effect of AMIs on the performance of MIMO-ARQ setups. Thus, we show how the encoding and decoding procedures change the AMIs of the RTD, INR and STC-based MIMO-ARQ schemes. Also, the results are used to sort ARQ schemes based on their AMIs.

A. RTD MIMO-ARQ protocol

Utilizing RTD, we have $\mathbf{X}_m = \mathbf{X}_1, \forall m$, and the receiver performs MRC of the received signals. In other words, each round of RTD is equivalent to adding L_r antennas at the receiver [11], [12], [16], [17]. Therefore, the equivalent channel model at the end of the m -th round, i.e., (1), is rephrased as

$$\mathbf{Y}_{(m)}^{\text{RTD}} = \mathbf{H}_{(m)}^{\text{RTD}} \mathbf{X}_1 + \mathbf{Z}_{(m)}^{\text{RTD}},$$

$$\mathbf{H}_{(m)}^{\text{RTD}} = \underbrace{[\mathbf{H}_1^T \dots \mathbf{H}_m^T]^T}_{m \text{ times}}, \mathbf{Y}_{(m)}^{\text{RTD}}, \mathbf{Z}_{(m)}^{\text{RTD}} \in \mathcal{C}^{mL_r \times 1} \quad (17)$$

which is an $mL_r \times L_t$ MIMO channel. Thus, according to the capacity of the MIMO channels [47], the AMI, i.e., the maximum decodable rate, at the end of the m -th round of the RTD protocol is obtained by

$$C_{(m)}^{\text{RTD}} = \frac{1}{m} \log \left| \mathbf{I}_{mL_r} + \mathbf{H}_{(m)}^{\text{RTD}} \mathbf{K} (\mathbf{H}_{(m)}^{\text{RTD}})^H \right|. \quad (18)$$

Here, the term $\frac{1}{m}$ is due to that as the data is repeated m times the maximum decodable rate is divided by m [29]–[31]. Finally, note that, as the RTD is a fixed-length coding ARQ scheme, we have $R_{(m)} = \frac{R}{m}$.

B. INR ARQ protocol

Using INR, new sub-codewords are (re)transmitted in the (re)transmissions and in each round the receiver combines all signals received up to the end of that round. Following the discussions in [57, chapter 15], [58, chapter 7], [59], the maximum decodable data rate of the INR protocol with different sub-codewords of lengths l_m and input covariance matrices \mathbf{K}_m is obtained by the TDMA-type equation

$$C_{(m)}^{\text{INR}} = \frac{\sum_{n=1}^m l_n \log |\mathbf{I}_{L_r} + \mathbf{H} \mathbf{K}_n \mathbf{H}^H|}{\sum_{n=1}^m l_n}$$

$$= R_{(m)} \sum_{n=1}^m \left(\frac{1}{R_{(n)}} - \frac{1}{R_{(n-1)}} \right) \log |\mathbf{I}_{L_r} + \mathbf{H} \mathbf{K}_n \mathbf{H}^H|. \quad (19)$$

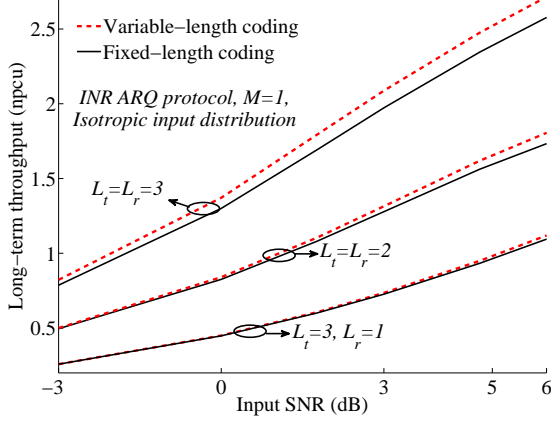


Figure 7. Effect of variable- and fixed-length coding on the LT throughput, isotropic input distribution $\mathbf{K} = \frac{\phi}{L_t} \mathbf{I}_{L_t}$, INR ARQ protocol with a maximum of $M = 1$ retransmission round. The rate parameters are optimized to maximize the LT throughput.

Specifically, with non-adaptive input covariance matrix $\mathbf{K}_m = \mathbf{K}$, $\forall m$, (19) is simplified to

$$C_{(m)}^{\text{INR}} = \log |\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h|, \forall m. \quad (20)$$

Thus, from (19)-(20) and as stated before, the INR is an optimal ARQ protocol (review Definition 1 introducing optimal ARQ protocol).

Although having high LT throughput and low outage probability, variable-length coding INR results in high *packet* complexity [27], [30], [31]. In order to reduce the complexity, fixed-length coding INR scheme can be considered where setting $l_m = l$, $\forall m$, in (19) leads to

$$C_{(m)}^{\text{INR, fixed-length}} = \frac{1}{m} \sum_{n=1}^m \log |\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}_n\mathbf{H}^h|. \quad (21)$$

Shown in Fig. 7 is the LT throughput of the MIMO-INR protocol for fixed- and variable-length coding schemes. According to the figure, variable-length coding is of less interest in the cases with small input SNR or number of transmit/receive antennas. However, sizable performance improvement is achieved by variable-length coding when the transmission power or the number of antennas increases.

The superiority of the INR over the RTD has been shown by [16], [27], [30], [31] for different performance metrics, and can be proved by comparing the AMIs as well. However, setting $R_{(m)} = \frac{R}{m}$ and non-adaptive input covariance matrix $\mathbf{K}_m = \epsilon \mathbf{K}'$, $\forall m$, where $\epsilon \rightarrow 0$, (20) is rephrased as

$$\lim_{\epsilon \rightarrow 0} C_{(m)}^{\text{INR}} = \epsilon \text{Tr}(\mathbf{H}\mathbf{K}'\mathbf{H}^h) \quad (22)$$

which is the same as in the RTD, where using (17) and (18) it can be written

$$\lim_{\epsilon \rightarrow 0} C_{(m)}^{\text{RTD}} = \frac{1}{m} \epsilon \text{Tr}(\mathbf{H}_{(m)}^{\text{RTD}} \mathbf{K}' (\mathbf{H}_{(m)}^{\text{RTD}})^h) = \epsilon \text{Tr}(\mathbf{H}\mathbf{K}'\mathbf{H}^h). \quad (23)$$

Thus, from Remark 1, the RTD and fixed-length INR lead to the same performance at low SNRs.

C. STC-based MIMO-ARQ protocols

Different kinds of STCs are normally used in MIMO-ARQ protocols, e.g., [16], [17]. Here, setting $M = 1$, we obtain the AMIs when the Alamouti, the CDD, the SMR or the AS codes are utilized in a 2×1 MISO-ARQ channel. Our reasons for selecting the 2×1 MISO setup are 1) the analytical results are tractable and 2) many STCs, such as the Alamouti code [20], were first designed for the 2×1 MISO channel.

