



CHALMERS

Chalmers Publication Library

Green Communication via Power-optimized HARQ Protocols

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 Vehicular Technology (ISSN: 0018-9545)

Citation for the published paper:

Makki, B. ; Graell i Amat, A. ; Eriksson, T. (2014) "Green Communication via Power-optimized HARQ Protocols". IEEE Transactions on Vehicular Technology, vol. 63(1), pp. 161-177.

<http://dx.doi.org/10.1109/TVT.2013.2274287>

Downloaded from: <http://publications.lib.chalmers.se/publication/179853>

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)

Green Communication via Power-optimized HARQ Protocols

Behrooz Makki, Alexandre Graell i Amat, *Senior Member, IEEE* and Thomas Eriksson

Abstract—Recently, efficient use of energy has become an essential research topic for green communication. This paper studies the effect of optimal power controllers on the performance of delay-sensitive communication setups utilizing hybrid automatic repeat request (HARQ). The results are obtained for repetition time diversity (RTD) and incremental redundancy (INR) HARQ protocols. In all cases, the optimal power allocation, minimizing the outage-limited average transmission power, is obtained under both continuous and bursting communication models. Also, we investigate the system throughput in different conditions. The results indicate that the power efficiency is increased substantially, if adaptive power allocation is utilized. For instance, assume Rayleigh-fading channel, a maximum of two (re)transmission rounds with rates $\{1, \frac{1}{2}\}$ nats-per-channel-use and an outage probability constraint 10^{-3} . Then, compared to uniform power allocation, optimal power allocation in RTD reduces the average power by 9 and 11 dB in the bursting and continuous communication models, respectively. In INR, these values are obtained to be 8 and 9 dB, respectively.

Index terms: Green communication, hybrid automatic repeat request, outage probability, adaptive power allocation, block-fading channels, feedback

I. INTRODUCTION

The first and the most important requirement in many wireless applications is that the data must be decodable at the receiver. This problem is often studied under the topic of *outage-limited* data transmission [1]–[5]; An outage happens when the data can not be decoded at the receiver. With an outage constraint, the data transmission is successful if the codewords are decodable, regardless of whether they lead to maximum system throughput or not.

Assuming delay-sensitive communication over block-fading channels, it is well-accepted that infinite power is required to achieve zero outage probability for all channel conditions [1]–[5]. This is because the channel may fall in deep fading conditions and so infinitely large powers are needed for *channel inversion*, i.e., for compensating the channel *bad* conditions. In these cases, the strict outage requirements are replaced by more relaxed probabilistic constraints [1]–[5], where “a service is acceptable as long as the data is always decoded with some probability ϵ .” Here, ϵ is a parameter representing the system outage tolerance.

Due to the fast growth of wireless networks and the data-intensive applications produced by smart phones, green communication via improving the power efficiency is becoming in-

creasingly important for wireless communication. The network data volume is expected to increase by a factor of 2 every year, associated with 16 – 20% increase of energy consumption, which contributes about 2% of global CO_2 emissions [6]. Hence, minimizing the power consumption is a very important design consideration, and power-efficient data transmission schemes must be taken into account for the wireless networks [7]. Green radio has thereby been proposed as an effective solution and is becoming the mainstream for future wireless network design [8]–[14].

From another perspective, hybrid automatic repeat request (HARQ) is a well-known approach applied in today’s wireless networks to increase the data transmission reliability and efficiency [5], [15]–[29]. The main idea behind the HARQ techniques is to reduce the data outage probability or increase the throughput by retransmitting the data that underwent *bad* channel conditions. Consequently, it is expected that joint implementation of adaptive power controllers and HARQ protocols improve the power efficiency of outage-limited communication systems.

Outage-limited power allocation is an interesting problem which was previously studied by, e.g., Caire *et. al.*, [2] under perfect channel state information (CSI) assumption. Also, [3], [4] investigated the same problem in the presence of quantized CSI. In [5], Wu and Jindal studied the outage-limited performance of HARQ protocols in block-fading channels. Here, the results were obtained with no power adaptation and under the assumption that the channel changes in each retransmission round. The outage-limited power allocation problem for the repetition time diversity (RTD) and fixed-length coding incremental redundancy (INR) HARQ protocols were investigated in [23] and [24], respectively. Finally, assuming that the channel changes in each retransmission, [29] found the optimal retransmission power for basic automatic repeat request (ARQ) protocols where it was shown that the power should be increasing in every retransmission.

Power allocation in HARQ schemes has also been considered in [15] and [18], where the power-limited throughput optimization problem of different HARQ protocols was studied. In [25], the optimal HARQ-based power allocation for minimizing the required number of retransmission rounds was obtained in a down-link wideband code division multiple access (WCDMA) system. Assuming partial CSI at the transmitter, [26] and [27] studied the power allocation problem in HARQ systems when the power is changed according to the received CSI and by a linear programming approach under a buffer cost constraint, respectively. Finally, [28] investigated the power allocation between the source and the relay in a relay channel where, while the powers are fixed in the

Behrooz Makki, Alexandre Graell i Amat and Thomas Eriksson are with Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, Email: {behrooz.makki, alexandre.graell, thomase}@chalmers.se

Alexandre Graell i Amat was supported by the Swedish Agency for Innovation Systems (VINNOVA) under the P36604-1 MAGIC project.

retransmission rounds, there is a sum power constraint on the relay and the transmitter.

This paper investigates the power allocation problem in HARQ-based systems under different outage probability constraints. Utilizing different HARQ protocols, the goal is to determine the optimal power controllers satisfying different outage probability constraints. Also, the same setup is valid for the inverse problem where for a given transmission power, the goal is to minimize the outage probability. The results are obtained for the RTD and the INR protocols, under both continuous and bursting communication models and with the average power definition given in, e.g., [2], [4], [15], [17], [18], [30]–[32].

Along with the outage probability, the long-term throughput is another metric demonstrating the system long-term performance. We obtain the long-term throughput and give comparisons between the outage-limited average power and throughput of the RTD and INR protocols.

The outage-limited power allocation problem setup of the paper has not been investigated in [5], [15], [18], [21], [22], [25]–[28]. Also, the paper is different from [29] as 1) in contrast to [29] that deals with basic ARQ, we study the HARQ protocols and 2) the results are obtained under different block-fading channel assumptions. As mentioned before, [23], [24] have also studied the outage-limited power allocation problem in HARQ protocols. However, there are fundamental differences between the definition of the average power in [23], [24] and the definition considered in our work. In particular, the definition of power used in [23], [24] corresponds to the definition of the energy (although in these papers it is referred as power), while in this paper we consider the more common definition of the average power, as in, e.g., [2], [4], [15], [17], [18], [30]–[32]. The difference in the definition of the average power makes the problem solved in this paper completely different from the one addressed in [23], [24], leading to different analytical and numerical results, as well as to different conclusions. Also, some other contributions of our work compared to [23], [24] (and also [5], [15], [18], [21], [22], [25]–[29]) are: 1) solving the problem for both continuous and bursting communication models, 2) focusing on variable-length INR HARQ, instead of a fixed-length coding scheme, and 3) studying the long-term throughput in each case.

The main idea of the paper can be explained as follows. With an outage probability constraint, the initial transmission(s) of the HARQ scheme is set to have a small power. If the channel is *bad* the data can not be decoded and is retransmitted with higher powers. On the other hand, if the channel experiences good conditions, this gambling brings high return. With this scheme, which is shown to be optimal in terms of outage-limited average transmission power, the HARQ helps us to exploit the good channel conditions properly, and decrease the average transmission power for a given outage probability constraint. Moreover, we illustrate power adaptation schemes based on reinforcement algorithms which, depending on the fading condition, can lead to performance improvement in the HARQ protocols.

Minimizing the average transmission power with an outage

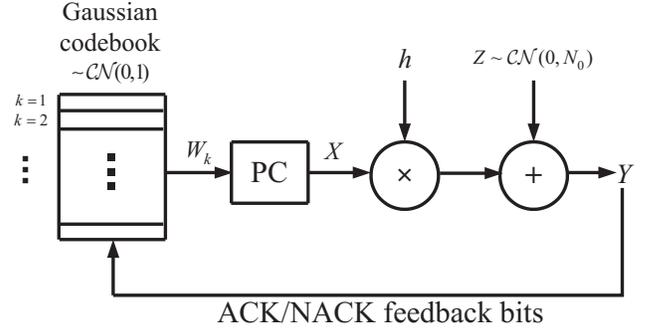


Figure 1. System model. The codewords are constructed according to the standard Gaussian codes. Therefore, we have $\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=1}^L |W_k[i]|^2 \leq 1, \forall k$, where L is the length of the codewords. Both fixed-length (RTD) HARQ and variable-length (INR) HARQ schemes are studied in Sections IV and V, respectively. The block PC represents the power controller where the codewords are rescaled based on the number of (re)transmission round. The goal is to determine the optimal power controllers under different outage probability constraints. Finally, the HARQ feedback bits are assumed to be received by the transmitter error- and delay-free.

probability constraint, our results show that: 1) for the RTD HARQ protocol, the transmission powers must be increasing in every retransmission. 2) Also, for the INR HARQ protocol, the transmission energies must increase in every retransmission. 3) For sufficiently large number of retransmissions and independent of the fading distribution, the optimal retransmission powers of the RTD and fixed-length coding INR schemes can be determined via a specific sequence of numbers converging to a geometric sequence. Note that these results are in contrast with the results in [23], [24], which is due to differences in the objective function of our work and [23], [24]¹. 4) Finally, optimal power allocation leads to considerable performance improvement in HARQ protocols. For instance, assume Rayleigh-fading channels, a maximum of two (re)transmission rounds with rates $\{1, \frac{1}{2}\}$ nats-per-channel-use (npcu)² and an outage probability constraint 10^{-3} . Then, compared to uniform power allocation, optimal power allocation in the RTD HARQ scheme reduces the average power by 9 and 11 dB in the bursting and continuous communication models, respectively. In the INR, these values are obtained to be 8 and 9dB, respectively. Also, depending on the channel condition, the power efficiency can be improved by reinforcement-based power allocation schemes, at the cost of higher implementation complexity and sensitivity to imperfect feedback signal.

II. SYSTEM MODEL

Channel assumptions: As illustrated in Fig. 1, we consider a block-fading channel where the power-limited input message X multiplied by the random variable h is summed with an i.i.d. complex Gaussian noise $Z \sim \mathcal{CN}(0, N_0)$ resulting in the output

$$Y = hX + Z. \quad (1)$$

¹In [23], [24], where the goal is to minimize the transmission energy, the optimal (re)transmission powers are found to be neither increasing nor decreasing.

²A nat is a unit of information, based on the natural logarithm [17], [18], [32], [33]. The results can be mapped to the bit unit, if the logarithmic terms are presented in base 2.

Let us define $g = |h|^2$ as the *channel gain* random variable. The channel gain remains constant for a duration of L_c channel uses (cu)³, generally determined by the channel coherence time, and then changes independently according to the fading probability density function (pdf) $f_G(g)$. Furthermore, with no loss of generality, we consider $N_0 = 1$.

It is assumed that there is perfect instantaneous knowledge about the channel gain at the receiver, which is an acceptable assumption under block-fading conditions [1]–[4], [29]–[31], [35], [36]. Also, the fading is assumed to be constant over the transmission of one packet, where a packet is defined as the transmission of a codeword along with all its possible retransmissions. This model has been considered in many other papers, e.g. [2], [4], [17], [18], [23], [24], [30]–[32], and is a good model for stationary or slow-moving users⁴. Also, it corresponds to the worst case, since no time diversity can be exploited if the channel is fixed within a packet period. Later, in Sections VII and VIII, we relax the block-fading assumption and extend the results to the case with fast-fading conditions. All results are presented in natural logarithm basis and, as each transmission experiences an AWGN channel, the results are restricted to Gaussian input distributions. Finally, the main focus is on two stop-and-wait HARQ protocols:

- 1) **Repetition time diversity.** This scheme belongs to the *diversity combining* category of HARQ protocols [15], [18], [23], [34] where the same data is repeated in the (re)transmission rounds and, in each round, the receiver performs maximum ratio combining of all received signals.
- 2) **Incremental Redundancy.** The INR belongs to the category of *code combining* protocols [15], [17], [18], [24], [34]. Here, a codeword is sent with an aggressive rate in the first round. Then, if the receiver cannot decode the initial codeword, further parity bits are sent in the next retransmission rounds and in each round the receiver decodes the data based on all received signals.

