
Eric Jonsson School of Engineering and Computer Science

2004-7

Single-Source Single-Relay Cooperative ARQ Protocols in TDM Radio Networks

Isabella Cerutti, *et al.*

For more information about this item go to: <http://hdl.handle.net/10735.1/2682>

**Single-Source Single-Relay Cooperative
ARQ Protocols in TDM Radio Networks**

Isabella Cerutti, Andrea Fumagalli, and Puja Gupta

**Technical Report UTD/EE-01/2006
April 2006**

Single-Source Single-Relay Cooperative ARQ Protocols in TDM Radio Networks

Isabella Cerutti, Andrea Fumagalli, and Puja Gupta

Abstract

In conventional (non-cooperative) automatic repeat request (ARQ) protocols for radio networks, the corrupted data frames that cannot be correctly decoded at the destination are retransmitted by the source. In cooperative ARQ protocols, data frame retransmissions may be performed by a neighboring node (the relay) that has successfully overheard the source's frame transmission. One advantage of the latter group of ARQ protocols is the spatial diversity provided by the relay.

The first delay model for cooperative ARQ protocols is derived in this paper. The model is analytically derived for a simple set of retransmission rules that make use of both uncoded and coded cooperative communications in slotted radio networks. The model estimates the delay experienced by Poisson arriving frames, whose retransmissions (when required) are performed also by a single relay. Saturation throughput, data frame latency, and buffer occupancy at both the source and relay are quantified and compared against two non-cooperative ARQ protocols.

I. INTRODUCTION

Wireless networks are enjoying a widespread diffusion thanks to a variety of solutions that are increasingly deployed in the field, e.g., cellular, ad-hoc, wireless LAN, and sensor networks. The main drive behind this expansion is the virtually endless list of enabled applications [1], [2].

One peculiar characteristic of the radio medium is its inherent broadcast nature. Beside the intended destination, a signal transmitted by a source may be received also by other neighboring nodes that are within earshot. Traditionally, this phenomenon is treated as interference, i.e., the received signal is discarded by the nodes that are not the intended destination.

The quality of the signal received by the destination (and other nodes) depends on various factors, e.g., path loss, fading, and noise. Automatic repeat request (ARQ) protocols are used to guarantee reliable data delivery over the radio channel. The ARQ protocol specifies how data frames that are not successfully received and decoded by the destination must be retransmitted until they are successfully delivered. Most of the available ARQ protocols require the source to retransmit the data frames unsuccessfully delivered to the destination [3]. As other neighboring nodes do not take part in the data frame retransmission process, these ARQ protocols are referred to as *non-cooperative*.

This research was supported in part by NSF Grants No. ANI-0082085, ECS-0225528, CNS-0435429, and the Italian Ministry of University (MIUR) under FIRB project "Enabling platforms for high-performance computational grids oriented to scalable virtual organizations" (contract n. RBNE01KNFP).

I. Cerutti is with Center of Excellence for Communication Networks Engineering (CEIRC), Scuola Superiore Sant'Anna in Pisa, Italy (Email: isabella.cerutti@sssup.it). A. Fumagalli and P. Gupta are with OpNeAR Laboratory at Erik Jonsson School of Engineering and Computer Science, The University of Texas at Dallas, in Richardson, TX, USA (Emails: {andrea.fumagalli, pkg021000}@utdallas.edu).

Cooperative ARQ ($C - ARQ$) protocols permit nodes — other than the source and the destination — to actively help deliver the data frame correctly. The rationale is that a node(s) which is within earshot from the source and the destination may cooperate. This node is referred to as the *relay*. The relay makes use of the received signal from the source to improve the overall capacity of the source-destination radio channel as follows. When the source's data frame transmission is not successful, the relay is invited to take part in the frame retransmission process. By doing so, the destination can rely on data frames that are transmitted by both the source and the relay, possibly yielding a better overall reception quality. The essence of the idea lies in that the destination benefits from data frames arriving via two statistically independent paths, i.e., spatial diversity.