While the AMIs of the considered STCs are summarized in Table 1, because of the mathematical similarity, we only explain the procedure for deriving the AMIs of the CDD code; with a 2×1 MISO setup, the mother code of the CDD is $\mathbf{X} = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}$. In the original CDD, the codeword \mathbf{X} is sent to the receiver via two channel uses. Using ARQ, however, the data is sent in two rounds as follows. In the first round, $\mathbf{X}_1 = [x_1 \ x_2]^T$ is sent via one channel use. Thus, the received signal is $Y_{(1)} = [h_1 \ h_2][x_1 \ x_2]^T + Z_{(1)}$ and the AMI is obtained by $C_{(1)}^{\text{CDD}} = \log(1 + \frac{\phi}{2}[h_1 \ h_2][h_1 \ h_2]^h) = \log(1 + \frac{\phi}{2}\chi_2)$. Here, the last equality is for Rayleigh-fading channels, on which we focus, and $\chi_k = \sum_{i=1}^k |h_i|^2$ is a random variable following the Chi-square distribution with $2k$ degrees of freedom. If the data is not decoded at the end of the first round, one more channel use is used to send $\mathbf{X}_2 = [x_2 \ x_1]^T$ and the receiver combines the two received signals. Hence, the equivalent channel is

$$\mathbf{Y}_{(2)} = [h_1 \ h_2] \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix} + \mathbf{Z}_{(2)} \equiv \begin{bmatrix} h_1 & h_2 \\ h_2 & h_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \mathbf{Z}_{(2)}. \quad (24)$$

In simple words, using CDD in two channel uses is equivalent to transmitting the codeword $\mathbf{X}_1 = [x_1 \ x_2]^T$ in the 2×2 MIMO with channel matrix $\begin{bmatrix} h_1 & h_2 \\ h_2 & h_1 \end{bmatrix}$. Thus, the AMI of the second round is

$$\begin{aligned} C_{(2)}^{\text{CDD}} &= \frac{1}{2} \log \left| \mathbf{I}_2 + \frac{\phi}{2} \begin{bmatrix} h_1 & h_2 \\ h_2 & h_1 \end{bmatrix} \begin{bmatrix} h_1 & h_2 \\ h_2 & h_1 \end{bmatrix}^h \right| \\ &= \frac{1}{2} \log \left| \mathbf{I}_2 + \frac{\phi}{2} \begin{bmatrix} \chi_2 & 2\text{Re}\{h_1 h_2^*\} \\ 2\text{Re}\{h_1 h_2^*\} & \chi_2 \end{bmatrix} \right| \\ &= \frac{1}{2} \log((1 + \frac{\phi}{2}\chi_2)^2 - \phi^2 \text{Re}\{h_1 h_2^*\}^2). \end{aligned} \quad (25)$$

The results for the other STCs considered in Table 1 are obtained with the same procedure. The effect of power allocation on the performance of STC-based MIMO-ARQ setups is studied in Section VI.

Finally, it is worth noting that with $L_t \times 1$ MISO channel $\mathbf{H} = [h_1 \dots h_{L_t}]$ and Rayleigh-fading condition the AMIs of the RTD and INR protocols, i.e., (18) and (20) are rephrased as

$$C_{(m)}^{\text{RTD}} = \frac{1}{m} \log(1 + \frac{m\phi}{L_t} \chi_{L_t}), \chi_{L_t} = \sum_{i=1}^{L_t} |h_i|^2, \quad (26)$$

and

$$C_{(m)}^{\text{INR}} = \log(1 + \frac{\phi}{L_t} \chi_{L_t}), \quad (27)$$

respectively, where $f_{\chi_{L_t}}(x) = \frac{1}{\Gamma(L_t)} x^{L_t-1} e^{-x}$, $x \geq 0$, and $\Gamma(\cdot)$ is the Gamma function. Here, (26) follows from (18) and $|\mathbf{I}_n +$

Table I
THE EQUIVALENT CHANNEL MODELS AND THE AMIS OF DIFFERENT STC-BASED ARQ SCHEMES.

Mother STC method	Codeword in round # 1	Equivalent channel model	AMI of round # 1
AS: $\begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ 0 \end{bmatrix}$	$[h_1 \ 0]$	$C_{(1)}^{\text{AS}} = \log(1 + \phi\chi_1)$
SMR: $\begin{bmatrix} x_1 & x_1 \\ x_2 & x_2 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$	$[h_1 \ h_2]$	$C_{(1)}^{\text{SMR}} = \log(1 + \frac{\phi}{2}\chi_2)$
Alamouti: $\begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$	$[h_1 \ h_2]$	$C_{(1)}^{\text{Alamouti}} = \log(1 + \frac{\phi}{2}\chi_2)$
CDD: $\begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$	$[h_1 \ h_2]$	$C_{(1)}^{\text{CDD}} = \log(1 + \frac{\phi}{2}\chi_2)$
Mother STC method	Codeword in round # 2	Equivalent channel model	AMI of round # 2
AS: $\begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}$	$\begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}$	$[h_1 \ h_2]$	$C_{(2)}^{\text{AS}} = \frac{1}{2} \log(1 + \phi\chi_2)$
SMR: $\begin{bmatrix} x_1 & x_1 \\ x_2 & x_2 \end{bmatrix}$	$\begin{bmatrix} x_1 & x_1 \\ x_2 & x_2 \end{bmatrix}$	$\begin{bmatrix} h_1 & h_2 \\ h_1 & h_2 \end{bmatrix}$	$C_{(2)}^{\text{SMR}} = \frac{1}{2} \log(1 + \phi\chi_2)$
Alamouti: $\begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$	$\begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$	$\begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}$	$C_{(2)}^{\text{Alamouti}} = \log(1 + \frac{\phi}{2}\chi_2)$
CDD: $\begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}$	$\begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}$	$\begin{bmatrix} h_1 & h_2 \\ h_2 & h_1 \end{bmatrix}$	Eq. (25)

$x\mathbf{J}_n| = 1 + nx$ where \mathbf{J}_n is the $n \times n$ ones matrix. Interestingly, (26) indicates that the same performance is achieved by the MISO-RTD and a SISO-RTD model with fading pdf $f_{\chi_{L_1}}$.