Objective function: The main goal of this paper is to design the optimal power controllers, i.e., the block PC in Fig. 1, such that a given outage probability constraint is satisfied with minimum average transmission power. The codebooks are assumed to be constructed by the standard complex Gaussian codes, which have been shown to be optimal for power-limited data transmission in AWGN channels with long codewords [33]. That is, each codeword is constructed according to $\mathcal{CN}(0, 1)$. Then, the codewords are rescaled to have different powers, determined based on the number of (re)transmission round. We assume the codewords lengths and rates to be previously designed based on, e.g., data structure and coding

³To be in harmony with the literature, e.g., [17], [18], [32], [34], channel use (cu) is considered as the unit for all temporal parameters, including the channel coherence time, the length of a fading block and the codewords length. For every given duration of a channel use, the results can be mapped to other time units, e.g., seconds, via a scaling factor.

⁴As discussed in, e.g., [2], the information theoretical results of block-fading channels match the results of actual codes for practical code lengths, e.g., ≈ 100 channel uses, and provide appropriate performance bounds for systems with smaller code lengths. Also, please see [37]–[40] for mappings between the block- and continuous-fading channels (With continuous-fading the channel changes in each channel use).

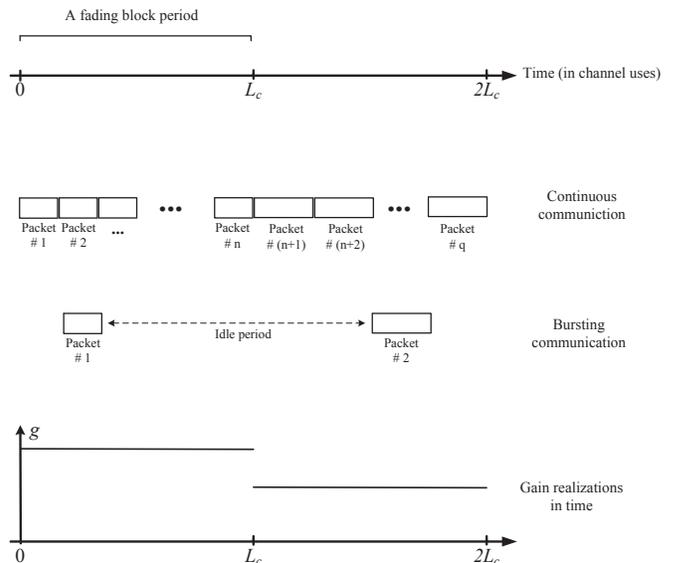


Figure 2. Description of different data communication models. With bursting (continuous) communication model, one (many) packet(s) is sent within each fading block. A packet is defined as the transmission of a codeword along with all its possible retransmission rounds.

complexity. Therefore, the transmission rates are assumed to be out of our control throughout the paper. However, the results are valid for any given codewords rates and number of retransmissions. Hence, this assumption does not affect the generality of the arguments. A number of cases in which the problem setup is applicable are explained in the following.

In general, the problem formulation used here is appropriate for fixed-rate delay-sensitive applications such as voice over internet protocol (VoIP) and fixed-rate video streaming applications [3]. The setup becomes more interesting when we remember that many practical ARQ schemes are designed to operate at fixed (re)transmission rates [34]. Also, the power control strategy resulting from this problem formulation attempts to fix the transmission rate by compensating for the channel fading and in this sense resembles the power control mechanisms used in WCDMA system standards [3], [41].

III. COMMUNICATION MODELS AND DEFINITIONS

We consider both continuous and bursting communication schemes, as illustrated in Fig. 2. Under the continuous communication model, it is assumed that there is an unlimited amount of information available at the transmitter and it is always active [17], [18], [30]. In this way, multiple packets, each packet containing multiple HARQ rounds, are transmitted within one fading block of length L_c . When the channel is good, many packets can be sent within the fading block, while only few can be transmitted within the same period for bad channels.

Under the bursting communication model, on the other hand, it is assumed that there is a long idle period between the packet transmissions. Therefore, while the HARQ retransmission rounds of each packet experience the same gain realization, the channel changes independently from one packet to another. To be more clear, all the L_c channel uses of a fading block are utilized in the continuous communication model.

This is because data is continuously transmitted, regardless of whether it is decoded or not. In the bursting communication model, on the other hand, only one packet is sent within each block that, depending on the channel conditions, can be decoded in different (re)transmission rounds. Therefore, the number of channel uses in each block is different. In the following, the long-term throughput and the long-term average transmission power are defined and derived for the continuous communication model. Then, all arguments about performance analysis in the bursting communication model are presented in Section VI. Section VIII presents the simulation results for both cases.

A. Long-term throughput

The long-term throughput (in nats-per-channel-use) is defined as

$$\eta \doteq \lim_{k \rightarrow \infty} \frac{Q^{(k)}}{\tau^{(k)}} = \frac{E\{\tilde{Q}\}}{E\{\tilde{\tau}\}} \quad (2)$$

where $Q^{(k)}$ and $\tau^{(k)}$ are the total successfully decoded information nats and the total number of channel uses at the end of k data transmission times. Then, \tilde{Q} and $\tilde{\tau}$ are the number of successfully decoded information nats and channel uses in each block, respectively, and $E\{\cdot\}$ is the expectation operator [19].

Assuming continuous communication, the long-term throughput can be calculated as follows. Let $R(g)$ be the instantaneous data rate of the HARQ approach for a given gain realization g . The total number of information nats that can be decoded in each state is obtained by $Q(g) = L_c R(g)$. Consequently, the long-term throughput is simplified to

$$\eta = \frac{E\{L_c R(g)\}}{L_c} = E\{R(g)\} = \bar{R} \quad (3)$$

where \bar{R} is the channel average rate [17], [18], [30], [34], [42].

B. Long-term average transmission power

The long-term average transmission power is defined as

$$\varphi = \lim_{k \rightarrow \infty} \frac{\xi^{(k)}}{\tau^{(k)}} = \frac{E\{\tilde{\xi}\}}{E\{\tilde{\tau}\}} \quad (4)$$

where $\xi^{(k)}$ and $\tilde{\xi}$ are the total energy consumed in k data transmission times and within a block period, respectively [2], [4], [15], [17], [18], [30], [31]. We denote by $P(g)$ the transmission power random variable of an HARQ scheme for a channel gain realization g . Then, the average transmission power for the continuous communication model is obtained by

$$\varphi = \frac{E\{L_c P(g)\}}{L_c} = E\{P(g)\} = \bar{P}. \quad (5)$$

In the bursting communication model, on the other hand, the denominator of, e.g., (2) is not constant. Hence, the long-term throughput and the average transmission power should be directly calculated based on (2) and (4), respectively (please see Section VI for more details).

IV. RTD HARQ PROTOCOLS WITH AN OUTAGE PROBABILITY CONSTRAINT

Assuming a continuous communication model, this section first studies the optimal power controllers in RTD HARQ-based schemes constrained to different outage probability constraints. Later, the long-term throughput is investigated.

We consider a maximum of M data retransmission rounds, i.e., each codeword is (re)transmitted a maximum of $M + 1$ times, determined by the system delay requirements. Implementing the *Gaussian* codes, the original RTD-based codeword is constructed by encoding Q information nats into a codeword of length L , $L \ll L_c$, and rate $R = \frac{Q}{L}$. The codeword is rescaled in each (re)transmission round to have power P_m , $m = 1, \dots, M + 1$. Representing the codeword transmitted at the m -th (re)transmission round by $\{X_m[i], i = 1, \dots, L\}$ we have

$$P_m = \frac{1}{L} \sum_{i=1}^L |X_m[i]|^2. \quad (6)$$

The (re)transmission continues until an acknowledgement (ACK) feedback bit is received by the transmitter or the maximum permitted retransmission rounds are used. At the end of the m -th (re)transmission round, the receiver performs maximum ratio combining of the m received signals. This process effectively increases the received signal-to-noise ratio (SNR) to $g \sum_{n=1}^m P_n$ and reduces the data rate to $\frac{R}{m}$. Define $I_m \doteq \frac{1}{m} \log(1 + g \sum_{n=1}^m P_n)$ as the instantaneous mutual information and $\Upsilon_m \doteq \{I_m \geq \frac{R}{m}\}$ as the event that the instantaneous mutual information exceeds the equivalent transmission rate at the m -th (re)transmission round. The data is successfully decoded at the m -th retransmission round if 1) the receiver has not decoded the message in the previous (re)transmissions, i.e., $I_n < \frac{R}{n} \forall n < m$, and 2) using the m -th retransmission round it can decode the information, that is, $I_m \geq \frac{R}{m}$. Then, as $\Upsilon_m \subset \Upsilon_n, n \leq m$, we have

$$\begin{aligned} \Pr\{S_m\} &= \Pr\{\hat{\Upsilon}_1, \dots, \hat{\Upsilon}_{m-1}, \Upsilon_m\} \\ &= \Pr\left\{\log(1 + g \sum_{n=1}^{m-1} P_n) < R \leq \log(1 + g \sum_{n=1}^m P_n)\right\} \\ &= F_G\left(\frac{e^R - 1}{\sum_{n=1}^{m-1} P_n}\right) - F_G\left(\frac{e^R - 1}{\sum_{n=1}^m P_n}\right), \\ \Pr\{\hat{S}_{M+1}\} &= \Pr\left\{\log(1 + g \sum_{n=1}^{M+1} P_n) < R\right\} \\ &= F_G\left(\frac{e^R - 1}{\sum_{n=1}^{M+1} P_n}\right). \end{aligned} \quad (7)$$

Here, S_m is the event that the data is decoded at the end of the m -th round, \hat{U} denotes the complement of the event U and F_G is the channel gain cumulative distribution function (cdf). Consequently, $\Pr\{\hat{S}_{M+1}\}$ represents the probability that the data is lost while all retransmission rounds have been used.

On the other hand, the average transmission power at the end of the m -th (re)transmission round is

$$P^{(m)} = \frac{1}{mL} \sum_{n=1}^m \sum_{i=1}^L |X_n[i]|^2 = \frac{1}{m} \sum_{n=1}^m P_n. \quad (8)$$

Also, independent of the message decoding status, the average transmission power over the packet is $P^{(M+1)} = \frac{1}{M+1} \sum_{n=1}^{M+1} P_n$ if all possible retransmission rounds are used.

Therefore, assuming a continuous communication model, the overall long-term average transmission power is determined by

$$\begin{aligned} \bar{P} &= \sum_{m=1}^{M+1} P^{(m)} \Pr\{S_m\} + P^{(M+1)} \Pr\{\hat{S}_{M+1}\} = \\ &= \sum_{m=1}^{M+1} \left(\frac{1}{m} \sum_{n=1}^m P_n \right) \Pr\{S_m\} + \left(\frac{1}{M+1} \sum_{n=1}^{M+1} P_n \right) \Pr\{\hat{S}_{M+1}\} \\ &= \sum_{m=1}^{M+1} \left(\frac{1}{m} \sum_{n=1}^m P_n \right) \times \\ &\quad \Pr \left\{ \log(1 + g \sum_{n=1}^{m-1} P_n) < R \leq \log(1 + g \sum_{n=1}^m P_n) \right\} \\ &+ \left(\frac{1}{M+1} \sum_{k=1}^{M+1} P_k \right) \Pr \left\{ \log(1 + g \sum_{n=1}^{M+1} P_n) < R \right\} \\ &= \sum_{m=1}^{M+1} \left(\frac{1}{m} \sum_{n=1}^m P_n \right) \left(F_G \left(\frac{e^R - 1}{\sum_{n=1}^{m-1} P_n} \right) - F_G \left(\frac{e^R - 1}{\sum_{n=1}^m P_n} \right) \right) \\ &+ \left(\frac{1}{M+1} \sum_{k=1}^{M+1} P_k \right) F_G \left(\frac{e^R - 1}{\sum_{n=1}^{M+1} P_n} \right). \end{aligned} \quad (9)$$

For a given channel gain realization g , the data can not be decoded if and only if $R > \log(1 + g \sum_{n=1}^{M+1} P_n)$. Therefore, the outage probability constraint $\Pr\{\text{outage}\} \leq \epsilon$ can be represented as

$$\Pr \left\{ \log(1 + g \sum_{n=1}^{M+1} P_n) < R \right\} \leq \epsilon \Rightarrow \sum_{n=1}^{M+1} P_n \geq \frac{e^R - 1}{F_G^{-1}(\epsilon)} \quad (10)$$

where F_G^{-1} is the inverse function of the channel gain cdf. Intuitively, higher transmission power or more retransmission rounds are needed when the outage probability constraint gets harder, i.e., ϵ becomes smaller.

