Power allocation in repetition time diversity hybrid automatic repeat request feedback

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 Wireless Communications and Networking Conference, WCNC 2012, Paris, 1-4 April 2012 (ISSN: 1525-3511)

Citation for the published paper:

http://dx.doi.org/10.1109/WCNC.2012.6214182

Downloaded from: http://publications.lib.chalmers.se/publication/163181

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.
Power Allocation in Repetition Time Diversity
Hybrid Automatic Repeat Request Feedback

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

Abstract—This paper addresses the problem of optimal power allocation for hybrid automatic repeat request (HARQ) feedback over slowly-fading channels. We mainly focus on the repetition time diversity HARQ scheme where the results are obtained for both continuous and bursting communication models. Moreover, the effect of an outage probability constraint on the system data transmission efficiency is studied under different transmission power constraints. Simulation results show that 1) for Nakagami fading channels, the optimal HARQ-based (re)transmission powers maximizing the system throughput should be decreasing in every (re)transmission round, 2) higher rates are achieved in the continuous communication, when compared with the bursting model, and 3) HARQ feedback leads to considerable performance improvement even in outage-limited conditions.

I. INTRODUCTION

Automatic repeat request (ARQ) is a well-known approach applied in today’s wireless networks to increase the data transmission reliability in the absence of channel state information at the transmitter (CSIT) [1]–[8]. In a general ARQ approach, the transmitter considers some initial transmission rate and power with no pre-knowledge about the channel quality. Then, with the help of ARQ, the decoding status at the receiver will be reported back to the transmitter via one bit feedback. The feedback indicates successful decoding of the received signal by an acknowledgement (ACK) bit and failed decoding by a negative acknowledgement (NACK) bit. Based on the received feedback, it is decided by the transmitter whether to retransmit the data or to move on to the next codeword. In basic ARQ approaches the same data is retransmitted and the receiver decodes the message based on the received signal in each time slot. Hybrid ARQ (HARQ) protocols, on the other hand, are more advanced methods where, while utilizing both forward error correction and error detection, the receiver combines all received representations of a message.

Among different HARQ schemes, repetition time diversity (RTD) [3]–[6] has been introduced as one of the best fixed-length coding HARQ techniques where the same codeword is retransmitted in each retransmission and the receiver performs maximum ratio combining of all received signals. Hence, it leads to an acceptable implementation complexity, as not only the code lengths are the same in all retransmissions but also no new parity symbols need to be generated in the retransmission rounds.

In wireless devices with limited power resources, optimal power allocation in the HARQ retransmission rounds is a key point for increasing the system data transmission efficiency. In [9] the optimal power allocation for different HARQ protocols was investigated under perfect CSIT assumption. Also, [10] presented a linear programming approach for determining the optimal transmission power in the retransmission rounds. Here, it is assumed that partial CSIT is available and the power values are selected from a set of finite power levels. Seo and Lee [11] determined the optimal transmission power allocation approach for the basic ARQ schemes where the data is decoded based on the received signal in each transmission round. Lee, et. al, [12] investigated the power allocation in a relay channel where, while the powers are fixed in the retransmissions, there is a sum power constraint on the relay and the transmitter. Finally, considering a down-link W-CDMA system, [13] studied the optimal power allocation strategies such that the total number of ARQ-based retransmission rounds required for successfully receiving the data is minimized.

In this perspective, considering slowly-fading channels, this paper investigates the effect of optimal power allocation on the data transmission efficiency of the RTD HARQ protocol. The goal is to determine the optimal transmission powers and rates maximizing the system throughput under power-limited conditions. The results are obtained under both continuous and bursting communication models. Also, we evaluate the effect of an outage probability constraint on the system performance. Simulation results show that 1) HARQ feedback results in considerable performance improvement even in outage-limited conditions, 2) in comparison with the bursting model, higher rates are achievable in the continuous communication model, 3) for Nakagami fading channels and under both bursting and continuous communication models, the optimal transmission powers maximizing the system throughput should be decreasing in every retransmission round and 4) the HARQ-based system throughput is less affected by the outage probability constraint, when compared with the case where there is no feedback to the transmitter.