Remark 2: (Sorting the ARQ protocols) Using Table 1, we can write

$$\begin{cases} C_{(1)}^{\text{RTD}} = C_{(1)}^{\text{SMR}} = C_{(1)}^{\text{CDD}} = C_{(1)}^{\text{Alamouti}} = C_{(1)}^{\text{INR, fixed-length}}, \forall \mathbf{H} \\ C_{(2)}^{\text{RTD}} = C_{(2)}^{\text{SMR}} \leq C_{(2)}^{\text{CDD}} \leq C_{(2)}^{\text{Alamouti}} = C_{(2)}^{\text{INR, fixed-length}}, \forall \mathbf{H} \end{cases} \quad (28)$$

which, according to Remark 1, leads to

$$\theta^{\text{RTD}} = \theta^{\text{SMR}} \prec \theta^{\text{CDD}} \prec \theta^{\text{Alamouti}} = \theta^{\text{INR, fixed-length}} \quad (29)$$

with \prec defined in Remark 1 (Clearly, $\theta^{\text{INR, fixed-length}} \prec \theta^{\text{INR, variable-length}}$). However, letting $\phi \rightarrow 0$ (low SNRs) in Table 1, (29) is simplified to

$$\theta^{\text{RTD}} = \theta^{\text{SMR}} = \theta^{\text{CDD}} = \theta^{\text{Alamouti}} = \theta^{\text{INR, fixed-length}}. \quad (30)$$

The remark demonstrates the usefulness of the STC techniques; as also seen in Table 1, the developed STC methods are mainly based on repeating a permuted version of the initial sub-codeword. Thus, their encoding and decoding complexity is comparable with the ones in the RTD protocol, and less than the complexity in the INR. On the other hand, proper selection of the STC approach (for instance, the Alamouti code) results in appropriate performance of the MIMO-ARQ scheme, comparable with performance in the INR protocol.

Finally, the simulation results for different MIMO-ARQ techniques are presented later in Section VI, where the system performance is obtained for both adaptive and non-adaptive power allocation approaches. The results, which are presented in Fig. 8, emphasize the validity of Remark 2.

VI. ON THE EFFECT OF POWER ALLOCATION

With imperfect CSIT, the optimal power allocation for the MIMO channels is still an open problem. For the MIMO-ARQ systems power allocation can be done within the

(re)transmissions (temporal power allocation), between the antennas (spatial power allocation) or both. We present some discussions on the two first cases. As demonstrated, the discussions on the AMIs are useful for deriving the (sub)optimal power allocation strategies.

A. Temporal power allocation

Let the input power in the m -th (re)transmission round be ϕ_m . If the data is decoded at the end of the m -th round, and not before, the total consumed energy is $\xi_{(m)} = \sum_{n=1}^m l_n \phi_n$. Also, the consumed energy is found as $\xi_{(M+1)} = \sum_{n=1}^{M+1} l_n \phi_n$ if an outage occurs. In this way, (6) can be used to find the average transmission power $\Phi = \frac{\mathbb{E}\{\xi\}}{\mathbb{E}\{L\}}$ [27]–[31], [59], i.e., the ratio of the expected consumed energy and the expected number of channel uses, by

$$\begin{aligned} \Phi &= \frac{1}{\mathbb{E}\{\tilde{L}\}} \left(\sum_{m=1}^{M+1} \left(\sum_{n=1}^m l_n \phi_n \right) \Pr(\bar{A}_0, \dots, \bar{A}_{m-1}, A_m) \right. \\ &\quad \left. + \left(\sum_{n=1}^{M+1} l_n \phi_n \right) \Pr(\bar{A}_0, \dots, \bar{A}_{M+1}) \right) \\ &\stackrel{(a)}{=} \frac{\sum_{m=1}^{M+1} \phi_m \left(\frac{1}{R_{(m)}} - \frac{1}{R_{(m-1)}} \right) \Pr(C_{(m-1)} < R_{(m-1)})}{\sum_{m=1}^{M+1} \left(\frac{1}{R_{(m)}} - \frac{1}{R_{(m-1)}} \right) \Pr(C_{(m-1)} < R_{(m-1)})}. \end{aligned} \quad (31)$$

Here, (a) is obtained with the same procedure as in (6). Considering (31), an average power constraint $\Phi \leq \phi$ and (6)–(11), we can optimize different metrics, e.g., outage probability and LT throughput, via power allocation.

More specifically, the equivalent channel matrix of the RTD protocol in (17) is rephrased as $\mathbf{H}_{(m)}^{\text{RTD}} = [\sqrt{\phi_1} \mathbf{H}^T \dots \sqrt{\phi_m} \mathbf{H}^T]^T$, if the input signal is scaled in the retransmissions. For instance, with a temporally-adaptive isotropic Gaussian input distribution $\mathbf{K}_m = \frac{\phi_m}{L_1} \mathbf{I}_{L_1}$ the AMI (18) is found as $C_{(m)}^{\text{RTD}} = \frac{1}{m} \log |\mathbf{I}_{mL_r} + \Theta_m \otimes \mathbf{H} \mathbf{H}^H|$ where for

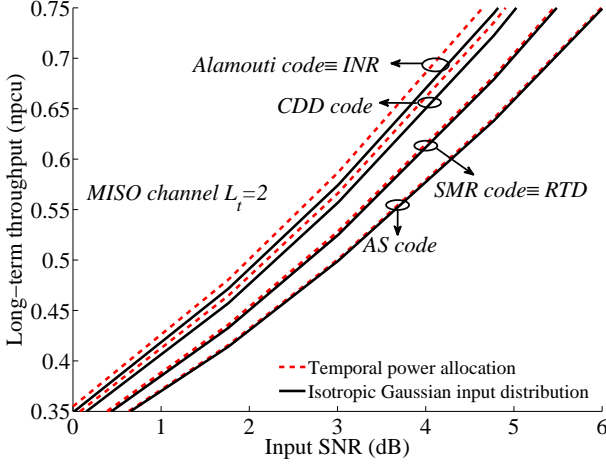


Figure 8. The LT throughput for different STCs. MISO channel $L_t = 2$, $M = 1$. With temporal (resp. isotropic input distribution) we have $\mathbf{K}_m = \frac{\phi_m}{L_t} \mathbf{I}_{L_t}$ (resp. $\mathbf{K}_m = \frac{\phi}{L_t} \mathbf{I}_{L_t}$, $m = 1, 2$). The terms ϕ_m , $m = 1, 2$, are optimized based on the average power constraint (31). The same LT throughput as in the Alamouti code (resp. SMR code) is achieved by the MISO-INR (resp. MISO-RTD) with $M = 1$.

the $L_t \times 1$ RTD-based MISO setup we have $C_{(m)}^{\text{RTD}} = \frac{1}{m} \log(1 + (\sum_{n=1}^m \phi_n) \chi_{L_t})$. Here, $\Theta_m(i, j) = \sqrt{\phi_i \phi_j}$, $i, j = 1, \dots, m$, and \otimes denotes the Kronecker product.