A. Power allocation

Communication systems may have different power allocation capabilities. For instance, due to, e.g., hardware or complexity limitations, there are cases where, independently of the channel conditions, the data must be transmitted at a fixed power P , i.e., $P_m = P \forall m$, which is normally called *short-term* power allocation or data transmission with a peak power constraint [1]–[4], [15]–[18], [30], [31]. In this case, the minimum transmission power satisfying an RTD-based outage probability constraint (10) is found as

$$P = \frac{e^R - 1}{(M+1)F_G^{-1}(\epsilon)}. \quad (11)$$

Under the more relaxed long-term (battery-limited) power allocation scenario, the transmitter can adapt the power in each retransmission round. In this case, using (9), the power allocation problem can be stated as

$$\begin{aligned} \min_{P_1, \dots, P_{M+1}} \quad & \bar{P} \\ \text{subject to} \quad & \sum_{n=1}^{M+1} P_n \geq \frac{e^R - 1}{F_G^{-1}(\epsilon)} \end{aligned} \quad (12)$$

which can be solved numerically. The following theorem shows that, independent of the number of retransmission rounds, the initial transmission rate and the fading distribution, the transmission powers must be increasing in every RTD-based retransmission round.

Theorem 1: In RTD HARQ protocols with an outage probability constraint, the optimal transmission powers, minimizing the average transmission power, must be increasing in every retransmission.

Proof: The two successive power terms P_k and P_{k+1} in (10) are interchangeable⁵. However, using (9), the contributions of the power terms P_{k+1} and P_k on the average transmission power are respectively found as

$$\begin{aligned} c(P_{k+1}) &= P_{k+1} a_{k+1}, \\ a_{k+1} &= \sum_{j=k+1}^{M+1} \frac{1}{j} \left(F_G \left(\frac{e^R - 1}{\sum_{n=1}^{j-1} P_n} \right) - F_G \left(\frac{e^R - 1}{\sum_{n=1}^j P_n} \right) \right) \\ &\quad + \frac{1}{M+1} F_G \left(\frac{e^R - 1}{\sum_{n=1}^{M+1} P_n} \right) \end{aligned} \quad (13)$$

and

$$\begin{aligned} c(P_k) &= P_k a_k \\ &= P_k \left(a_{k+1} + \frac{1}{k} \left(F_G \left(\frac{e^R - 1}{\sum_{n=1}^{k-1} P_n} \right) - F_G \left(\frac{e^R - 1}{\sum_{n=1}^k P_n} \right) \right) \right) > P_k a_{k+1} \end{aligned} \quad (14)$$

i.e., $a_k \geq a_{k+1}$. Here, $c(P_k)$ is the contribution of the power term P_k on the average transmission power, i.e., $\bar{P} = \sum_{k=1}^{M+1} c(P_k)$. Then, using (13) and (14), it is obvious that, in the optimal case, we have $P_k \leq P_{k+1}$. This is particularly because with the same powers $P_k = P_{k+1}$ (or $P_k \geq P_{k+1}$) the k -th power term has more contribution on the average transmission power than the $(k+1)$ -th term. Therefore, in order to have minimum average transmission power, the powers should preferably be given to the last retransmission rounds. ■

There is an interesting intuition behind Theorem 1 (see also Fig. 3); implementing RTD HARQ, the initial transmission is set to have a small power. If the channel is *bad* the data can not be decoded and is retransmitted with higher powers. On the other hand, if the channel is *good*, this gambling brings high return. In other words, the HARQ helps us to exploit the good channel conditions properly. Moreover, the outage probability constraint, i.e., (10), does not imply any preference between the order of the power terms. Finally, it is worth noting that in [29] it was shown that the same conclusion is valid for basic ARQ schemes where, in each retransmission round, the data is decoded based on the signal received in that time slot, regardless of the previously received signals. However, the conclusion of the theorem is not valid in the problem formulation of [23], where the transmission energy is minimized in outage-limited conditions.

Remark: Using $\Pr\{\Upsilon_m\} = \Pr\{\log(1 + g \sum_{n=1}^m P_n) \geq R\}$, $\Pr\{\hat{\Upsilon}_1, \dots, \hat{\Upsilon}_{m-1}, \Upsilon_m\} = \Pr\{\Upsilon_m\} - \Pr\{\Upsilon_{m-1}\}$ and some manipulations in (9), it can be written

$$\bar{P} = \sum_{m=1}^M \left(\frac{\sum_{n=1}^m P_n - m P_{m+1}}{m(m+1)} \right) \Pr\{\Upsilon_m\} + \frac{1}{M+1} \sum_{m=1}^{M+1} P_m.$$

According to Theorem 1, we have $\sum_{n=1}^m P_n - m P_{m+1} \leq 0$. Moreover, using the exponential Chebyshev's inequality, $\Pr\{X \geq x\} \leq e^{-tx} E(e^{tX}), \forall t > 0$ [43], with $t = 1$, the probability $\Pr\{\Upsilon_m\}$ can be upper bounded by $\Pr\{\Upsilon_m\} = \Pr\{\log(1 + g \sum_{n=1}^m P_n) \geq R\} \leq e^{-R} (1 + \lambda \sum_{n=1}^m P_n)$ where

⁵As seen in the following, power allocation strategies affect the long-term throughput. However, this point is not important in outage-limited data transmission scenario, as the throughput is not an objective function in this case.

$\lambda \doteq E\{G\}$. In this way, (10) can be used to bound the average power by

$$\bar{P} \geq e^{-R} \sum_{m=1}^M \left(\frac{\sum_{n=1}^m P_n - mP_{m+1}}{m(m+1)} \right) \left(1 + \lambda \sum_{n=1}^m P_n \right) + \frac{e^R - 1}{(M+1)F_G^{-1}(\epsilon)}.$$

Due to the Chebyshev's inequality, the bound is not tight, particularly at low rates. However, it leads to the following interesting conclusions: 1) the average power scales with $F_G^{-1}(\epsilon)$ at least inversely linear and 2) with high initial rates R , where the first term of the bound vanishes, the outage-limited average power scales with R exponentially. Here, it is interesting to note that, in general, the outage probability is a nonlinear function of the retransmission rates/powers and the average transmission power.

A simple power allocation algorithm: Generally, (12) is not a convex optimization problem. Thus, although implementable, gradient-based algorithms are not efficient in this case. To tackle this problem, we propose an iterative algorithm, stated in Algorithm 1, to solve (12). The proposed algorithm has been shown to be efficient in complex optimization problems dealing with local minima issues [32], [44]. However, since the problem is nonconvex, we may need to run the algorithm for several iterations/initial settings, when the number of optimization parameters, i.e., P_m 's, increases.

Algorithm 1 Power allocation optimization

- I. For a given outage probability constraint $\Pr\{\text{outage}\} \leq \epsilon$, initial transmission rate R and the fading pdf f_G , consider J , e.g. $J = 20$, randomly generated vectors $\Lambda^j = [P_1^j, P_2^j, \dots, P_M^j]$ such that $P_m^j \geq 0$.
 - II. For each vector, do the following procedures
 - 1) Determine the last retransmission power P_{M+1}^j according to (10). If $P_{M+1}^j < 0$, eliminate the j -th vector.
 - 2) Determine the average power \bar{P}^j based on (9).
 - III. Find the vector which results in the lowest average transmission power, i.e., Λ^i where $\bar{P}^i \leq \bar{P}^j, \forall j = 1, \dots, J$.
 - IV. $\Lambda^1 \leftarrow \Lambda^i$.
 - V. Generate $b \ll J$, e.g., $b = 5$, vectors $\Lambda^{j,\text{new}}, j = 1, \dots, b$ around Λ^1 . These vectors should also satisfy the constraints introduced in I.
 - VI. $\Lambda^{j+1} \leftarrow \Lambda^{j,\text{new}}, j = 1, \dots, b$.
 - VII. Regenerate the remaining vectors $\Lambda^j, j = b+2, \dots, J$ randomly such that $\Lambda^j = [P_1^j, P_2^j, \dots, P_M^j]$ and $P_m^j \geq 0$.
 - VIII. Go to II and continue until convergence.
-

B. Long-term throughput

Provided that the data is decoded at the end of the m -th (re)transmission round, the transmission rate is $\frac{R}{m}$. Therefore, using (7), the system average rate (or equivalently, throughput) is found as

$$\eta = \sum_{m=1}^{M+1} \frac{R}{m} \Pr\{S_m\} = \sum_{m=1}^{M+1} \frac{R}{m} \left(F_G\left(\frac{e^R - 1}{\sum_{n=1}^{m-1} P_n}\right) - F_G\left(\frac{e^R - 1}{\sum_{n=1}^m P_n}\right) \right). \quad (15)$$

The transmission parameters are determined based on (12). Finally, implementing short-term power allocation, i.e., $P_m = P \forall m$, (15) is simplified to

$$\eta = \sum_{m=1}^{M+1} \frac{R}{m} \left(F_G\left(\frac{e^R - 1}{(m-1)P}\right) - F_G\left(\frac{e^R - 1}{mP}\right) \right). \quad (16)$$

V. INR HARQ PROTOCOLS WITH AN OUTAGE PROBABILITY CONSTRAINT

Assuming the continuous communication model, this section investigates the performance of the INR HARQ protocol in the presence of an outage probability constraint. Performance analysis for the bursting communication model is presented in the Section VI.

Considering a maximum of $M+1$ INR-based HARQ rounds, Q information nats are encoded into a *mother* codeword of length $l^{(M+1)}$. Then, the codeword is punctured into $M+1$ codewords with powers P_m and strictly decreasing rates

$$R^{(m)} = \frac{Q}{\sum_{j=1}^m l_j}, \quad m = 1, \dots, M+1. \quad (17)$$

Here, l_m and $R^{(m)}$ are the channel uses and the *equivalent* transmission rate in the m -th time slot, respectively. Moreover, $l^{(m)} = \sum_{k=1}^m l_k$ denotes the total number of channel uses at the end of the m -th slot. Also, $\xi_n = l_n P_n$ is the energy of the signal transmitted at the m -th (re)transmission round and $\xi^{(m)} = \sum_{n=1}^m l_n P_n$ denotes the sum energy consumed in the first m time slots.

The (re)transmission stops at the end of the m -th retransmission round if the data is successfully decoded at the m -th round and not before. Therefore, implementing random coding and typical set-based decoding, the results of [2], [33, chapter 15] can be used where $\Pr\{S_m\}$ is simplified to the time division multiple access (TDMA)-type equation

$$\Pr\{S_m\} = \Pr \left\{ R^{(m)} \leq \frac{\sum_{n=1}^m l_n \log(1+gP_n)}{\sum_{j=1}^m l_j} \cap R^{(m-1)} > \frac{\sum_{n=1}^{m-1} l_n \log(1+gP_n)}{\sum_{j=1}^{m-1} l_j} \right\} \quad (18)$$

$$\Pr\{\hat{S}_{M+1}\} = \Pr \left\{ R^{(M+1)} > \frac{\sum_{n=1}^{M+1} l_n \log(1+gP_n)}{\sum_{j=1}^{M+1} l_j} \right\}.$$

Note that, based on (17), we have

$$\frac{l_n}{\sum_{j=1}^m l_j} = R^{(m)} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right), \quad R^{(0)} \doteq \infty \quad (19)$$

and so $\Pr\{S_m\}$ is found as a function of $R^{(n)}$'s.