II. SYSTEM MODEL

We consider a slowly-fading communication setup where the power-limited input message $X$ multiplied by the random variable $h$ is summed with an independent and identically distributed (iid) complex Gaussian noise $Z \sim \mathcal{CN}(0, N_0)$ resulting in the output

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

Motivated by the transmission of training sequences, it is assumed that there is perfect instantaneous knowledge about the channel gain at the receiver, which is an acceptable assumption under slowly-fading conditions \([5]–[11]\). Also, the HARQ feedback bits are assumed to be delivered at the transmitter error-free and with negligible delay. Furthermore, as each transmission experiences an AWGN channel, all results are restricted to Gaussian input distributions. Finally, all results are presented in natural logarithm basis and in all simulations the throughput is presented in nats-per-channel-use (ncpu). This is a good model for networks with stationary or slow-moving users such as wireless local area networks (WLANs) \([14]\). Particularly, since long-block-length capacity-approaching codes can be implemented in such systems, the results can provide realistic insight about the performance bounds of the considered schemes\(^1\).

**Evaluation yardstick:** Our figure of merit is the long-term throughput. Let the total number of channel uses and the total number of successfully decoded information nats up to the end of the \( k \)-th time slot be denoted by \( \tau^{(k)} \) and \( Q^{(k)} \), respectively. In this way, the long-term throughput is defined as \([16]\)

\[
\eta = \lim_{k \to \infty} \frac{Q^{(k)}}{\tau^{(k)}} = \frac{E\{Q\} }{E\{\tau\}} = \frac{E\{R\} }{E\{T\} } \tag{2}
\]

where \( Q \) is the number of successfully decoded information nats in each transmission and \( \tau \) is the number of channel uses. Also, \( (a) \) is based on the renewal-reward theorem where \( \bar{R} \) represents the reward, i.e., the number of nats successfully decoded per channel use when the transmission of a packet\(^2\) ends and \( \bar{T} \) is the number of HARQ-based (re)transmissions needed to complete the packet transmission, normally denoted inter-renewal time \([3], [4], [16]\).

Here, we consider two, namely, continuous and bursting, communication schemes \([6]–[8]\). Under the continuous communication model, it is assumed that there is an infinite amount of information available at the transmitter and it is always active. 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. Therefore, the long-term throughput can be calculated as follows. Let \( R(g) \) be the instantaneous data rate of the RTD HARQ approach for a given gain realization \( g \). Then, 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

\[
\eta = \frac{E\{L_c R(g)\}}{L_c} = E\{R(g)\} = \bar{R} \tag{3}
\]

where \( \bar{R} \) is the channel average rate \([6]–[8], [17], [18]\).

Under the bursting communication model, on the other hand, it is assumed that there is a long idle period between the packets transmission. Therefore, while the HARQ retransmission rounds of each packet experience the same gain realization, the channel changes independently from one packet to another. In this case, the denominator of (2) is not constant and, as discussed in the following, should be calculated separately.

**Power constraint:** In general, the average transmission power is obtained by \([15]\)

\[
\phi = \frac{E\{|\xi|\}}{E\{\tau\}} \tag{4}
\]

where \( E\{|\xi|\} \) is the expected energy consumed within a packet transmission period. Then, defining \( P(g) \) as the transmission power random variable for a gain realization \( g \), the average power in the continuous communication model is found as

\[
\phi = \frac{E\{L_c P(g)\}}{L_c} = E\{P(g)\} = \bar{P}. \tag{5}
\]

However, as seen in the sequel, this argument is not valid under the bursting communication assumption and the average power should be calculated directly based on (4).