For INR, we have $C_{(m)}^{\text{INR}} = R_{(m)} \sum_{n=1}^m (\frac{1}{R_{(n)}} - \frac{1}{R_{(n-1)}}) \log |\mathbf{I}_{L_t} + \frac{\phi_n}{L_t} \mathbf{H} \mathbf{H}^H|$ if adaptive isotropic Gaussian input distribution is considered in (19). Moreover, considering the $L_t \times 1$ MISO-INR system the AMI is rephrased as $C_{(m)}^{\text{INR}} = R_{(m)} \sum_{n=1}^m (\frac{1}{R_{(n)}} - \frac{1}{R_{(n-1)}}) \log(1 + \frac{\phi_n}{L_t} \chi_{L_t})$. Also, we can scale the columns of the mother codes of Table 1 with terms $\sqrt{\phi_m}$ and use the same procedure as before to determine the AMIs of the considered STCs in the presence of temporal power allocation. Finally, note that many results such as [60, eq. (23)] and [61, eq. (16), (18)] can be extended to bound the AMIs of the MIMO-ARQ and, consequently, the system performance, in the presence of temporal power allocation. However, due to space limits, we do not go into details.

Fig. 8 demonstrates the LT throughput of different MISO-ARQ techniques in the presence of temporally-adaptive and non-adaptive power allocation. The results emphasize the validity of Remark 2 where, for each SNR, we have $\eta^{\text{AS}} \leq \eta^{\text{SMR}} \leq \eta^{\text{CDD}} \leq \eta^{\text{Alamouti}}$ and the difference between the LT throughput of different schemes decreases at low SNR. Also, adaptive power allocation is observed to be effective for the Alamouti and CDD codes while its effect is negligible in the SMR and AS codes.

B. Spatial power allocation

Instead of power allocation over time, the adaptation can be between the antennas where, while the overall transmit power is constant in different retransmissions, in each round the antennas experience different transmit powers. This is still a difficult problem to solve analytically. However, the AMI

discussions are helpful for numerical solution of the problem as follows.

Lemma 2: Consider two possible input covariance matrices \mathbf{K} and \mathbf{K}' . The matrix \mathbf{K} outperforms \mathbf{K}' , in terms of AMIs and, consequently, in performance metrics such as (6)-(11), if $F_{\Upsilon}(x) \leq F_{\Upsilon'}(x)$, $\forall x > 0$. Here, Υ and Υ' are the equivalent SISO channel SNR random variables, defined in Example 3, that are obtained from \mathbf{K} and \mathbf{K}' , respectively.

Proof. Lemma 2 is a consequence of Lemma 1; for simplicity, the proof is given for an optimal ARQ protocol, e.g., INR, with $C_{(m)} = \log |\mathbf{I}_{L_t} + \mathbf{H} \mathbf{K} \mathbf{H}^H|$, $\forall m = 1, \dots, M + 1$, while the same arguments hold for non-optimal ARQ protocols as well. With $F_{\Upsilon}(x) \leq F_{\Upsilon'}(x)$, $\forall x > 0$, we have $\Pr(C_{(m)} = \log(1 + \Upsilon) \leq R_{(m)}) \leq \Pr(C'_{(m)} = \log(1 + \Upsilon') \leq R_{(m)})$, $\forall R_{(m)}$. Therefore, since the performance of the ARQ schemes is a decreasing function of $\Pr(C_{(m)} \leq R_{(m)})$, the lemma is proved immediately. \square

According to Example 3 on equivalent SISO- and MIMO-ARQ models and Lemma 2, the iterative optimization algorithm stated in Algorithm 1 can be utilized to determine the sub-optimal covariance matrix optimizing the system performance. The algorithm, which roots from machine learning concepts [30], is simple and very efficient compared to exhaustive search. Also, although it has been described for a constant covariance matrix, it can be adapted in a greedy fashion for the cases with different covariance matrices in the (re)transmissions. Finally, the sub-optimality of the derived results is due to structure of the considered covariance matrices (see Step I of the algorithm) and the fact that Lemma 2 provides a sufficient but not necessary constraint for optimizing the input covariance matrix.

Algorithm 1 Sub-optimal power allocation

- I. For a given power constraint $\text{Tr}(\mathbf{K}) = \phi$, consider J , e.g. $J = 20$, randomly generated covariance matrices \mathbf{K}^j , $\text{Tr}(\mathbf{K}^j) = \phi$, $\mathbf{K}^j(n, m) = 0$, $n \neq m$, where $\mathbf{K}^j(n, m)$ is the (n, m) -th element of \mathbf{K}^j .
- II. For each matrix, find the cdf of the AMIs, e.g., $C_{(m)}^j = \log |\mathbf{I}_{L_t} + \mathbf{H} \mathbf{K}^j \mathbf{H}^H|$, denoted by $F_{C_{(m)}^j}$.
- III. Find the covariance matrix which results in the best AMI cdf, i.e., \mathbf{K}^i where $F_{C_{(m)}^i}(x) \leq F_{C_{(m)}^j}(x)$, $\forall j = 1, \dots, J, x$. Keep the best cdf of the previous iteration, if such a matrix is not found.
- IV. $\mathbf{K}^1 \leftarrow \mathbf{K}^i$.
- V. Generate $b \ll J$, e.g., $b = 5$, matrices $\mathbf{K}^{j, \text{new}}$, $j = 1, \dots, b$ around \mathbf{K}^1 . These matrices, which are generated by adding small random numbers to \mathbf{K}^1 , should also satisfy the constraints introduced in I. Then, $\mathbf{K}^{j+1} \leftarrow \mathbf{K}^{j, \text{new}}$, $j = 1, \dots, b$.
- VI. Regenerate the remaining matrices \mathbf{K}^j , $j = b + 2, \dots, J$ randomly such that the constraints in I are satisfied.
- VII. Go to II and continue until convergence (when no improvement is observed after many, e.g., 10^4 , iterations).

Setting $R = 1$, Fig. 9 shows the effect of power allocation on the LT throughput of MIMO-INR scheme. For temporal

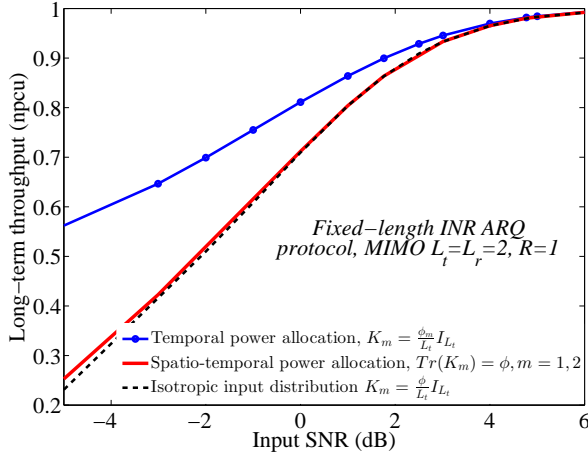


Figure 9. LT throughput vs the input SNR. MIMO-ARQ setup $L_t = L_r = 2$, $M = 1$, fixed-length INR ARQ protocol $R = 1$. With temporal (resp. isotropic input distribution) we have $\mathbf{K}_m = \frac{\phi_m}{L_t} \mathbf{I}_{L_t}$ (resp. $\mathbf{K}_m = \frac{\phi}{L_t} \mathbf{I}_{L_t}$, $m = 1, 2$) and ϕ_m 's are related together via (31). With sub-optimal spatio-temporal power allocation, we have $\text{Tr}(\mathbf{K}_m) = \phi$, $m = 1, 2$, $\mathbf{K}(i, j) = 0$, $i \neq j$, i.e., the sub-optimal diagonal matrices \mathbf{K}_m , $m = 1, 2$, are obtained through Algorithm 1 when it is adapted in a greedy fashion. The results are for fixed (non-optimized) initial rate $R = 1$.