If the data transmission stops at the end of the m -th time slot, the equivalent transmission power is

$$P^{(m)} = \frac{\xi^{(m)}}{\sum_{j=1}^m l_j} = \frac{\sum_{n=1}^m l_n P_n}{\sum_{j=1}^m l_j} = R^{(m)} \sum_{n=1}^m P_n \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \quad (20)$$

where the last equality is based on (19). Consequently, using (18) and (20), the average transmission power and the average transmission energy per packet are respectively found as

$$\bar{P} = \sum_{m=1}^{M+1} P^{(m)} \Pr\{S_m\} + P^{(M+1)} \Pr\{\hat{S}_{M+1}\} \quad (21)$$

and

$$\bar{\xi} = \sum_{m=1}^{M+1} \xi^{(m)} \Pr\{S_m\} + \xi^{(M+1)} \Pr\{\hat{S}_{M+1}\}. \quad (22)$$

A. Power allocation

An outage probability constraint $\Pr\{\text{outage}\} \leq \epsilon$ implies that the rate $R^{(M+1)} = \frac{Q}{L^{(M+1)}}$ should be, with probability $1 - \epsilon$, supported in the last (re)transmission round, i.e.,

$$\Pr\left\{\frac{\sum_{n=1}^{M+1} l_n \log(1 + gP_n)}{\sum_{j=1}^{M+1} l_j} < R^{(M+1)}\right\} = \Pr\left\{\sum_{n=1}^{M+1} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}}\right) \log(1 + gP_n) < 1\right\} \leq \epsilon. \quad (23)$$

Therefore, the INR-based outage-limited power allocation problem can be formulated as

$$\begin{aligned} & \min_{P_1, \dots, P_{M+1}} \left\{ \sum_{m=1}^{M+1} P^{(m)} \Pr\{S_m\} + P^{(M+1)} \Pr\{\hat{S}_{M+1}\} \right\} \\ & \text{subject to } \Pr\left\{\frac{\sum_{n=1}^{M+1} l_n \log(1 + gP_n)}{\sum_{j=1}^{M+1} l_j} < R^{(M+1)}\right\} \leq \epsilon. \end{aligned} \quad (24)$$

The following theorem discusses the optimal adaptive power allocation in the INR protocol with an outage probability constraint.

Theorem 2: For INR HARQ protocols with an outage probability constraint, the transmission energies must be increasing in every retransmission.

Proof: The contributions of the k -th and the $(k+1)$ -th retransmission rounds on the outage probability constraint (23) are $\frac{l_k \log(1 + gP_k)}{\sum_{j=1}^{M+1} l_j}$ and $\frac{l_{k+1} \log(1 + gP_{k+1})}{\sum_{j=1}^{M+1} l_j}$, respectively. Thus, the parameters (l_m, P_m) and (l_{m+1}, P_{m+1}) are interchangeable, in the sense that the outage probability stays the same if the parameters are switched.

However, this is not the case if we instead study the contributions to the average power. The contributions of the $(k+1)$ -th and the k -th retransmission rounds on the average transmission power are

$$\begin{aligned} c(\xi_{k+1}) &= \xi_{k+1} b_{k+1} = P_{k+1} l_{k+1} b_{k+1}, \\ b_{k+1} &= \sum_{n=k+1}^{M+1} \frac{1}{\sum_{j=1}^{M+1} l_j} \Pr\{S_n\} + \frac{1}{\sum_{j=1}^{M+1} l_j} \Pr\{\hat{S}_1, \dots, \hat{S}_{M+1}\} \end{aligned} \quad (25)$$

and

$$\begin{aligned} c(\xi_k) &= \xi_k b_k = P_k l_k b_k \\ &= P_k l_k \left(b_{k+1} + \frac{1}{\sum_{j=1}^k l_j} \Pr\{S_k\} \right) \geq \xi_k b_{k+1}, \end{aligned} \quad (26)$$

respectively. Here, $c(\xi_k)$ is the contribution of the energy term ξ_k on the average transmission power, that is, $\bar{P} = \sum_{k=1}^{M+1} c(\xi_k)$. Therefore, from (25) and (26), it is obvious that in order to minimize the average transmission power subject to an outage probability constraint, the transmission powers and

code lengths should be designed such that the transmission energies are increasing with the number of retransmissions, i.e., $\xi_k \leq \xi_{k+1}, \forall k$. (For simulations, the readers are referred to Fig. 7.) ■

Remark: The difference between the performance of different stop-and-wait HARQ protocols is in the way the probability terms $\Pr\{S_m\}$, $m = 1, \dots, M+1$, are calculated, e.g., [15] (Also, please see RTD and INR schemes as examples.) Moreover, the proofs of Theorems 1 and 2 are valid independent of how the probabilities $\Pr\{S_m\}$ are determined. Thus, the same conclusion can also be proved for the other stop-and-wait HARQ protocols such as basic ARQ for which the validity of the argument has been previously shown in [29].

As stated before, an outage probability constraint can be satisfied with either a few high-power retransmissions or by many low-SNR retransmission rounds. The following theorem demonstrates the optimal power allocation rule for the second case.

Theorem 3: Assume fixed-length coding for the INR. For sufficiently large number of retransmissions, the optimal outage-limited (re)transmission powers follow a specific sequence of numbers which converges to a geometric sequence.

Proof: With fixed-length coding for the INR, i.e., $l_i = L, \forall i$, (17) and (20) lead to $R^{(m)} = \frac{R}{m}$ and $P^{(m)} = \frac{1}{m} \sum_{n=1}^m P_n$ where $R = \frac{Q}{L}$ is the initial codeword rate. Therefore, for large M , (18) and (21) can be used to rephrase the average transmission power (21) as

$$\begin{aligned} \bar{P} &= R \sum_{m=1}^{M+1} \frac{1}{mZ^{(m)}} \left(F_G(Z^{(m-1)}) - F_G(Z^{(m)}) \right) \\ &+ \frac{R}{(M+1)Z^{(M+1)}} F_G(Z^{(M+1)}). \end{aligned} \quad (27)$$

Here, (27) is obtained by 1) defining $Z^{(m)} \doteq \frac{R}{\sum_{n=1}^m P_n}$, 2) the fact that the outage probability constraint can be satisfied with low transmission powers when the number of retransmission rounds increases and 3) $\log(1+x) \rightarrow x$ for small values of x . The outage probability constraint determines $Z^{(M+1)}$ as in (23). Taking the derivative with respect to $Z^{(m)}$, however, results in

$$\begin{aligned} \frac{\partial \bar{P}}{\partial Z^{(m)}} &= R \left(\frac{F_G(Z^{(m)}) - F_G(Z^{(m-1)})}{mZ^{(m)^2}} \right. \\ &- \left. \frac{f_G(Z^{(m)})}{mZ^{(m)}} + \frac{f_G(Z^{(m)})}{(m+1)Z^{(m+1)}} \right) \\ &\stackrel{(a)}{=} R f_G(Z^{(m)}) \left(\frac{1}{(m+1)Z^{(m+1)}} - \frac{Z^{(m-1)}}{mZ^{(m)^2}} \right) \end{aligned}$$

where (a) follows from $(F_G(Z^{(m)}) - F_G(Z^{(m-1)})) \rightarrow f_G(Z^{(m)})(Z^{(m)} - Z^{(m-1)})$ for large values of M . Therefore, setting $\frac{\partial \bar{P}}{\partial Z^{(m)}} = 0$, the optimal power allocation rule is found by the sequence of numbers

$$Z^{(m)} = \sqrt{\frac{m+1}{m} Z^{(m-1)} Z^{(m+1)}}. \quad (28)$$

Particularly, (28) converges to a geometric sequence since $Z^{(m)} \rightarrow \sqrt{Z^{(m-1)} Z^{(m+1)}}$ when m increases. Finally, it is interesting to note that the power allocation rule (28) is

independent of the channel pdf, initial transmission rate and outage probability constraint. ■

Remark: Using (7), (18) and $\log(1+x) \rightarrow x$ for small values of x , we can show that the performance of the RTD and fixed-length coding INR protocols converges when the outage probability constraint gets relaxed or the number of retransmission rounds increases. Thus, the results of Theorem 3 are valid for the RTD as well.

Note that the conclusion of Theorems 2 and 3, which solve the optimal energies minimizing the average power, are different from the results in [24], where the expected energy is the optimization criterion.

Finally, considering short-term power allocation, i.e., $P_m = P \forall m$, the minimum power satisfying the outage probability constraint (23) is obtained by

$$\Pr \left\{ \log(1+gP) < R^{(M+1)} \right\} \leq \epsilon \Rightarrow P = \frac{e^{R^{(M+1)}} - 1}{F_G^{-1}(\epsilon)}.$$

B. Long-term throughput

Given that the data (re)transmission successfully ends at the end of the m -th time slot, the rate $R^{(m)}$ is received at the receiver. Therefore, we can use (18) to find the system throughput as

$$\begin{aligned} \eta &= \sum_{m=1}^{M+1} R^{(m)} \Pr\{S_m\} \\ &= \sum_{m=1}^{M+1} R^{(m)} \Pr \left\{ \sum_{n=1}^{m-1} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \log(1+gP_n) < 1 \leq \sum_{n=1}^m \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \log(1+gP_n) \right\}. \end{aligned} \quad (29)$$

Also, assuming short-term power constraint (29) is simplified to

$$\begin{aligned} \eta &= \sum_{m=1}^{M+1} R^{(m)} \Pr \left\{ R^{(m)} \leq \log(1+gP) < R^{(m-1)} \right\} \\ &= \sum_{m=1}^{M+1} R^{(m)} \left(F_G \left(\frac{e^{R^{(m-1)}} - 1}{P} \right) - F_G \left(\frac{e^{R^{(m)}} - 1}{P} \right) \right). \end{aligned} \quad (30)$$

With the same arguments as in [15], [18], it can be shown that, compared to the RTD scheme, better performance is achieved in the INR HARQ approach. Moreover, we can prove that the data is decoded with less (or equal) number of INR-based retransmission rounds, when compared with the RTD. Hence, the expected feedback load and the expected number of retransmission rounds, which are of interest in limited-feedback and delay-sensitive systems, respectively, are less in the INR. However, the superiority of the INR scheme over the RTD is at the cost of higher complexity in both the transmitter and the receiver. This is because the INR HARQ requires variable-length coding at the transmitter, producing new parity bits in each retransmission round and proper combination of the received signals at the receiver.

VI. PERFORMANCE ANALYSIS FOR THE BURSTING COMMUNICATION MODEL

In this section, the optimal power allocation problem is studied for the bursting communication model, where there is a long idle period between two successive packet transmissions.

A. RTD HARQ protocol

Utilizing the RTD HARQ scheme, the consumed energy is $\xi^{(m)} = L \sum_{n=1}^m P_n$, if the data (re)transmission ends at the end of the m -th (re)transmission round. Therefore, the expected energy within a packet transmission period is obtained by

$$\begin{aligned} E\{\tilde{\xi}\} &= L \sum_{m=1}^{M+1} \left(\sum_{n=1}^m P_n \right) \Pr\{S_m\} + \left(L \sum_{n=1}^{M+1} P_n \right) \Pr\{\hat{S}_{M+1}\} \\ &\stackrel{(b)}{=} L \left(P_1 + \sum_{m=2}^{M+1} P_m F_G \left(\frac{e^R - 1}{\sum_{n=1}^{m-1} P_n} \right) \right) \end{aligned} \quad (31)$$

where (b) follows from (7) and some straightforward calculations. On the other hand, mL channel uses are spent in the first m rounds. Also, independent of the message decoding status, there will be $(M+1)L$ channel uses if all possible retransmission rounds are used. Therefore, the expected number of channel uses and the average transmission power are respectively found as

$$\begin{aligned} E\{\tilde{\tau}\} &= L \sum_{m=1}^{M+1} m \Pr\{S_m\} + (M+1)L \Pr\{\hat{S}_{M+1}\} \\ &\stackrel{(c)}{=} L \left(1 + \sum_{m=1}^M F_G \left(\frac{e^R - 1}{\sum_{n=1}^m P_n} \right) \right) \end{aligned} \quad (32)$$

and

$$\varphi = \frac{E\{\tilde{\xi}\}}{E\{\tilde{\tau}\}} = \frac{P_1 + \sum_{m=2}^{M+1} P_m F_G \left(\frac{e^R - 1}{\sum_{n=1}^{m-1} P_n} \right)}{1 + \sum_{m=1}^M F_G \left(\frac{e^R - 1}{\sum_{n=1}^m P_n} \right)} \quad (33)$$

where (c) follows from (7). In this way, using (10), the power allocation problem can be formulated as

$$\begin{aligned} \min_{P_1, \dots, P_{M+1}} & \left\{ \frac{P_1 + \sum_{m=2}^{M+1} P_m F_G \left(\frac{e^R - 1}{\sum_{n=1}^{m-1} P_n} \right)}{1 + \sum_{m=1}^M F_G \left(\frac{e^R - 1}{\sum_{n=1}^m P_n} \right)} \right\} \\ \text{subject to} & \sum_{m=1}^{M+1} P_m \geq \frac{e^R - 1}{F_G^{-1}(\epsilon)} \end{aligned} \quad (34)$$

which can be solved numerically (Simulation results can be found in Figs. 4-6).