Focusing on cooperative communications, it must be noted that a number of results has been published on this topic. A recent survey on cooperative radio communications can be found in [4]. Initial work on cooperative communications on the Gaussian relay channel is reported in [5]. The relay's role is to assist one source, i.e., *single-source cooperation*. More recent works [4], [6], [7], [8], [9], [10] have extended the concept of cooperative communications by taking into account fading and allowing two sources to cooperate with one another at the same time. This case is referred to as *double-source cooperation*, whereby each source interleaves the transmission of its own data frames with the retransmission of the other source's data frames. These works can be divided into three categories, according to the method used to realize cooperative communications [4]. In *detect-and-forward* methods [8], [9], the relay detects and retransmits the frame whenever possible. In *amplify-and-forward* methods [6], the relay amplifies the received signal and retransmits it. Both of these methods use retransmission of the exact copy of the data frame. In *coded cooperation* methods [11], [7], cooperation is achieved in the framework of channel coding. The approach in [11], [7] shows the feasibility of coded cooperation and evaluates the benefits, in terms of reduction of bit and frame error probability, when using various codes and cooperation levels.

Most of these results focus on the physical layer aspects of cooperation. Only few works have considered the related ARQ protocol aspects [10], [12], [13], [14]. In [10], the signal-to-noise ratio (SNR) gain and average number of retransmissions of a single-source cooperative ARQ protocol is studied. In [12], the performance of different cooperative protocols is derived in terms of outage probability and SNR gain and compared against non-cooperative protocol performance. In [14], both the saturation throughput and latency of three double-source cooperative ARQ protocols are studied. In [13], a relaying protocol for multiple relays, operating over orthogonal time slots, is proposed as a generalization of hybrid ARQ protocols. Throughput, energy consumption, and outage probability of the proposed protocol are compared against multihop protocol performance. These studies are based on event-driven simulations. Analytically derived delay models for cooperative ARQ protocols are not available.

The objective of this paper is to present the first delay model for single-source and single-relay cooperative ARQ ($C - ARQ$) protocols. Two protocols are considered that belong to the *stop and wait* family combined with either *type I* or *type II* hybrid retransmission techniques. A simple set of retransmission rules are defined to minimize signaling and control overhead in the network, hardware and algorithm complexity at the nodes, and changes required in the existing ARQ protocols to introduce cooperation.

The delay model is derived assuming a collision-free slotted radio network in which multiple

nodes share the radio medium using time division multiple access (TDMA). It is assumed that for a given source-destination pair only one preselected relay is used. The channel characteristics are fully known, i.e., bit and frame error probability. Data frames are generated at the source using a Poisson arrival process. The delay model is derived using the second moment of the service time.

With the derived delay model some of the advantages and disadvantages of cooperative ARQ protocols may be quantified. A performance gain analysis and a case study are presented in Section V and VI, respectively. The performance of the two cooperative ARQ protocols is compared against two non-cooperative ARQ protocols, i.e., type I hybrid-ARQ ($H - ARQ$) [15] and type II $H - ARQ$ [16], [17]. Saturation throughput, expected frame latency, and expected buffer occupancy at both the source and relay are evaluated and compared. Various scenarios of offered load, radio channel attenuation, and geographical distribution of the nodes are considered. The numerical results obtained from the delay model help illustrate under what conditions the cooperative ARQ protocols yield superior network performance. It is expected that the derived delay model may provide useful insights into the use of cooperative ARQ protocols in a number of wireless TDMA solutions. These include ETSI HIPERLAN/2 [18], IEEE 802.16 (WiMax) [19], and TDMA protocols over IEEE 802.11 (WiFi) [20].

II. SYSTEM ASSUMPTIONS

In describing and studying the cooperative ARQ protocols, the following simplifying assumptions are made. Relaxation of these assumptions does not alter the fundamental behavior of the protocols and is omitted in the paper to maintain the protocol description simple.