power allocation, we have $\mathbf{K}_m = \frac{\phi_m}{L_t} \mathbf{I}_{L_t}$ where ϕ_m 's are related together through (31). For the sub-optimal spatio-temporal power allocation we have adapted Algorithm 1 in a greedy fashion such that $\text{Tr}(\mathbf{K}_m) = \phi$, $m = 1, 2$. Compared to isotropic Gaussian input distribution, substantial LT throughput increment is achieved by temporal power allocation at low SNRs. With Rayleigh-fading and the considered parameters, however, the sub-optimal spatio-temporal power allocation of Algorithm 1 is only marginally useful. At high SNRs, an isotropic Gaussian input distribution is optimal, in harmony with the literature. Finally, the gain of adaptive power allocation decreases if, along with the powers, the rate R is optimized as well.

VII. RELAXING THE SYSTEM MODEL ASSUMPTIONS

Throughout the paper, the results were obtained for a noise-free feedback channel and under the quasi-static channel assumption where the fading coefficients remain fixed during a packet period. In this section, we relax these assumptions. We illustrate that the techniques developed previously can be used in many cases with different forward/feedback channel models. Also, it is shown that the performance of MIMO-ARQ protocols with different temporal fading variations can be mapped to each other, under certain conditions.

A. On temporal variations of the fading coefficients

Suppose the fast-fading condition where T different fading realizations $\mathbf{H}(t)$, $t = (m-1)T+1, \dots, mT$, are experienced during the m -th (re)transmission round. In this case, the same procedure as in (21) can be used to obtain the AMIs of an

optimal (INR) ARQ protocol as⁷

$$C_{(m)}^{\text{INR, fast-fading}} = \frac{1}{mT} \sum_{n=1}^m \sum_{t=(n-1)T+1}^{nT} \log |\mathbf{I}_{L_r} + \mathbf{H}(t) \mathbf{K}_n \mathbf{H}(t)^H|. \quad (32)$$

Also, note that setting $T = 1$ (resp. $\mathbf{H}(t) = \mathbf{H}$, $\forall t = 1, \dots, (M+1)T$) in (32) provides the results with slow-fading (resp. quasi-static) channel condition (see Section II). Here, the important point is that the discussions of Section III are independent of the fading model. In other words, the only parameters that are affected by the fading model are the AMIs while the performance metrics, e.g., (6)-(11), are independent of the fading model as soon as the pdfs of the AMIs are given. Thus, the discussions on using the AMIs as intermediate scheme-dependent parameters for the analysis of MIMO-ARQ protocols are also valid for the slow- and fast-fading models.

It is well-known that with the same fading pdf the performance of ARQ protocols improves as the channel temporal variation increases, because more time diversity is exploited by the ARQ. In the following, however, we compare the performance of ARQ protocols in different fading conditions. The results are of interest because they provide connections between the papers considering one of the quasi-static, slow- or fast-fading models.

Theorem 2: Assume non-adaptive power allocation. The performance of an optimal (INR) ARQ scheme in the fast-fading channel with T fading realizations in a (re)transmission round is worse than the performance in a slow-fading channel which follows the average characteristic of the fast-fading channel in each round. However, the difference between the two cases diminishes at low SNRs.

Proof. See the appendix. \square

In practice the channel does not remain constant, even at low speeds, although it is approximated to be fixed. The theorem shows that such theoretical approximations overestimate the data transmission efficiency of practical schemes. Finally, while Theorem 2 connects the fast- and slow-fading models, we close the discussions with Theorem 3 that deals with quasi-static and slow-fading models.

Theorem 3: A MIMO-ARQ protocol with a slow-fading channel leads to the same performance as the one with a quasi-static channel using random power allocation and a different average power which is found by averaging on the random power allocation.

Proof. See the appendix. \square

In words, the theorem means that, although the channel remains fixed during a packet period of the quasi-static channel, we can use random power allocation to provide the same randomness as the one which is experienced in each (re)transmission round of the slow-fading channel. However, the average transmission power should be scaled appropriately.

⁷For simplicity, the results are given mainly for an optimal (INR) ARQ protocol, while they can be extended for many other schemes as well.

B. ARQ with an unreliable feedback channel

Suppose an unreliable feedback channel where the error probability for decoding the ARQ feedback bits is p_b . With the same procedure as in, e.g., (3)-(8), the outage probability and the LT throughput are obtained as

$$\Pr(\text{Outage}) = p_b \sum_{m=1}^M (1 - p_b)^{m-1} \Pr(C_{(m)} < R_{(m)}) + (1 - p_b)^M \Pr(C_{(M+1)} < R_{(M+1)}), \quad (33)$$

$$\eta = \frac{1}{\alpha_M} \left(1 - \left\{ p_b \sum_{m=1}^M (1 - p_b)^{m-1} \Pr(C_{(m)} < R_{(m)}) + (1 - p_b)^M \Pr(C_{(M+1)} < R_{(M+1)}) \right\} \right),$$

$$\begin{aligned} \alpha_M &= \sum_{m=1}^M \frac{(1 - p_b)p_b^{m-1}}{R_{(m)}} + \frac{p_b^M}{R_{(M+1)}} \\ &+ (1 - p_b)^{M-1} (2p_b - 1) \left(\frac{1}{R_{(M)}} - \frac{1}{R_{(M+1)}} \right) \Pr(C_{(M)} < R_{(M)}) \\ &+ (2p_b - 1) \sum_{m=1}^{M-1} \left((1 - p_b)^{m-1} \Pr(C_{(m)} < R_{(m)}) \times \right. \\ &\quad \left. \left(\frac{1}{R_{(m)}} - \left\{ \sum_{j=m+1}^M \frac{(1 - p_b)p_b^{j-(m+1)}}{R_{(j)}} + \frac{p_b^{M-m}}{R_{(M+1)}} \right\} \right) \right). \end{aligned} \quad (34)$$