1) *Long-term throughput:* Provided that the receiver can decode the data, Q nats are received by the receiver in each packet. Therefore, the expected received information nats in each packet is

$$\begin{aligned} E\{\tilde{Q}\} &= Q \left(1 - \Pr \left\{ R \geq \log(1+g \sum_{n=1}^{M+1} P_n) \right\} \right) \\ &= Q \left(1 - F_G \left(\frac{e^R - 1}{\sum_{n=1}^{M+1} P_n} \right) \right). \end{aligned} \quad (35)$$

Consequently, using (32) and as $R = \frac{Q}{L}$, the system throughput is obtained by

$$\eta = \frac{R(1 - F_G(\frac{e^{R-1}}{\sum_{n=1}^{M+1} P_n}))}{1 + \sum_{m=1}^M F_G(\frac{e^{R-1}}{\sum_{n=1}^m P_n})} \quad (36)$$

where the transmission powers are determined by (34). Finally, implementing short-term power allocation, i.e., $P_m = P \forall m$, the system throughput is rephrased as

$$\eta = \frac{R(1 - F_G(\frac{e^{R-1}}{(M+1)P}))}{1 + \sum_{m=1}^M F_G(\frac{e^{R-1}}{mP})}. \quad (37)$$

B. INR HARQ protocol

Using (18) and (22), the expected energy consumed within an INR-based packet transmission period is found as

$$E\{\tilde{\xi}\} = P_1 l_1 + \sum_{n=2}^{M+1} P_n l_n \Pr \left\{ \sum_{j=1}^{n-1} \frac{l_j \log(1 + gP_j)}{\sum_{i=1}^{n-1} l_i} < R^{(n-1)} \right\} \quad (38)$$

Also, with the same arguments as for (32), the expected number of channel uses in each packet is

$$\begin{aligned} E\{\tilde{\tau}\} &= \sum_{m=1}^{M+1} \left(\sum_{n=1}^m l_n \right) \Pr\{S_m\} + \left(\sum_{n=1}^{M+1} l_n \right) \Pr\{\hat{S}_{M+1}\} \\ &\stackrel{(d)}{=} l_1 + \sum_{n=2}^{M+1} l_n \Pr \left\{ \sum_{j=1}^{n-1} \frac{l_j \log(1 + gP_j)}{\sum_{i=1}^{n-1} l_i} < R^{(n-1)} \right\} \end{aligned} \quad (39)$$

where (d) comes from (18). Hence, from (17), the average INR-based transmission power is obtained by

$$\begin{aligned} \varphi &= \frac{\frac{P_1}{R^{(1)}} + \sum_{n=2}^{M+1} P_n \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \Theta_n}{\frac{1}{R^{(1)}} + \sum_{n=2}^{M+1} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \Theta_n}, \\ \Theta_n &\doteq \Pr \left\{ \sum_{j=1}^{n-1} \left(\frac{1}{R^{(j)}} - \frac{1}{R^{(j-1)}} \right) \log(1 + gP_j) < 1 \right\}. \end{aligned} \quad (40)$$

Replacing (40) in (24), the power allocation problem can be solved numerically (Please see Fig. 6).

1) *Long-term throughput*: Q information nats are received if the data is decoded at any retransmission round. Therefore, the INR-based expected received information nats in each packet is

$$\begin{aligned} E\{\tilde{Q}\} &= Q(1 - \Pr\{\hat{S}_{M+1}\}) \\ &= Q \Pr \left\{ \sum_{n=1}^{M+1} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \log(1 + gP_n) \geq 1 \right\} \end{aligned} \quad (41)$$

and, using (17) and (39), the long-term throughput is obtained by

$$\eta = \frac{E\{\tilde{Q}\}}{E\{\tilde{\tau}\}} = \frac{\Pr \left\{ \sum_{n=1}^{M+1} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \log(1 + gP_n) \geq 1 \right\}}{\frac{1}{R^{(1)}} + \sum_{n=2}^{M+1} \left(\frac{1}{R^{(n)}} - \frac{1}{R^{(n-1)}} \right) \Theta_n}. \quad (42)$$

Finally, assuming short-term power constraint (42) simplifies to

$$\eta = \frac{1 - F_G\left(\frac{e^{R(M+1)} - 1}{P}\right)}{\sum_{m=1}^{M+1} \frac{F_G\left(\frac{e^{R(m-1)} - 1}{P}\right) - F_G\left(\frac{e^{R(m)} - 1}{P}\right)}{R^{(m)}} + \frac{F_G\left(\frac{e^{R(M+1)} - 1}{P}\right)}{R^{(M+1)}}}. \quad (43)$$

Due to the denominators of (33) and (40), it is difficult to prove the validity of Theorems 1, 2 and 3 under bursting communication model, although they seem intuitively valid. However, the following theorem studies the power allocation problem for simplest cases of RTD and INR HARQ feedback where the validity of the previous arguments is proved for the case of $M = 1$ bit ARQ feedback.

Theorem 4: Considering one bit feedback in the RTD and the INR HARQ protocols under bursting communication model, the following assertions are valid if $R^{(m)} = \frac{R}{m}$, $m = 1, 2$, where $R^{(m)}$ and $\frac{R}{m}$ are the INR- and RTD-based equivalent (re)transmission rates in the m -th round, respectively:

- In both schemes, the transmission powers should be increasing with the number of retransmission rounds.
- In order to satisfy an outage probability constraint, less average transmission power is required in the INR HARQ protocol⁶.

Proof: Please see Appendix A. ■

Finally, note that if there is an additional constraint on the peak transmission power, the adaptive power controllers can be easily updated by the same procedure as in [31, Section III.B] which was developed for throughput maximization in communication setups utilizing quantized CSI.

VII. DISCUSSIONS

In this section, we present some discussions on the system model assumptions and possible extensions of the paper.

A. On temporal variations of the fading coefficients

Block-fading is an appropriate model for the stationary or slow-moving users [2], [17], [18], [23], [24], [30]. For the users with high speed or long codewords, compared to the channel coherence time, one can consider the fast-fading models where each codeword spans a fading block, and the channel changes in each (re)transmission round [15]⁷. In this case, while the average transmission power, i.e., (9), (21), (33) and (40), is obtained with the same procedure as before, the probability terms $\Pr\{S_m\}$ and $\Pr\{\hat{S}_{M+1}\}$ are replaced by

$$\Pr\{S_m\} = \begin{cases} \Pr\left\{ \log(1 + \sum_{n=1}^{m-1} g_n P_n) < R \right. \\ \left. \leq \log(1 + \sum_{n=1}^m g_n P_n) \right\}, & \text{For RTD} \\ \Pr\left\{ \sum_{n=1}^{m-1} \log(1 + g_n P_n) < R \right. \\ \left. \leq \sum_{n=1}^m \log(1 + g_n P_n) \right\}, & \text{For INR} \end{cases}$$

⁶As stated before, the same argument as in part (b) can be proved for arbitrary number of retransmissions, if the continuous communication model is considered.

⁷Under fast-fading channel conditions, the INR protocol is studied with fixed-length coding because the length of the codewords is the same as the fading block length.

$$\Pr\{\hat{S}_{M+1}\} = \begin{cases} \Pr\{\log(1 + \sum_{n=1}^{M+1} g_n P_n) < R\} & \text{For RTD} \\ \Pr\{\sum_{n=1}^{M+1} \log(1 + g_n P_n) < R\} & \text{For INR} \end{cases}$$

where g_m is the channel realization at the m -th round. Here, the interesting point is that the general conclusions of the paper, e.g., Theorems 1, 2 and 4, hold for the fast-fading scenario as well, since the arguments are valid for every given probability terms $\Pr\{S_m\}$ and $\Pr\{\hat{S}_{M+1}\}$, independent of how they are found (For simulation results with fast-fading conditions, please see Figs. 12-14.).

Finally, with block-fading conditions and, e.g., a continuous communication model, it may occur that two different channel realizations are experienced during the last packet transmission of a fading block. In this case, the results of the fast- and block-fading models can be combined to study the system performance during the transmission of that packet. However, with sufficiently long fading blocks, we can ignore the effect of that single packet, as in [2], [17], [18], [23], [24], [30].

B. On the effect of the feedback channel

Throughout the paper, the feedback signal is supposed to be error-free. This is an acceptable assumption in many scenarios because the feedback bits are sent at very low rates, compared to the forward link, and with an appropriate coding they can be decoded error-free [2], [17]–[19], [23], [24], [31]. Also, the energy consumption in the feedback link is normally much less than the one in the forward link, such that it is ignored in the optimization problem [1]–[5], [23], [24]. A more accurate extension of the paper, however, is to take the power consumption in the feedback channel into account. In that case, there is a tradeoff between the feedback channel reliability and the power consumption in the feedback/forward links; with a low-power feedback signal the feedback bits may be decoded erroneously and, consequently, lead to waste of energy/throughput resources in the forward/feedback links. Thus, depending on the feedback channel fading model, power adaptation capability and the coding scheme of the feedback bits, a wide range of results can be observed in the cases with an imperfect feedback channel.

C. Performance improvement via reinforcement algorithms

In harmony with practical HARQ protocols [15]–[29], we considered the same set of (re)transmission rates and powers for each packet transmission. However, for the continuous communication model, the system data transmission efficiency can be improved via implementation of a reinforcement algorithm [45]⁸. In that setup, not only the (re)transmission rates/powers are different in each round but also the whole parameters are adapted in each packet period based on the decisions in the previous packet transmissions. Thus, the performance is improved, as the system adaptation capability increases (For simulation results, please see Fig. 14). However, it

⁸Reinforcement algorithms refer to the schemes where the results of the previous trials and a reward-punishment rule are used for parameter setting in the next steps [45]. Due to the long idle period between successive packet transmissions, the reinforcement schemes are not useful in the bursting communication model.

is worth noting that implementation of reinforcement schemes increases the implementation complexity substantially. Also, many practical HARQ schemes are designed to operate at fixed (re)transmission rates/powers, for which reinforcement algorithms are not feasible. Finally, the reinforcement-based schemes suffer from the error propagation problems where erroneous decoding of a single feedback bit, which may occur in some practical scenarios, affects many successive packet transmissions.

D. Sub-packeting approach

In this paper, we concentrate on a *complete packeting* scheme where Q nats are encoded into a, e.g., RTD-based codeword of length L and the whole codeword is repeated in the retransmissions [46]. An alternative scheme is the *sub-packeting* approach where each codeword is constructed by concatenation of N parallel error detection encoders, producing subcodewords of length $\frac{L}{N}$, and one forward error correction encoder [46, Fig. 2]. In this case, only the erroneous subcodewords, and not the whole codeword, are retransmitted in the retransmission rounds which leads to energy saving and outage probability/transmission delay reduction. The sub-packeting scheme is of interest in the scenario of bursty errors and when short length cyclic redundancy check (CRC) codes are used, and is more flexible than the complete packeting approach. Therefore, an interesting extension of the paper is to study the effect of adaptive power allocation on the performance of the sub-packeting approach when the channel experiences very-fast-fading conditions, e.g., [46], and there is burst error probability. Here, the same approach as before can be used for power adaptation between the subcodewords.