Consider a network, in which M sources are in close proximity and share the same channel frequency using time division multiplexing. Due to the nodes' proximity, the radio signal propagation time is considered to be negligible when compared to the data frame transmission time. Time is divided into time frames. Each time frame is divided into M time slots. Each node is scheduled to transmit in a given time slot. During a time slot, a data frame and its acknowledgment control frame (ACK) are transmitted by the scheduled source and destination, respectively. Data frames carry some degree of redundancy to enable error detection and correction at the receiver. Transmitted data frames are always received. However, the data frame payload may not be decoded correctly at the receiver due to transmission errors. For this reason, an ARQ protocol must be used to provide reliable delivery of data frames to the destination.

TABLE I
ARQ PROTOCOL SUMMARY

ARQ Protocol Type	Without incremental redundancy transmissions	With incremental redundancy transmissions
Without cooperation	Type I $H - ARQ$	Type II $H - ARQ$
With cooperation	Type I $C - ARQ$	Type II $C - ARQ$

III. THE ARQ PROTOCOLS

Four stop and wait ARQ protocols (Table I) are defined in this section. The two non-cooperative ARQ protocols are used in the performance benchmarking of the two cooperative ARQ protocols.

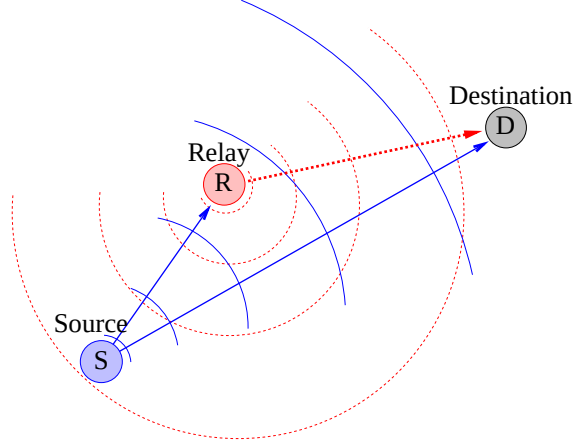


Fig. 1. Three cooperating nodes

In describing the four ARQ protocols it is assumed that source S is scheduled to transmit to destination D in the first time slot of each time frame. In the cooperative ARQ protocols, node R is the only relay assigned to help S (single-relay)¹. This case is illustrated in Fig. 1. Any of the $M - 1$ nodes (other than S), that are scheduled to transmit during the time frame, may be chosen to be R , as long as it can “eavesdrop” the transmission from S to D . While the explicit selection of R is not addressed in the paper, the performance model that is derived may be adopted to help make such selection in practice. Also, it is assumed that while cooperation takes place, R does not generate traffic of its own. It simply helps S deliver its own data frames to D successfully (single-source).

Simple retransmission rules for the ARQ protocols are chosen to contain both protocol and signaling complexity. In addition, this choice facilitates the assessment of the resulting benefits.

In the protocol description that follows, SN indicates the data frame sequence number. RN indicates the ACK request number. IN indicates the sequence number of the incremental redundancy frame, i.e., the frame carrying additional redundancy bits for data frame SN=IN.

A. Two Non-Cooperative ARQ Protocols

In this section, two non-cooperative ARQ protocols are briefly described, i.e., type I and type II $H - ARQ$. The two non-cooperative protocols are used in Sections V and VI to assess the performance gain of the cooperative ARQ protocols described in Section III-B.

In the non-cooperative ARQ protocols, R is not required.

¹The case in which multiple relays assist the same source is outside the scope of the paper.

1) *Type I H – ARQ Protocol*: Type I *H – ARQ* protocol makes use of both error detection and error correction capabilities. The following retransmission rules are used.