Here, to derive (33)-(34) we have used the fact that the data is correctly decoded at the end of the m -th round (and not before) if 1) all previous feedback bits have been correctly decoded by the transmitter (with probability $(1 - p_b)^{m-1}$), 2) the receiver has not decoded the data before, i.e., $C_{(n)} < R_{(n)}, \forall n < m$, and 3) (re)transmitting the data in the m -th slot, the receiver can decode the codeword, i.e., $C_{(m)} \geq R_{(m)}$. Also, the data (re)transmission is, either successfully or not, stopped at the m -th round with probability

$$\begin{aligned} \Pr(S_m) &= \sum_{n=1}^m (1 - p_b)^n p_b^{m-n} \Pr(\bar{A}_1, \dots, \bar{A}_{n-1}, A_n) \\ &+ \Pr(\bar{A}_1, \dots, \bar{A}_m) p_b (1 - p_b)^{m-1}, m = 1, \dots, M \\ \Pr(S_{M+1}) &= \sum_{n=1}^{M+1} (1 - p_b)^n p_b^{M+1-n} \Pr(\bar{A}_1, \dots, \bar{A}_{n-1}, A_n) \\ &+ \Pr(\bar{A}_1, \dots, \bar{A}_{M+1}) (1 - p_b)^M, \end{aligned} \quad (35)$$

and the number of channel uses is $l_{(m)}$ in that case⁸. Thus, as shown in, e.g., (33)-(34), the main conclusions of the paper are valid for noisy ARQ schemes as well, because for a given set of rates and powers the system performance is only dependent to the $C_{(m)}$'s. Finally, Fig. 10 shows the outage probability for different feedback and forward channel conditions. Better system performance and less robustness to feedback channel noise is observed when the fading channel temporal variability increases. Also, although not seen in the figure, the sensitivity to feedback channel noise increases with the number of transmit/receive antennas.

⁸The other metrics are obtained with the same procedure as in (33)-(34).

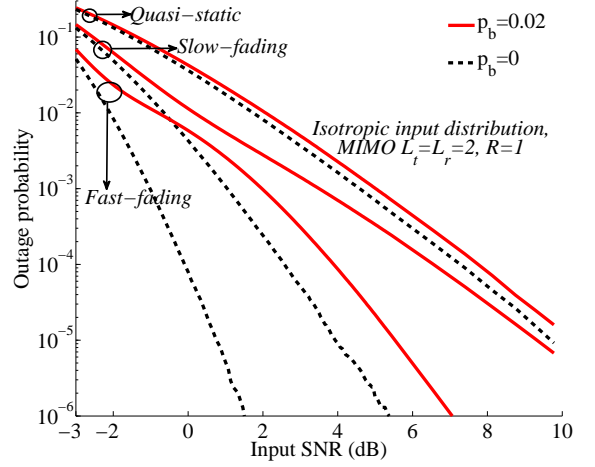


Figure 10. Outage probability vs input SNR. MIMO channel $L_t = L_r = 2$, $M = 1$, fixed-length INR ARQ protocol, $R = 1$, $\mathbf{K}_m = \frac{\phi}{T} \mathbf{I}_{L_t}$, $m = 1, 2$. For the fast-fading scenario, the results are obtained with $T = 2$ different channel realizations in each sub-codeword transmission.

VIII. CONCLUSION

We showed that the data transmission efficiency of the MIMO-ARQ protocols can be computed as a function of parameters which are scheme-dependent, and not metric-dependent. Then, the results were used to study various aspects of the MIMO-ARQ in different forward/feedback channel models. We showed that the empirical pdf of the channel SNR may be different from the true one and no ARQ protocol reaches the channel ergodic capacity, if the data is transmitted in a bursty fashion and the channel is quasi-static. The performance of ARQ protocols can be sorted based on their AMIs and MIMO-ARQ can be mapped to an equivalent SISO-ARQ setup. Better system performance and less robustness to feedback channel noise is observed when the fading channel temporal variability increases. Also, the performance of MIMO-ARQ protocols with different temporal fading variations can be mapped to each other, under certain conditions.

Along with studying the effect of feedback delay, the following points would be interesting extensions of the paper; 1) data transmission with partial CSIT is still an open-problem for large-scale MIMO. Therefore, further analysis on large-scale MIMO-ARQ is a topic which is expected to result in outstanding outcomes. 2) Many optimization problems of the MIMO-ARQ setup are nonconvex problems. Thus, linearization techniques and AMIs approximations can be used to make the problem convex. Here, the results of, e.g., [60], [61] can be of great help. 3) Compensating the effect of nonlinear power amplifiers in MIMO-ARQ is an attractive problem for which the results of [48], [62] can be supportive. 4) Finally, we showed the variable-length coding to be an efficient technique improving the performance of the MIMO-ARQ setups at high SNRs/number of antennas. Therefore, DMDT analysis in the presence of variable-length coding ARQ techniques is an interesting issue which, to the best of authors' knowledge, has not been studied yet.

IX. APPENDIX

A. Proof of Theorem 1

Defining $F(x) = \Pr(C_{(m)} \leq x)$, $C_{(m)} = \log |\mathbf{I}_{L_r} + \mathbf{H}\mathbf{K}\mathbf{H}^h|$, $\forall m$, as the cdf of the AMIs in an optimal, e.g., INR or the Alamouti STC, ARQ model, the LT throughput (7) can be rewritten as

$$\eta = \frac{1 - F(R_{(M+1)})}{\sum_{m=1}^{M+1} \frac{1}{R_{(m)}} (F(R_{(m-1)}) - F(R_{(m)})) + \frac{F(R_{(M+1)})}{R_{(M+1)}}}.$$

Therefore, setting $\frac{\partial \eta}{\partial R_{(m)}} = 0$, we have

$$\begin{aligned} \frac{F(R_{(m)}) - F(R_{(m-1)})}{R_{(m)}^2} - \frac{f(R_{(m)})}{R_{(m)}} + \frac{f(R_{(m)})}{R_{(m+1)}} &= 0 \\ \xrightarrow{(y)} \frac{-R_{(m-1)}}{R_{(m)}^2} + \frac{1}{R_{(m+1)}} &= 0 \Rightarrow R_{(m)}^2 = R_{(m-1)}R_{(m+1)}. \end{aligned} \quad (36)$$

Thus, for high m 's, the optimal equivalent data rates can be approximated by a geometric sequence. Note that (y) follows from the approximation $(F(R_{(m)}) - F(R_{(m-1)})) \rightarrow f(R_{(m)})(R_{(m)} - R_{(m-1)})$ when m increases and the equivalent data rates $R_{(m)} = \frac{Q}{\sum_{i=1}^m l_i}$ decrease. Then, from (36), the equivalent rates $R_{(m)}$ can be used to approximate the optimal, in terms of LT throughput, values of l_m and the transmission rate in each round⁹.