**A. Ultimate system performance**

Assuming perfect CSIT, it is well accepted, e.g., \([6], [7], [18], [19]\), that the system throughput can be calculated by

\[
\eta_{\text{perfect}} = \int_0^\infty f_G(g) \log (1 + gP(g)) \, dg \tag{6}
\]

where \( P(g) \) is the optimal power allocation function determined based on the power constraint. For instance, considering an average transmission power constraint \( \bar{P} = \int_0^\infty f_G(g) P(g) \, dg \leq P \), the optimal power allocation function is determined based on the Lagrange objective function \( \Upsilon = \eta_{\text{perfect}} - \lambda \bar{P} \) which leads to

\[
\frac{\partial \Upsilon}{\partial P(g)} = 0 \Rightarrow T(g) = \begin{cases} 0, & g < \beta^* \\ \frac{1}{\beta^*} - \frac{1}{g}, & g \geq \beta^* \end{cases} \tag{7}
\]

Here, \( P \) denotes the average transmission power constraint, \( \beta \) is the Lagrange multiplier and \( \beta^* \) is the power allocation threshold obtained by

\[
\beta^* = \arg \left\{ \int_\beta^\infty \left( \frac{1}{\beta} - \frac{1}{g} \right) f_G(g) \, dg = P \right\}. \tag{8}
\]

Considering (6)-(8) and the Nakagami-2 fading channel model \( f_G(g) = \frac{2\gamma g^{2\gamma - 1} e^{-\frac{g}{\omega}}}{\sqrt{\pi} \Gamma(\frac{1}{2})} \), on which we focus, the system throughput with perfect CSIT is found as

\[
\eta_{\text{perfect}} = \int_{\beta^*}^\infty f_G(g) \log \left( \frac{\omega}{P} \right) \, dg = \frac{1}{2w} E\{\left( \frac{2}{w} \right)^{\beta^*} \} - \frac{1}{2} e^{-\frac{\omega}{4w^2}} - \left( \frac{1}{\beta^*} - \frac{\gamma - 2}{\omega} \right) e^{-\frac{\gamma - 2}{\omega}} = \bar{P} \tag{9}
\]
where \( \text{Ei}(x) = \int_x^{\infty} \frac{e^{-u}}{u} \, du \), \( x \geq 0 \) is the exponential integral function, \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt \) denotes the error function and \( w \) is the Nakagami-2 pdf parameter, normally determined based on the path loss and shadowing between the terminals.

With no CSIT, on the other hand, the data is transmitted at a fixed rate \( R \) which is decoded if the channel realization supports the rate, i.e., \( R \leq \log(1 + gP) \). Therefore, representing the gain cumulative distribution function (cdf) by \( F_G(g) \), the no-CSIT system throughput is obtained by

\[
\eta_{\text{ns}} = \max_R \left( 1 - F_G\left( \frac{R - 1}{P} \right) \right)
\]

which can be solved numerically.

III. SYSTEM THROUGHPUT UNDER CONTINUOUS COMMUNICATION ASSUMPTION

We consider a maximum of \( M \) data retransmission rounds, i.e., each codeword is (re)transmitted a maximum of \( M + 1 \) times. The original codeword is constructed by encoding \( Q_0 \) information nats into a codeword of length \( L_0 \), \( L_0 \ll L_c \) and rate \( R_0 = \frac{Q_0}{L_0} \). The codeword is rescaled in each (re)transmission round to have power \( P_m, m = 1, \ldots, M + 1 \). Therefore, representing the codeword transmitted at the \( m \)-th (re)transmission round by \( \{X_m[i], i = 1, \ldots, L_0\} \) we have

\[
P_m = \frac{1}{L_0} \sum_{i=1}^{L_0} |X_m[i]|^2.
\]