S sends the next data frame, e.g., $SN=i$. Three cases are possible. If data frame $SN=i$ is received and decoded correctly at *D*, $ACK\ RN=i+1$ is sent. At this point, a new data frame may be transmitted. If data frame $SN=i$ is received but not decoded correctly at *D*, $ACK\ RN=i$ is sent, requesting *S* to retransmit data frame $SN=i$. If data frame $SN=i$ is not received by *D*, a timeout is used at *S* to retransmit data frame $SN=i$. The timeout value is chosen to be one time frame.

2) *Type II H – ARQ Protocol*: Type II *H – ARQ* protocol follows the same retransmission rules of type I *H – ARQ* protocol. The difference is that a frame containing incremental redundancy bits is transmitted by *S* when the data frame is not received correctly by *D*. At *D*, the data frame and the incremental redundancy frame are jointly decoded for improved performance [15]. The following retransmission rules are used.

S sends the next data frame, e.g., $SN=i$. If data frame $SN=i$ is received and decoded correctly at *D*, $ACK\ RN=i+1$ is sent. At this point, a new data frame may be transmitted. If data frame $SN=i$ is received but not decoded correctly at *D*, $ACK\ RN=i$ is sent, requesting retransmission. In response, *S* sends incremental redundancy frame $IN=i$. When incremental redundancy frame $IN=i$ is received, *D* attempts to jointly decode the two combined frames, i.e., $SN=i$ and $IN=i$, as their combination produces a stronger redundancy code than each individual frame does alone. At the end of a transmission sequence, irrespective of its success or unsuccess, all frames stored at *D* are discarded. If *D* is able to jointly decode frames $SN=i$ and $IN=i$ correctly, $ACK\ RN=i+1$ is sent, and a new data frame may be transmitted. If *D* is not able to jointly decode the two frames, $ACK\ RN=i$ is sent. *S* alternates the transmission of frames $SN=i$ and $IN=i$ until $ACK\ RN=i+1$ is received. A retransmission timeout of one time frame is used at *S* to avoid deadlock.

B. Two Single-Source Single-Relay Cooperative ARQ Protocols

Type I and type II cooperative ARQ (*C – ARQ*) protocols are described in this section.

1) *Type I C – ARQ Protocol*: In type I *C – ARQ* protocol, when it is required, *R* transmits an exact replica of the data frame received from *S*. The following four sequences of data frame and ACK exchange are possible. Each sequence is described with the help of a figure.

- a) Fig. 2(a): *D* successfully receives the data frame transmitted by *S*. Data frame $SN=i$ is transmitted by *S* and acknowledged by *D* with the transmission of $ACK\ RN=i+1$. In the next time frame, *S* may transmit data frame $SN=i+1$.
- b) Fig. 2(b): *D* successfully receives the data frame with the help of *R*. Data frame $SN=i$ is transmitted by *S*. It is not correctly decoded by *D*. However, it is correctly received and decoded by *R*. *D* sends a (re)transmission request to *R* using $ACK\ RN=i$. Then *R* transmits data frame $SN=i$. The frame is correctly received by *D*, which sends control frame $RN=i+1$ to *S*. In the next time frame, *S* may begin a new sequence and transmit data frame $SN=i+1$.
- c) Fig. 2(c): *D* does not receive successfully the data frame due to some transmission error(s) detected in the data frame from *R*. This sequence begins in a way similar to the previous one. This time, however, data frame $SN=i$ transmitted by *R* is not correctly received by *D*. *D* sends $ACK\ RN=i$ to *S* which begins a new transmission sequence of data frame $SN=i$.