VIII. SIMULATION RESULTS

Simulation results are given for Rayleigh-fading channels $f_G(g) = \lambda e^{-\lambda g}$, $g \geq 0$ where we set $\lambda = 1$. Also, the initial transmission rate is set to $R = 1$, unless otherwise stated. We consider fixed-length coding for the INR, i.e., $R^{(m)} = \frac{R}{m}$. Further, in all figures, the powers are presented in dB. In Figs. 3-11, we consider the block-fading condition; For the bursting model, a different gain realization is experienced in each packet period. With a continuous communication model, many packets are sent in each fading block. In Figs. 12-14, we study the system performance under fast-fading assumption where the channel changes in each retransmission round.

Utilizing the RTD HARQ protocol, Figs. 3-5 investigate the effect of the outage probability constraint on the required transmission power under continuous and bursting communication models. The figures emphasize the validity of Theorems 1 and 4. Optimal power allocation brings sizable gains; with an outage probability 10^{-3} and $M = 1$, optimal power allocation reduces the average power by 9 and 11 dB in the bursting and continuous communication models, respectively, compared to short-term power allocation (Figs. 3-5). Finally, Fig. 5 studies the power efficiency, which is defined as $\Delta\varphi = \varphi^{\text{short-term}} - \varphi^{\text{long-term}}$ with $\varphi^{\text{short-term}}$ and $\varphi^{\text{long-term}}$ being the average power achieved by short- and long-term

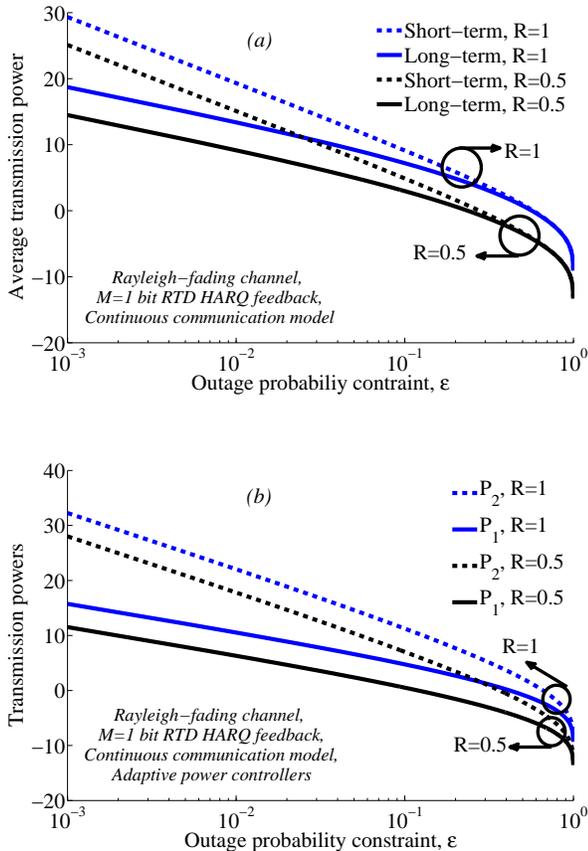


Figure 3. (a) Average transmission power and (b) transmission powers in the (re)transmission rounds as a function of outage probability constraint, Rayleigh-fading channel, $M = 1$ bit RTD HARQ protocol $R = 1$ or 0.5 , continuous communication. The required transmission power increases with the initial transmission rate R . Further, in the optimal case higher power is given to the second retransmission round, compared to the first round.

power allocation, respectively. As illustrated in the figures, the effect of long-term power allocation decreases under relaxed outage probability constraints and increases with the number of retransmission rounds.

Figs. 6 and 7 demonstrate the effect of an outage probability constraint on the required transmission power of the INR protocol and compare the results with the ones obtained for the RTD protocol. Here, the results show that 1) with medium outage probability constraints, the average transmission power of the INR HARQ scheme is about 1.2dB less than the one in the RTD protocol (Fig. 6). 2) Less transmission power is required to satisfy an outage probability constraint in the continuous communication model when compared with the bursting model. However, this difference diminishes as the outage probability constraint becomes more relaxed, i.e., ϵ becomes larger (Fig. 6). 3) Considerable performance improvement is achieved with adaptive power allocation in hard outage probability constraints, i.e., small ϵ 's. For instance, comparing with short-term power allocation at outage probability 10^{-3} and $M = 1$, adaptive power allocation leads to 8 and 9dB average power reduction in bursting and continuous communication models, respectively (Fig. 6). However, the

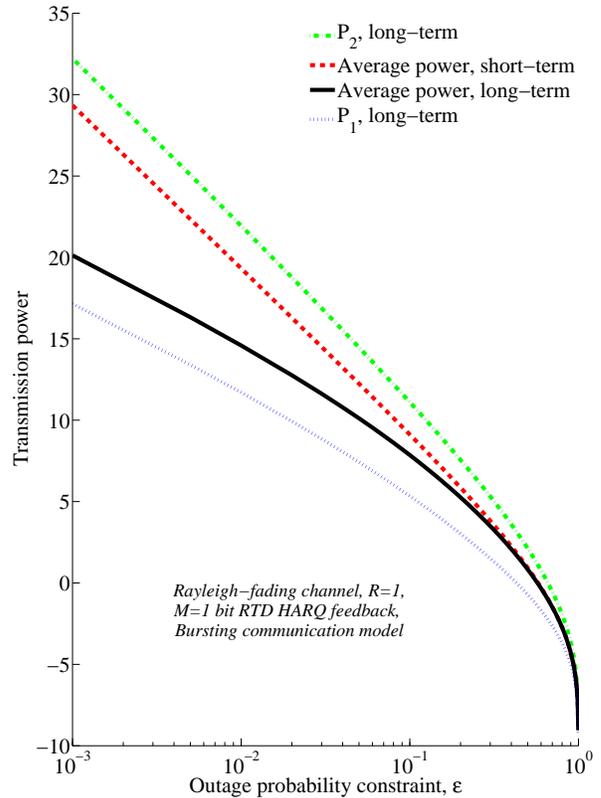


Figure 4. Transmission power vs. outage probability constraint, Rayleigh-fading channel, $M = 1$ bit RTD HARQ feedback $R = 1$, different power allocation scenarios. Under long-term power allocation condition, the optimal transmission power in the first round is less than the one in the second round.

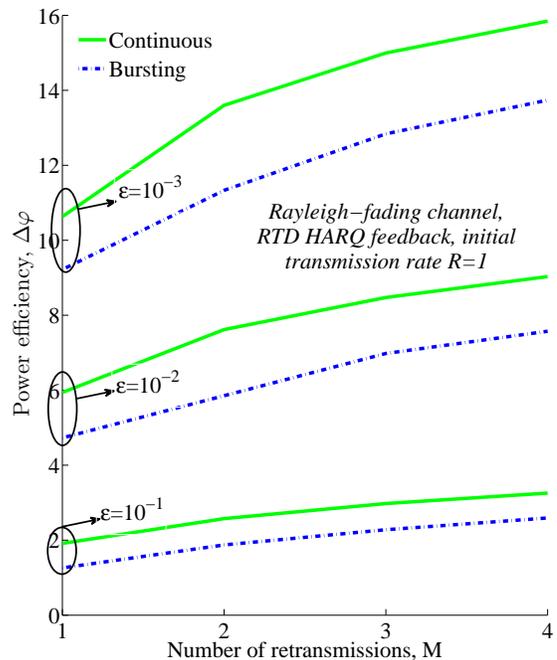


Figure 5. Power efficiency $\Delta\varphi = \varphi^{\text{short-term}} - \varphi^{\text{long-term}}$ for different number of retransmission rounds M . Rayleigh-fading channel, RTD HARQ feedback $R = 1$. The data is (re)transmitted a maximum of $(M + 1)$ rounds.

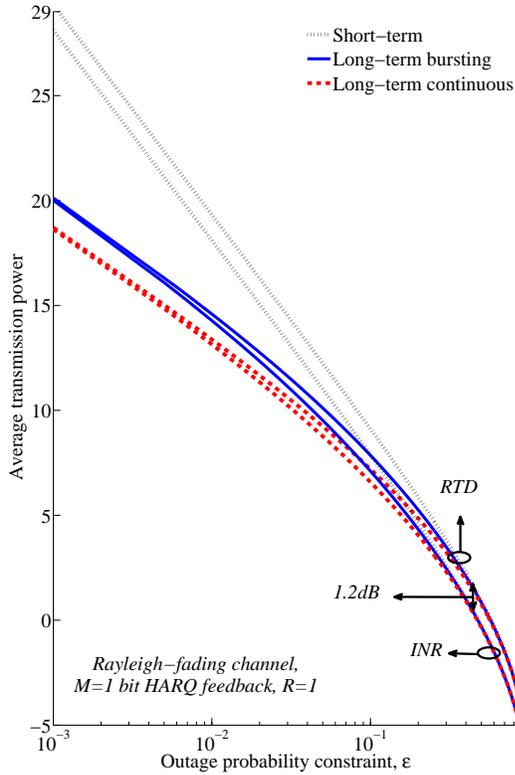


Figure 6. Average transmission power vs outage probability constraint, Rayleigh-fading channel, $M = 1$ bit HARQ feedback $R = 1$, different power allocation scenarios. With high outage probability constraints, the average transmission power of the INR HARQ scheme is 1.2dB less than the one in the RTD protocol. Further, less transmission power is required to satisfy an outage probability constraint in the continuous communication model when compared with the bursting model.

effect of power allocation diminishes under relaxed outage probability constraints (Figs. 6 and 7). Moreover, Figs. 3, 4 and 7 emphasize that with long-term power allocation and for both protocols, the optimal transmission power in the first round is less than the one in the second round, in harmony with Theorems 1, 2 and 4.

Considering $M = 1$ and 2 retransmission rounds, Fig. 8 shows the transmission power as a function of the outage probability constraint under short-term power allocation condition. In particular, Fig. 8 (Fig. 3) indicates that the required transmission power satisfying an outage probability constraint decreases (increases) with the number of retransmission rounds (the initial transmission rate R). Moreover, comparing Figs. 3-8, optimal power allocation is much more effective than increasing the number of retransmission rounds in improving the outage-limited performance of HARQ protocols.

Figures 9-11 investigate the system throughput versus the outage probability constraint in different conditions. Here, it is emphasized that with the same transmission power, less outage probability and higher throughput are achieved for the INR protocol, compared to RTD. Also, higher throughput and less (or equal) outage probability are achievable in the continuous model, in comparison with bursting model (Fig. 9 and 10). Moreover, considering the same powers, long-term

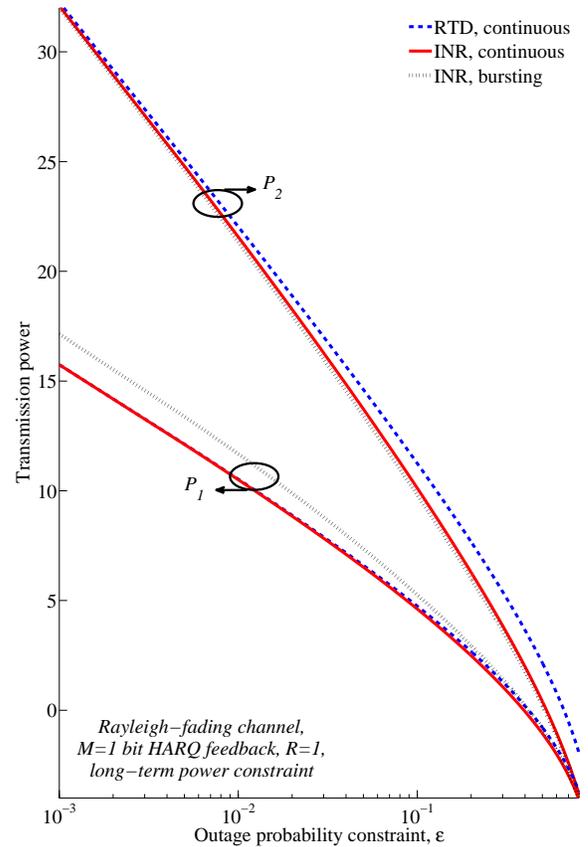


Figure 7. The first and the second rounds transmission power in different HARQ schemes and communication models, Rayleigh-fading channel, $M = 1$ bit HARQ feedback $R = 1$, adaptive power allocation scenario. The effect of power allocation diminishes under relaxed outage probability constraints, i.e., $\epsilon \rightarrow 1$.

power allocation wins (loses) in competition with short-term power allocation in terms of outage probability (throughput) (Fig. 10). This point is further evaluated in Fig. 11 where the relative throughput loss $\Delta\eta = \frac{\eta^{\text{short-term}} - \eta^{\text{long-term}}}{\eta^{\text{long-term}}} \%$ is plotted versus the power efficiency $\Delta\varphi$. Here, $\eta^{\text{short-term}}$ and $\eta^{\text{long-term}}$ denote the throughput achieved by short- and long-term power allocation, respectively. Interestingly, the relative throughput loss, due to outage-limited power allocation, diminishes at relaxed and hard outage probability constraints, because the required average power goes to zero and infinity, respectively. Also, the figures emphasize that different optimal power allocation strategies are obtained, depending on whether the outage probability or the system throughput is considered as the optimization criterion. That is, the power-limited outage minimization and the power-limited throughput maximization problems are not interchangeable.