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 \( \frac{R_0}{m} \). The (re)transmission continues until the codeword is correctly decoded by the receiver or the maximum number of retransmission rounds is reached. In this way, the results of, e.g., [3]–[6], can be used to show that the achievable rates random variable, i.e., \( R(g) \), is

\[
R(g) = \begin{cases} \frac{R_0}{m}, & \log(1 + g \sum_{n=1}^{m-1} P_n) < R_0 \leq \log(1 + g \sum_{n=1}^{m} P_n) \\ 0, & R_0 > \log(1 + g \sum_{n=1}^{M+1} P_n) \end{cases}
\]

Consequently, the system throughput in the continuous communication model is obtained by

\[
\eta = \sum_{m=1}^{M+1} \frac{R_0}{m} \Pr\left\{ \log(1 + g \sum_{n=1}^{m-1} P_n) < R_0 \right\}
\]

\[
= \sum_{m=1}^{M+1} \frac{R_0}{m} \left( 1 - F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{m-1} P_n} \right) \right).
\]

Moreover, the outage, defined as the event that the transmitted data is undecodable by the receiver after \( M + 1 \) (re)transmission rounds, happens if and only if \( R_0 > \log(1 + g \sum_{n=1}^{M+1} P_n) \). Therefore, the system outage probability is

\[
\pi = \Pr\{ R_0 > \log(1 + g \sum_{n=1}^{M+1} P_n) \} = F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{M+1} P_n} \right).
\]

The total transmission power up to the end of the \( m \)-th (re)transmission round is

\[
p(m) = \frac{1}{m L_0} \sum_{n=1}^{m} L_0 \left| X_n[i] \right|^2 = \frac{1}{m} \sum_{n=1}^{m} P_n.
\]

Therefore, the average transmission power is determined by

\[
P = \sum_{m=1}^{M+1} \frac{1}{m} \sum_{n=1}^{m} P_n \Pr\left\{ \log(1 + g \sum_{n=1}^{m-1} P_n) < R_0 \right\}
\]

\[
= \sum_{m=1}^{M+1} \frac{1}{m} \sum_{n=1}^{m} P_n \left( 1 - F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{m-1} P_n} \right) \right)
\]

Finally, considering (13) and (16), the power-limited throughput optimization problem can be stated as

\[
\max \left\{ R_0, P_{m, m=1, \ldots, M+1} \right\}
\]

subject to

\[
\sum_{m=1}^{M+1} \frac{R_0}{m} \left( 1 - F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{m-1} P_n} \right) \right) \leq P
\]

which based on the fading pdf can be solved numerically or analytically.

IV. SYSTEM THROUGHPUT UNDER BURSTING COMMUNICATION ASSUMPTION

If the receiver can decode the data at any retransmission round, the reward is \( R_0 \). Therefore, with a maximum of \( M + 1 \) (re)transmission rounds, the expected reward is

\[
E\{\tilde{R}\} = R_0 \left( 1 - \Pr\left\{ R_0 > \log(1 + g \sum_{n=1}^{M+1} P_n) \right\} \right)
\]

\[
= R_0 \left( 1 - F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{M+1} P_n} \right) \right).
\]

On the other hand, \( m \) time slots are spent if the data is decoded in the \( m \)-th (re)transmission round. Also, independent of the message decoding status, there will be \( M + 1 \) time slots if all possible retransmission rounds are used. In this way, the inter-renewal time is obtained by

\[
E\{\tilde{T}\} = \sum_{m=1}^{M+1} m \Pr\left\{ \log(1 + g \sum_{n=1}^{m-1} P_n) < R_0 \right\}
\]

\[
= 1 + \sum_{m=2}^{M+1} m \, F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{m-1} P_n} \right).
\]

and the throughput in the bursting model is rephrased as

\[
\eta = \frac{R_0 \left( 1 - F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{M+1} P_n} \right) \right)}{1 + \sum_{m=2}^{M+1} m \, F_G\left( \frac{e^{R_0} - 1}{\sum_{n=1}^{m-1} P_n} \right)}.
\]