B. Proof of Theorem 2

The proof follows from the following (in)equalities

$$\begin{aligned} C_{(m)}^{\text{INR, fast-fading, H}(t)} &= \frac{1}{mT} \sum_{n=1}^m \sum_{t=(n-1)T+1}^{nT} \log |\mathbf{I}_{L_r} + \mathbf{H}(t)\mathbf{K}\mathbf{H}(t)^h| \\ &\stackrel{(b)}{=} \frac{1}{mT} \sum_{n=1}^m \sum_{t=(n-1)T+1}^{nT} \log(1 + \Upsilon(t)) \\ &\stackrel{(c)}{\leq} \frac{1}{m} \sum_{n=1}^m \log(1 + V_n) = C_{(m)}^{\text{INR, slow-fading, } V_n}, \\ V_n &= \frac{1}{T} \sum_{t=(n-1)T+1}^{nT} \Upsilon(t). \end{aligned} \quad (37)$$

Here, (b) is obtained by defining the random variable $\Upsilon(t)$, $F_{\Upsilon(t)}(x) = F_{|\mathbf{I}_{L_r} + \mathbf{H}(t)\mathbf{K}\mathbf{H}(t)^h| - 1}(x)$, $\forall x$, (the same as in Example 3 on equivalent SISO- and MIMO-ARQ models) and (c) comes from the Jensen's inequality [57] and the concavity of the function $\log(1+x)$. Finally, $C_{(m)}^{\text{INR, slow-fading, } V_n}$ is the AMI in a slow-fading channel in which the channel quality follows the distribution f_V , $V = \frac{1}{T} \sum_{t=1}^T \Upsilon(t)$, i.e., the average of the fast-fading channel quality variations during a (re)transmission round. Then, according to Remark 1 and (37), better performance is observed in the *averaged* model. However, as $\log(1+x) \rightarrow x$ for small x , the inequality in (37) changes to equality and, consequently, the low-SNR system performance is the same in two models.

⁹For discussions on variable-length coding ARQ in slow-fading condition, see [55].

C. Proof of Theorem 3

For simplicity, we prove the theorem for an optimal MIMO-ARQ protocol with non-adaptive power allocation in the slow-fading channel case; the only difference between the two considered cases is in their AMIs which are obtained by

$$C_{(m)}^{\text{slow-fading, non-adaptive power}} = \frac{1}{m} \sum_{n=1}^m \log |\mathbf{I}_{L_r} + \mathbf{H}_n \mathbf{K} \mathbf{H}_n^h| \quad (38)$$

$$C_{(m)}^{\text{quasi-static, random power}} = \frac{1}{m} \sum_{n=1}^m \log |\mathbf{I}_{L_r} + \mathbf{H} \mathbf{K}_n \mathbf{H}^h|. \quad (39)$$

Thus, the performance of the two cases is the same if the random powers (\mathbf{K}_n in (39)) are selected via a specific distribution F_ε such that $F_{|\mathbf{I}_{L_r} + \mathbf{H}_n \mathbf{K} \mathbf{H}_n^h|}(x) = F_{|\mathbf{I}_{L_r} + \mathbf{H} \mathbf{K}_n \mathbf{H}^h|}(x)$, $\forall x$, i.e., the same randomness is experienced in the channel quality of the two cases. Here, the only point is that, representing the power term associated with \mathbf{K}_n in (39) by ϕ_n , the average power in the second case is found as

$$\Phi^{\text{quasi-static, random power}} = E_\varepsilon \left\{ \frac{\sum_{m=1}^{M+1} \phi_m \Pr(C_{(m-1)} < R_{(m-1)})}{\sum_{m=1}^{M+1} \Pr(C_{(m-1)} < R_{(m-1)})} \right\} \quad (40)$$

which is different from the transmission power in the first case, i.e., $\Phi^{\text{slow-fading, non-adaptive power}} = \text{Tr}(\mathbf{K})$. In (40), E_ε denotes the expectation on F_ε .

REFERENCES

- [1] 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.
- [2] T. Holliday, A. J. Goldsmith, and H. V. Poor, "Joint source and channel coding for MIMO systems: Is it better to be robust or quick?" *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1393–1405, April 2008.
- [3] A. Chuang, A. Guillén i Fábregas, 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, Sept. 2008.
- [4] K. D. Nguyen, L. K. Rasmussen, A. Guillén i Fábregas, and N. Letzepis, "Rate-diversity-delay tradeoff for ARQ systems over MIMO block-fading channels," in *AusCTW*, Feb. 2009, pp. 116–121.
- [5] H. Liu, L. Razoumov, N. Mandayam, and P. 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.
- [6] K. D. Nguyen, L. K. Rasmussen, A. Guillén i Fábregas, and N. Letzepis, "MIMO ARQ with multibit feedback: Outage analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 765–779, Feb. 2012.
- [7] M. Gidlund, "Design and performance of space-time block coded hybrid ARQ schemes for multiple antenna transmission," in *ISPACS*, Feb. 2009, pp. 1–4.
- [8] A. L. Toledo and X. Wang, "TCP performance over wireless MIMO channels with ARQ and packet combining," *IEEE Trans. Mobile Comput.*, vol. 5, no. 3, pp. 208–223, March 2006.
- [9] J. Lee, H. L. Lou, D. Toupakaris, E. W. Jang, and J. M. Cioffi, "Transceiver design for MIMO wireless systems incorporating hybrid ARQ," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 32–40, Jan. 2009.
- [10] C. Bai, W. A. Krzymień, and I. J. Fair, "Hybrid-ARQ for layered space time MIMO systems with channel state information only at the receiver," *IET Commun.*, vol. 4, no. 14, pp. 1765–1773, 2010.
- [11] D. Toupakaris, J. Lee, A. Matache, and H. L. Lou, "Performance of MIMO HARQ under receiver complexity constraints," in *GLOBECOM*, Dec. 2008, pp. 1–5.
- [12] E. W. Jang, J. Lee, H. L. Lou, and J. M. Cioffi, "On the combining schemes for MIMO systems with hybrid ARQ," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 836–842, Feb. 2009.