As mentioned before, block-fading is a coarse but effective theoretical model of the fading channels, particularly for the stationary/slow-moving users. In Figs. 12-14, we analyze the system performance with the more realistic correlated time-varying fading model of [47, eq. 12] (and the references therein) where the channel changes in each codeword transmission period according to a first-order Gauss-Markov process

$$h_{k+1} = \beta h_k + \sqrt{1 - \beta^2} \vartheta, \vartheta \sim \mathcal{CN}(0, 1).$$

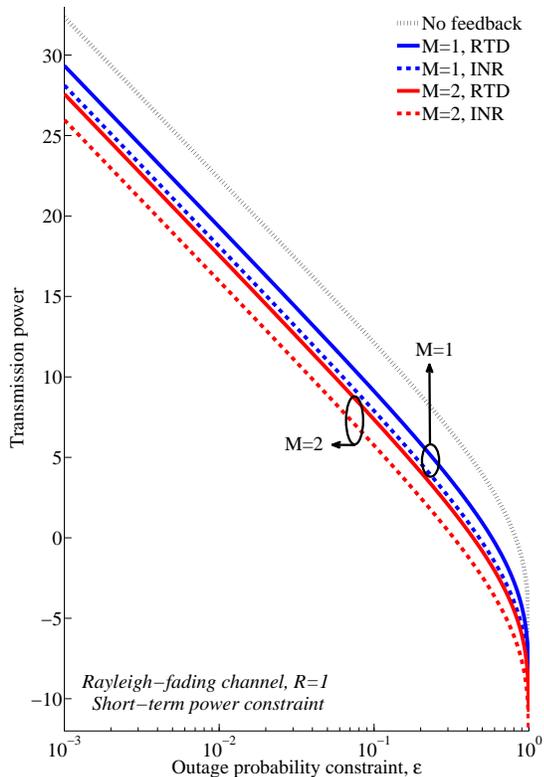


Figure 8. Transmission power vs outage probability constraint, short-term power constraint, Rayleigh-fading channel, $M = 1$ or 2 bits HARQ feedback, $R = 1$. The required transmission power satisfying an outage probability constraint decreases with the number of retransmission rounds.

Here, β is the correlation factor of fading realizations experienced in two successive codeword transmissions with $\beta = 0$ (resp. $\beta = 1$) representing the uncorrelated fast-fading (resp. block-fading) channel model. With this model, different channel realizations are experienced during a packet period. The simulation results indicate that the general conclusions of the paper, e.g., Theorems 1, 2 and 4, hold independently of the fading channel characteristics (Please see Subsection VII.A as well). For a given outage probability, the average transmission power increases with temporal correlation of the fading channel (Fig. 12). However, the gain of optimal power allocation is almost constant for different correlation conditions (Fig. 13). Finally, for further discussions about the fading channel models, the readers are referred to, e.g., [37]–[40], which present remarkable equivalences between the block-fading and continuous fading models, when the Doppler spectrum of the continuous fading model is bandlimited.

To investigate the effect of reinforcement-based schemes, we obtain the outage-limited average power of an RTD protocol using Algorithm 2. The algorithm follows the same structure as in adaptive quantization schemes [48, Section 11.6].⁹ For a given outage probability constraint, the updating steps $d_i, i = 1, \dots, 4$, and the initial transmission

⁹Algorithm 2 is presented for the case with a maximum of $M = 1$ retransmission. The extension of the results to the case with arbitrary number of retransmissions is straightforward.

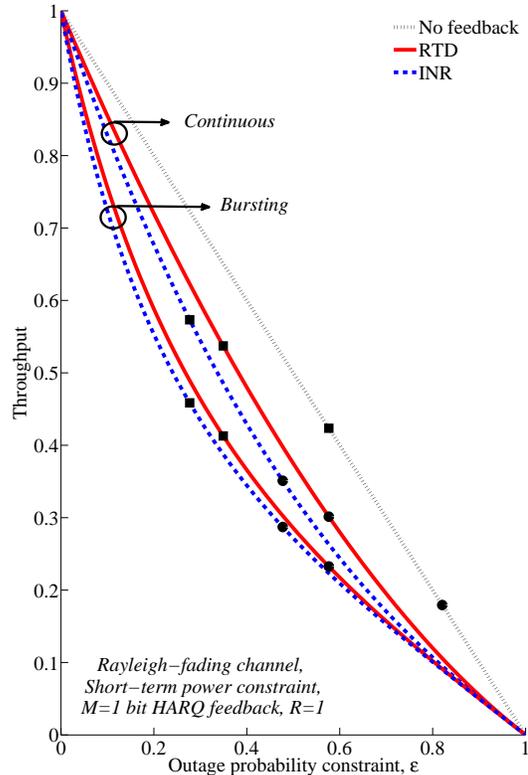


Figure 9. Throughput vs outage probability constraint, short-term power constraint, Rayleigh-fading channel, $M = 1$ bit HARQ feedback, $R = 1$. Circles (squares) represent the results with transmission power $P = 1$ ($P = 2$). With the same transmission power, less outage probability and higher throughput are achieved in the INR protocol, compared to RTD. Also, higher throughput is achievable in the continuous model, in comparison with the bursting model.

power P_{initial} of the algorithm are optimized by exhaustive search, to minimize the average power. The simulation results are presented in Fig. 14 which show that, depending on the fading condition, considerable performance improvement can be achieved by applying reinforcement-based techniques. However, as mentioned before, reinforcement-based schemes increase the implementation complexity substantially and are sensitive to feedback channel noise. Also, reinforcement algorithms are not useful in the bursting communication model or in the continuous communication model with uncorrelated fast-fading condition.

The intuition behind the better system performance in the continuous communication model (in comparison with the bursting) is that more packets are transmitted in *good* channels in continuous communication than in the bursting model. In other words, the good channels are more efficiently exploited in the continuous model. Finally, it is interesting to note that, comparing with the case when there is no feedback (no retransmission), considerable performance improvement is obtained with limited number of feedback bits (see Figs. 8-10).

IX. CONCLUSION

Considering delay-sensitive networks, this paper analyzed the performance of the RTD and INR HARQ protocols in outage-limited data transmission conditions. The results

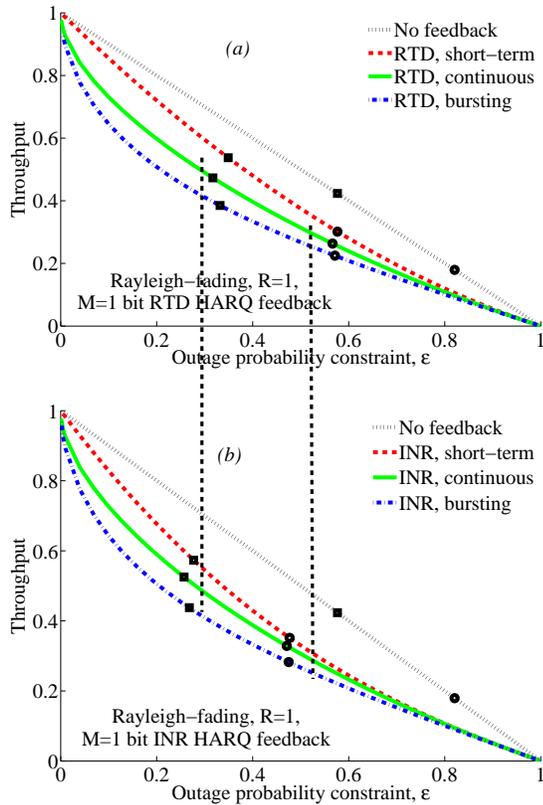


Figure 10. Throughput vs outage probability constraint in (a) RTD or (b) INR protocols, different power allocation conditions, Rayleigh-fading channel, $M = 1$ bit HARQ feedback, $R = 1$. Circles (squares) represent the results with average transmission power $P = 1$ ($P = 2$). In all schemes, INR outperforms the RTD in terms of throughput and outage probability, for a given average transmission power. Further, higher throughput and lower outage probability is achieved in the continuous communication model, in comparison to the bursting model. With the same average powers, long-term power allocation wins (loses) in competition with short-term power allocation in terms of outage probability (throughput).

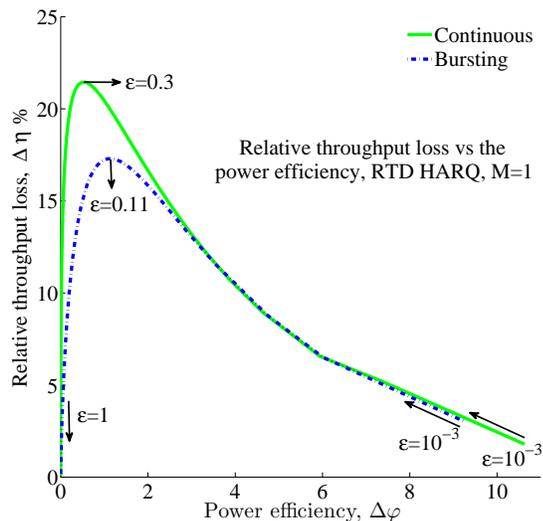


Figure 11. Relative throughput loss $\Delta\eta = \frac{\eta^{\text{short-term}} - \eta^{\text{long-term}}}{\eta^{\text{long-term}}} \%$ vs the power efficiency $\Delta\varphi$ dB. RTD HARQ protocol, Rayleigh-fading channel, $M = 1$ bit feedback, $R = 1$. Down right is the desired region in the figure.

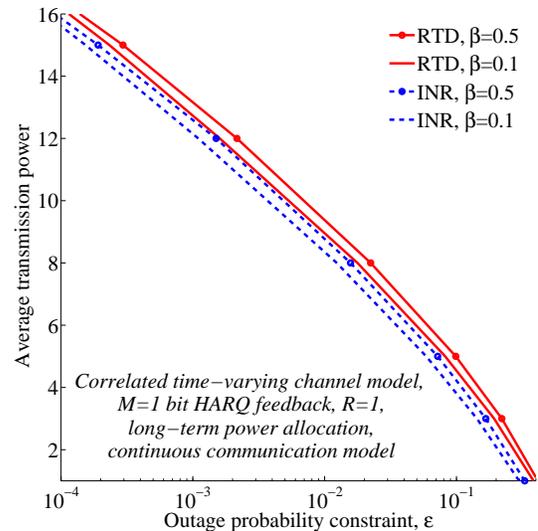


Figure 12. Performance analysis for the correlated time-varying fading model of [47], $M = 1$ bit feedback, $R = 1$, long-term power allocation, continuous communication model.