The outage probability is determined with the same procedure as before. However, as the transmission energy at the end of the \( m \)-th (re)transmission round is \( L_0 \sum_{n=1}^{m} P_n \), the expected
energy in a packet transmission period and the average power are obtained by
\[
E\{\xi\} = \sum_{m=1}^{M+1} (L_0 \sum_{n=1}^{m-1} P_n) \left( F_G(\frac{e^{R_0} - 1}{\sum_{m=1}^{M+1} P_n}) - F_G(\frac{e^{R_0} - 1}{\sum_{m=1}^{m} P_n}) \right) + (L_0 \sum_{n=1}^{M+1} P_n) F_G(\frac{e^{R_0} - 1}{\sum_{m=1}^{M+1} P_n})
\]
and
\[
\varphi = \frac{E\{\xi\}}{E\{\tau\}} = \frac{P + \sum_{m=2}^{M+1} P_m F_G(\frac{e^{R_0} - 1}{\sum_{m=1}^{M+1} P_n})}{1 + \sum_{m=2}^{M+1} F_G(\frac{e^{R_0} - 1}{\sum_{m=1}^{m} P_n})}
\]

respectively. Therefore, considering (20) and (22) the power-limited throughput optimization problem can be stated the same as in (17). Finally, it is worth noting that implementing uniform power allocation, i.e., \(P_m = P, \forall m\), (13) and (20) can be respectively rewritten as
\[
\eta = \sum_{m=1}^{M+1} \frac{R_0}{m} \left( F_G(\frac{e^{R_0} - 1}{m-1}P) - F_G(\frac{e^{R_0} - 1}{m}P) \right)
\]
and
\[
\eta = \frac{R_0(1 - F_G(\frac{e^{R_0} - 1}{(M+1)}P))}{1 + \sum_{m=2}^{M+1} F_G(\frac{e^{R_0} - 1}{(m-1)}P)}
\]

A. A simple iterative throughput optimization algorithm

Depending on the fading pdf and the number of retransmissions, it may be difficult to find the optimal transmission powers and rates analytically, particularly because, e.g., (17) is not a convex optimization problem. To tackle this problem, we propose an iterative algorithm stated in Algorithm 1.

**Algorithm 1 Throughput optimization**

I. For a given power constraint \(P\) and the fading pdf \(f_G\), consider \(J\) randomly generated vectors \(\Lambda^j = [R_0^j, P_1^j, P_2^j, \ldots, P_M^j]\) such that \(R_0^j > 0\) and \(P_m^j > 0, m = 1, \ldots, M\).

II. For each vector, do the following procedures

1) Determine the last retransmission power \(P_{M+1}^j\) according to (16) or (22). If \(P_{M+1}^j < 0\), eliminate the \(j\)-th vector.

2) Determine the throughput \(\eta^j\) based on (13) or (20).

III. Find the vector which results in the highest throughput, i.e., \(\Lambda^j\) where \(\eta^j \leq \eta^i, \forall j = 1, \ldots, J\).

IV. \(\Lambda^1 \leftarrow \Lambda^j\).

V. Generate \(b \ll J\) vectors \(\Lambda^{j,\text{new}}, j = 1, \ldots, b\) around \(\Lambda^1\). These vectors should also satisfy the constraints introduced in I.

VI. \(\Lambda^{j+1} \leftarrow \Lambda^{j,\text{new}}, j = 1, \ldots, b\).

VII. Regenerate the remaining vectors \(\Lambda^j, j = b + 2, \ldots, J\) randomly such that \(\Lambda^j = [R_0^j, P_1^j, P_2^j, \ldots, P_M^j]\) with \(R_0^j > 0\) and \(P_m^j > 0, m = 1, \ldots, M\).

VIII. Go to II and continue until convergence.

Although time-consuming, the proposed algorithm has been shown to be efficient in complex optimization problems dealing with large number of optimization parameters [20]. Also, as seen in the following, it can be easily adapted when considering other quality-of-service requirements.