- [13] G. Aniba and S. Aïssa, "Cross-layer designed adaptive modulation algorithm with packet combining and truncated ARQ over MIMO Nakagami fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1026–1031, April 2011.
- [14] Y. Xie and A. J. Goldsmith, "Diversity-multiplexing-delay tradeoffs in MIMO multihop networks with ARQ," in *ISIT*, June 2010, pp. 2208–2212.
- [15] A. Khina, Y. Kochman, U. Erez, and G. W. Wornell, "Incremental coding over MIMO channels," in *ITW*, Oct. 2011, pp. 400–404.
- [16] C. Shen and M. P. Fitz, "Hybrid ARQ in multiple-antenna slow fading channels: Performance limits and optimal linear dispersion code design," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5863–5883, Sept. 2011.
- [17] H. Zhao, H. Yan, and W. Wang, "An adaptive HARQ scheme based on quasi-orthogonal space-time coding for D-STTD system," in *GLOBECOM*, Nov. 2007, pp. 4000–4004.
- [18] C. Li and X. Wang, "Throughput analysis for parallel ARQ over correlated MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1322–1332, Sept. 2007.
- [19] P. Hesami and J. N. Laneman, "Low-complexity incremental use of multiple transmitters in wireless communication systems," in *Allerton*, Sept. 2011, pp. 1613–1618.
- [20] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [21] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.
- [22] 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, July 2002.
- [23] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, April 2003.
- [24] S. Nanda, R. Walton, J. Ketchum, M. Wallace, and S. Howard, "A high-performance MIMO OFDM wireless LAN," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 101–109, Feb. 2005.
- [25] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 129–136, Feb. 2005.
- [26] S. Y. Park and D. J. Love, "Hybrid ARQ protocol for multi-antenna multicasting using a common feedback channel," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1530–1542, June 2011.
- [27] D. Tuninetti, "On the benefits of partial channel state information for repetition protocols in block fading channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5036–5053, Aug. 2011.
- [28] B. Makki, A. Graell i Amat, and T. Eriksson, "HARQ feedback in spectrum sharing networks," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1337–1340, Sept. 2012.
- [29] 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, July 2001.
- [30] B. Makki and T. Eriksson, "On hybrid ARQ and quantized CSI feedback schemes in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 986–997, April 2012.
- [31] C. Shen, T. Liu, and M. P. Fitz, "On the average rate performance of hybrid-ARQ in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3339–3352, Nov. 2009.
- [32] B. Makki, A. Graell i Amat, and T. Eriksson, "Power allocation in repetition time diversity hybrid automatic repeat request feedback," in *WCNC*, April 2012, pp. 2329–2334.
- [33] P. Wu and N. Jindal, "Performance of hybrid-ARQ in block-fading channels: A fixed outage probability analysis," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1129–1141, April 2010.
- [34] L. Szczecinski, S. R. Khosravirad, P. Duhamel, and M. Rahman, "Rate allocation and adaptation for incremental redundancy truncated HARQ," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2580–2590, June 2013.
- [35] H. Sun, J. H. Manton, and Z. Ding, "Progressive linear precoder optimization for MIMO packet retransmissions," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 448–456, March 2006.
- [36] K. Zheng, H. Long, L. Wang, and W. Wang, "Linear space-time precoder with hybrid ARQ transmission," in *GLOBECOM*, Nov. 2007, pp. 3543–3547.
- [37] H. Sun, Z. Shi, C. Zhao, J. H. Manton, and Z. Ding, "Progressive linear precoder optimization for MIMO packet retransmissions exploiting channel covariance information," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 818–827, May 2008.
- [38] H. Huang and Z. Ding, "Ergodic capacity maximizing MIMO ARQ precoder design based on channel mean information," in *ITAW*, Feb. 2008, pp. 58–62.
- [39] K. Zheng, H. Long, L. Wang, W. Wang, and Y. I. Kim, "Design and performance of space-time precoder with hybrid ARQ transmission," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1816–1822, May 2009.
- [40] 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.
- [41] B. Makki and T. Eriksson, "Feedback subsampling in temporally-correlated slowly-fading channels using quantized CSI," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2282–2294, June 2013.
- [42] —, "On the average rate of quasi-static fading channels with ARQ and CSI feedback," *IEEE Commun. Lett.*, vol. 14, no. 9, pp. 806–808, Sept. 2010.
- [43] S. Lin and D. Costello, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [44] M. Zorzi and R. R. Rao, "On the use of renewal theory in the analysis of ARQ protocols," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1077–1081, Sept. 1996.
- [45] V. S. Bawa, "Optimal rules for ordering uncertain prospects," *J. of Financial Economics*, vol. 2, no. 1, pp. 95–121, March 1975.
- [46] T. V. K. Chaitanya and E. G. Larsson, "Outage-optimal power allocation for hybrid ARQ with incremental redundancy," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2069–2074, July 2011.
- [47] E. Telatar, "Capacity of multi-antenna Gaussian channels," Bell Labs, Tech. Rep., June 1995.
- [48] J. Qi and S. Aïssa, "Analysis and compensation of power amplifier nonlinearity in MIMO transmit diversity systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2921–2931, July 2010.
- [49] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [50] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *Allerton*, Sept. 2011, pp. 545–550.
- [51] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.
- [52] F. Berggren, O. Queseth, J. Zander, B. Asp, C. Jönsson, P. Stenungaard, N. Z. Kviselius, B. Thorngren, U. Landmark, and J. Wessel, "Dynamic spectrum access, phase 1: scenarios and research challenges," Sept. 2004, available: <http://www.queseth.se/olav/pubs/DSARepPhase1.pdf>.
- [53] I. F. Akyildiz, W. L. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, pp. 2127–2159, May 2006.
- [54] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.
- [55] L. Szczecinski, C. Correa, and L. Ahumada, "Variable-rate transmission for incremental redundancy hybrid ARQ," in *GLOBECOM*, 2010, pp. 1–5.
- [56] J. Tsao, D. Porat, and D. Tse, "Prediction and modeling for the time-evolving ultra-wideband channel," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 340–356, Oct. 2007.
- [57] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley Interscience, 1992.
- [58] A. El Gamal and Y.-H. Kim, "Lecture notes on network information theory," 2010, [Online]. Available at <http://arxiv.org/abs/1001.3404v4/>.
- [59] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, July 1999.
- [60] O. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.
- [61] Q. T. Zhang, X. W. Cui, and X. M. Li, "Very tight capacity bounds for MIMO-correlated Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 681–688, March 2005.
- [62] J. Qi and S. Aïssa, "On the power amplifier nonlinearity in MIMO transmit beamforming systems," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 876–887, March 2012.



Behrooz Makki was born in Tehran, Iran. He received the B.Sc. degree in Electrical Engineering from Sharif University of Technology, Tehran, Iran, and the M.Sc. degree in Bioelectric Engineering from Amirkabir University of Technology, Tehran, Iran, respectively. Behrooz received his PhD degree in Communication Engineering from Chalmers University of Technology, Gothenburg, Sweden. Since 2013, he is working as a Postdoc at Chalmers University. His current research interests include partial channel state information (CSI) feedback, hybrid

automatic repeat request, Green communication, correlated fading channels, spectrum sharing, relay and CoMP networks.



Thomas Eriksson received the M.S. degree in Electrical Engineering in 1990, and the Ph.D. degree in Information Theory in 1996, both at Chalmers University of Technology, Gothenburg, Sweden. After his degree, he worked for a year at AT&T Labs Research in Murray Hill, NJ, USA, and for a year at Ericsson AB in Kista, Stockholm, before he joined the department of Signals and Systems at Chalmers. He is now a full professor in the Department of Signals and Systems, where he is also vice head of department, responsible for the undergraduate

education.

A main research interest is so called “Dirty RF” processing, where imperfections in communication hardware are analyzed and compensated. To this area belongs, e.g., signal design for efficient amplifier architectures, digital predistortion of nonlinear amplifiers, compensation of I/Q imbalance, design of efficient MIMO transmitters, analysis and compensation of phase noise in oscillators etc.