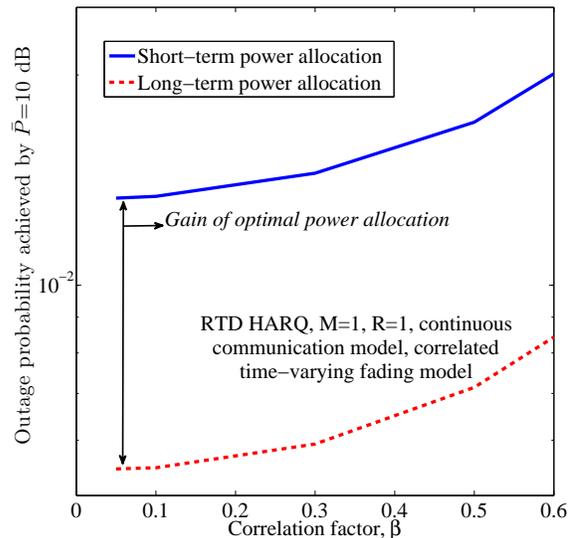


Figure 13. Outage probability achieved by average power $\bar{P} = 10$ dB and different correlation factors β . Correlated Rayleigh fading channel model of [47]. The difference between two curves, i.e., the gain of optimal power allocation is almost constant for different values of β .

emphasize that considerable performance improvement (in terms of both the required transmission power and the system throughput) is achieved with limited number of HARQ-based feedback bits. To minimize the outage-limited average power in the RTD and INR protocols, the transmission powers and energies must increase in every retransmission, respectively. Moreover, for sufficiently large number of retransmissions, the optimal (re)transmission powers can be determined via a specific sequence of numbers. This sequence, which is independent of the fading distribution, converges to a geometric sequence when the order of retransmission rounds increases. Also, adaptive power allocation between the data retransmission rounds is very beneficial in hard outage probability constraints while its effect diminishes in more relaxed outage-limited conditions. Depending on the

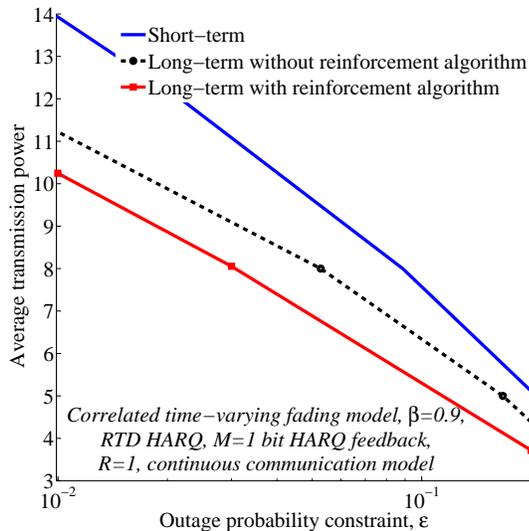


Figure 14. Outage-limited average power for different power allocation schemes, RTD HARQ, $M = 1$. Correlated Rayleigh fading channel model of [47], $\beta = 0.9$. For the reinforcement-based scheme, Algorithm 2 is used where the constants d_1, \dots, d_4 are optimized, in terms of average power, for every given outage probability.

Algorithm 2 ARQ-based data transmission by a reinforcement algorithm

- I. For a given initial transmission rate R , set the initial transmission power to $\check{P} = P_{\text{initial}}$.
 - II. Start a packet transmission with power \check{P} and do the following procedures
 - 1) If the codeword is decoded in round $m = 1$, $\check{P} \leftarrow (1 - d_1)\check{P}$, and go to II.
 - 2) If the codeword is not decoded in round $m = 1$, set $\check{P} \leftarrow (1 + d_2)\check{P}$ and retransmit the codeword. Continue as follows:
 - a) If the codeword is decoded in round $m = 2$, $\check{P} \leftarrow (1 - d_3)\check{P}$ and go to II.
 - b) If the codeword is not decoded in round $m = 2$, declare an outage, set $\check{P} \leftarrow (1 + d_4)\check{P}$ and go to II.
-

channel condition, the power efficiency can be improved by reinforcement-based power allocation schemes, at the cost of higher implementation complexity and sensitivity to imperfect feedback signal. Finally, studying the system performance with an imperfect feedback channel and implementation of sub-packeting schemes are interesting extensions of the paper.

APPENDIX A
PROOF OF THEOREM 4

Assuming $R^{(m)} = \frac{R}{m}$, Theorem 3 implies $l_1 = l_2$, i.e., the codeword lengths are the same in the INR protocol. Therefore, from (33) and (40), the average transmission power in both schemes can be rewritten as

$$\varphi = \frac{P_1 + P_2 F_G\left(\frac{e^R - 1}{P_1}\right)}{1 + F_G\left(\frac{e^R - 1}{P_1}\right)}. \quad (44)$$

To prove part (a), we consider two cases {Case 1 : $(P_1 = P + \Delta, P_2 = P)$ } and {Case 2 : $(P_1 = P, P_2 = P + \Delta)$ }, $\Delta \geq 0$

and show that less average transmission power is obtained in the second case. Note that based on (10) and (23) there is no preference between the transmission powers from the outage probability constraint perspective, as the powers are interchangeable. Then, based on the following inequalities

$$\begin{aligned} \varphi_{\text{case 1}} &= \frac{P + \Delta + P F_G\left(\frac{e^R - 1}{P + \Delta}\right)}{1 + F_G\left(\frac{e^R - 1}{P + \Delta}\right)} \geq \frac{P + (P + \Delta) F_G\left(\frac{e^R - 1}{P}\right)}{1 + F_G\left(\frac{e^R - 1}{P}\right)} = \varphi_{\text{case 2}} \\ &\Leftrightarrow (P + \Delta + P F_G\left(\frac{e^R - 1}{P + \Delta}\right))(1 + F_G\left(\frac{e^R - 1}{P}\right)) \\ &\quad \geq (P + (P + \Delta) F_G\left(\frac{e^R - 1}{P}\right))(1 + F_G\left(\frac{e^R - 1}{P + \Delta}\right)) \\ &\Leftrightarrow 1 \geq F_G\left(\frac{e^R - 1}{P + \Delta}\right) F_G\left(\frac{e^R - 1}{P}\right) \end{aligned} \quad (45)$$

it is obvious that less average transmission power is obtained in the second case, i.e., the optimal (re)transmission powers minimizing the average transmission power should be increasing with the number of retransmission rounds.

To prove part (b), we assume that the transmission powers P_1 and P_2 in (44) have been optimized for the RTD-based scheme, i.e., according to the constraint $\Pr\{\log(1 + g(P_1 + P_2)) < R\} \leq \epsilon$. Now, the same transmission power P_1 is considered for the first transmission round of the INR-based scheme which is not necessarily optimal for this protocol. Then, according to the property $\log(1 + ax) + \log(1 + by) \geq \log(1 + ax + by)$, $\forall x, y, a, b \geq 0$, it can be easily shown that less transmission power than P_2 is required to satisfy the INR-based outage probability constraint $\Pr\{\log(1 + gP_1) + \log(1 + gP_2) < R\} \leq \epsilon$. That is, defining

$$w \doteq \arg\{\Pr\{\log(1 + gP_1) + \log(1 + gx) < R\} = \epsilon\}$$

we have $w \leq P_2$. Therefore, as the average power in both schemes is found by (44), the optimal outage-limited average transmission power required in the INR HARQ scheme is less than the one required in the RTD protocol. Finally, the theorem can be further verified in Figs. 4-7.

REFERENCES

- [1] B. Makki and T. Eriksson, "Data transmission in the presence of noisy channel state feedback and outage probability constraint," in *ISITA*, Oct. 2010, pp. 458–463.
- [2] G. Caire, et. al, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, 1999.
- [3] B. Khoshnevis and W. Yu, "Bit allocation law for multiantenna channel feedback quantization: Single-user case," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2270–2283, May 2011.
- [4] S. Ekbatani, et. al, "Outage behavior of slow fading channels with power control using partial and erroneous CSIT," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6097–6102, Dec. 2010.
- [5] 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.
- [6] B. Gammage, et. al, "Gartner's top predictions for IT organizations and users, 2010 and beyond: A new balance," Gartner Report, Dec. 2009.
- [7] S. Cui, et. al, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.
- [8] Y. Chen, et. al, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.
- [9] G. Gur and F. Alagöz, "Green wireless communications via cognitive dimension: an overview," *IEEE Netw.*, vol. 25, no. 2, pp. 50–56, March 2011.
- [10] M. Ismail and W. Zhuang, "Network cooperation for energy saving in green radio communications," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 76–81, Oct. 2011.

- [11] B. Wang, et. al, "Green wireless communications: A time-reversal paradigm," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1698–1710, Sept. 2011.
- [12] W. Feng, et. al, "Green information and communication technology: energy efficiency in a motorway model," *IET Commun.*, vol. 4, no. 7, pp. 850–860, 2010.
- [13] C. Comaniciu, et. al, "Radio resource management for green wireless networks," in *VTC*, Sept. 2009, pp. 1–5.
- [14] T. Chen, et. al, "Energy efficiency metrics for green wireless communications," in *WCSP*, Oct. 2010, pp. 1–6.
- [15] 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.
- [16] R. Knopp, "Coding and multiple access over fading channels," Ph.D. dissertation, Ecole Polytech. Feder. Lausanne, Switzerland, 1997.
- [17] B. Makki and T. Eriksson, "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.
- [18] C. Shen, et. al, "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.
- [19] 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.
- [20] S. Sesia, et. al, "Incremental redundancy hybrid ARQ schemes based on low-density parity-check codes," *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1311–1321, Aug. 2004.
- [21] S. Park and D. 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.
- [22] S. M. Kim, et. al, "Optimal rate adaptation for hybrid ARQ in time-correlated rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 968–979, March 2011.
- [23] W. Su, et. al, "Optimal power assignment for minimizing the average total transmission power in hybrid-ARQ rayleigh fading links," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1867–1877, July 2011.
- [24] 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.
- [25] M. K. Chang, et. al, "Power control for packet-based wireless communication systems," in *WCNC*, vol. 1, March 2003, pp. 542–546.
- [26] N. Arulselvan and R. Berry, "Efficient power allocations in wireless ARQ protocols," in *WPMC*, vol. 3, Oct. 2002, pp. 976–980.
- [27] D. V. Djonin, et. al, "Joint rate and power adaptation for type-I hybrid ARQ systems over correlated fading channels under different buffer-cost constraints," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 421–435, Jan. 2008.
- [28] S. Lee, et. al, "Cooperative decode-and-forward ARQ relaying: Performance analysis and power optimization," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2632–2642, Aug. 2010.
- [29] H. Seo and B. G. Lee, "Optimal transmission power for single- and multi-hop links in wireless packet networks with ARQ capability," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 996–1006, May 2007.
- [30] S. Ekbatani, et. al, "Throughput maximization over slowly fading channels using quantized and erroneous feedback," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2528–2533, Sept. 2009.
- [31] T. T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 820–829, April 2007.
- [32] 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.
- [33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley Interscience, 1992.
- [34] C. Shen and M. 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.
- [35] T. V. K. Chaitanya and E. G. Larsson, "Optimal resource allocation for IR-HARQ," in *Swe-CTW*, Oct. 2011, pp. 74–79.
- [36] A. K. Karmokar, et. al, "Delay constrained rate and power adaptation over correlated fading channels," in *GLOBECOM*, vol. 6, Nov. 2004, pp. 3448–3453.
- [37] A. Lozano, et. al, "Fundamental limits of cooperation," arXiv:1204.0011v1 2012, available at: <http://arxiv.org/abs/1204.0011v1>.
- [38] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2939–2948, Oct. 2010.
- [39] A. Lozano and D. Porrat, "Non-peaky signals in wideband fading channels: Achievable bit rates and optimal bandwidth," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 246–257, Jan. 2012.
- [40] E. Biglieri, et. al, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [41] H. Holma and A. Toskala, "WCDMA for UMTS: HSPA evolution and LTE," New York: Wiley, 2007.
- [42] M. Effros, A. Goldsmith, and Y. Liang, "Capacity definitions of general channels with receiver side information," in *ISIT*, 2007, pp. 921–925.
- [43] I. N. Bronshtein, et. al, *Handbook of Mathematics*. Springer, 5th ed., 2007.
- [44] B. Makki, et. al, "Unaligned training for voice conversion based on a local nonlinear principal component analysis approach," *Neural comp. and appl.*, vol. 19, no. 3, pp. 437–444, Nov. 2009.
- [45] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT press, 1998.
- [46] Y. Zhou and J. Wang, "Optimum subpacket transmission for hybrid ARQ systems," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 934–942, May 2006.
- [47] B. Makki and T. Eriksson, "CSI feedback in correlated slow-fading channels," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1294–1297, Dec. 2011.
- [48] K. Sayood, *Introduction to Data Compression*. Morgan Kaufmann Publishers, 3rd ed., 2006.