V. SIMULATION RESULTS AND DISCUSSIONS

Simulation results are presented for Nakagami-2 fading channels \(f_G(g) = \frac{g^{K-1}e^{-\frac{g}{2}}}{\pi^\frac{K}{2} K\gamma^K}, g \geq 0\) where we set \(w = 1.3\). Considering a maximum of \(M = 1\) retransmission round, Fig. 1a evaluates the effect of optimal power allocation on the system performance in the continuous and the bursting communication models. Here, the results are presented using the throughput gain parameter which is defined as
\[K = \frac{\eta}{\eta_{\text{no}}}\]
i.e., the ratio of the HARQ-based throughput and the one obtained with no-CSIT. Then, Fig. 1b shows the optimal transmission powers \(P_1\) and \(P_2\) as a function of the average transmission power constraint \(P\). Here, the results emphasize that:

- Even with a single retransmission round, considerable throughput increment is achieved via optimal power allocation within the retransmission rounds. However, the effect of power allocation diminishes at high SNRs where the results converge to the ones with uniform power allocation. Moreover, in all cases the system throughput is improved, compared to the no-CSIT case, i.e., \(K > 1\).

- Even with no-CSIT the optimal transmission power in the first round is higher than the power in the second round. The intuition behind this point is interesting; implementing HARQ feedback, the data is first transmitted aggressively, i.e., with high power and rate. If the channel is not good, this aggressive transmission fails, and the HARQ can save it by retransmitting the data. On the other hand, if it is lucky that the channel is strong enough, such gambling brings high return. In this way, the strong channel realizations are (almost) fully exploited, i.e., there is (almost) no waste of the good channel. Finally, although not seen in the figure, the simulations indicate that the same conclusion is valid when the number of retransmission rounds increases.

- Using HARQ, the throughput increment in the continuous model is higher than the one in the bursting communication model. The intuition behind this point is that in the continuous model the good channel realizations are more efficiently exploited than in the bursting.

One of the most important applications of HARQ protocols is to reduce the outage probability. Therefore, as an additional quality-of-service constraint, it is interesting to add an outage probability constraint \(P \leq \gamma\) to the standard throughput optimization problem, e.g., (17). To evaluate the effect of this constraint, we can rephrase the second step of the proposed algorithm as

II. For each vector, do the following procedures

1) Determine the last retransmission power \(P_{M+1}^j\) according to (16) or (22). If \(P_{M+1}^j < 0\), eliminate the \(j\)-th vector.

As discussed in [21], Nakagami pdfs are appropriate models for the cases where the fading is not so severe.
2) Find the outage probability $\pi^j$ based on (14). If $\pi^j > \gamma$, eliminate the $j$-th vector.

3) Determine the throughput $\eta^j$ based on (13) or (20). Considering fixed average transmission power constraint $P = 1$, Fig. 2 studies the effect of an outage probability constraint on the system throughput under different uniform and optimal power allocation conditions. Here, the simulations indicate that:

- Although there is a significant performance degradation in the presence of hard outage probability constraints, the degradation diminishes when the constraint gets more relaxed.
- The HARQ-based system throughput is less affected by the outage probability constraint, when compared with the case where there is no feedback to the transmitter.

VI. CONCLUSION

This paper studied the effect of power allocation on the throughput of communication setups utilizing RTD HARQ protocol. Considering slowly-fading channels, the results were presented for both continuous and bursting communication models. Moreover, we evaluated the effect of an outage probability constraint on the system data transmission efficiency. Simulation results show considerable throughput increment via optimal power allocation within the HARQ-based retransmission rounds. Also, the optimal transmission powers maximizing the throughput are found to be decreasing in every retransmission round. Finally, compared to the case where there is no feedback to the transmitter, the system performance is less affected by an outage probability constraint when implementing an HARQ protocol.

REFERENCES