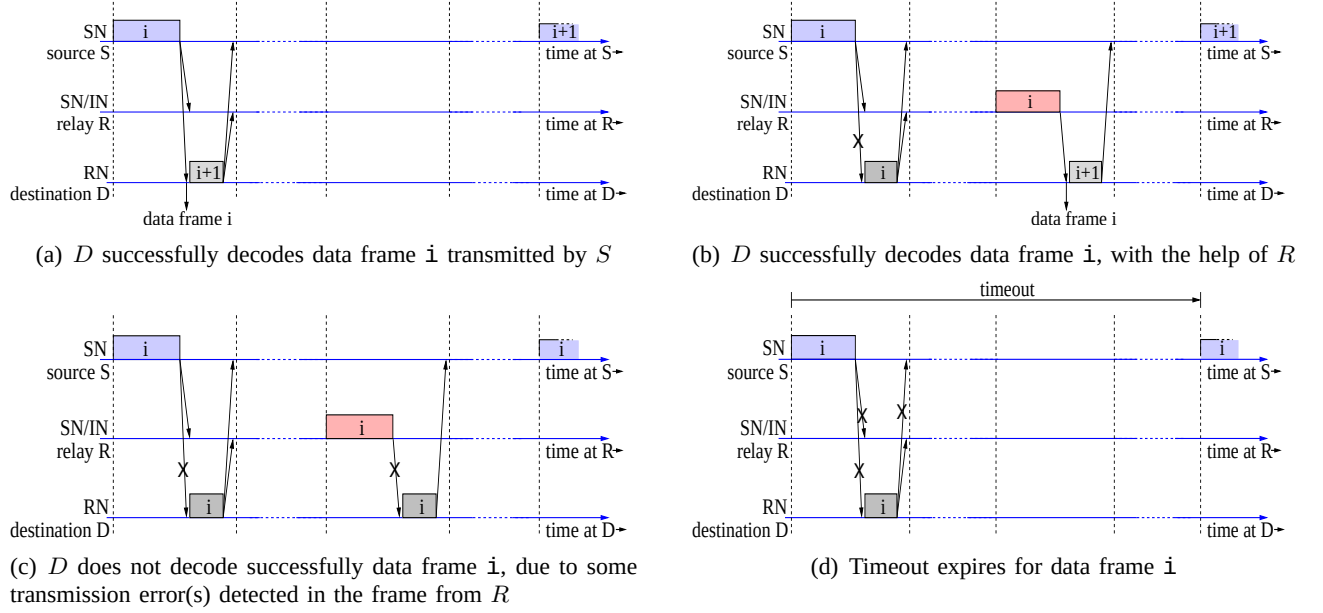


Fig. 2. Time chart

- d)** Fig. 2(d): timeout expires. For various reasons, S may not receive ACK from D . In this case, a timeout is used at S to avoid deadlock. In the example shown, data frame $SN=i$ is transmitted by S . It is not successfully decoded by D . D sends a (re)transmission request to R using $ACK\ RN=i$. However, not even R is able to decode successfully the data frame transmitted by S . Thus, it cannot cooperate, and the request from D is discarded. Upon expiration of the timeout, S begins a new transmission sequence of data frame $SN=i$. Observing that the transmission sequence lasts up to one entire time frame, a timeout of one time frame is chosen.

Recall that both S and R are assigned one distinct time slot in every time frame. For each transmission attempt made by S , R may help with its own transmission attempt. Thus, R needs not keep a copy of the data frame from S for more than one time frame.

2) Type II $C - ARQ$ Protocol: The type II $C - ARQ$ protocol follows the same retransmission rules of type I $C - ARQ$ protocol that are shown in Figs. 2(a)-2(d). The difference is that the frame transmitted by R contains incremental redundancy bits, instead of the exact replica of the data frame transmitted by S . Incremental redundancy frame $IN=i$ is computed at R , after receiving and decoding data frame $SN=i$ from S successfully. When data frame $SN=i$ from S is not decoded successfully at R , the latter cannot cooperate (sequence c). The decoding procedure at D is the same as the one used in type II $H - ARQ$ protocol. However, this time S and R alternate their transmissions, each transmitting frame $SN=i$ and $IN=i$, respectively, until $ACK\ RN=i+1$ is received. A retransmission timeout of one time frame is used at S to avoid deadlock.

For each of the four ARQ protocols, a delay model is derived in the next section.

IV. ANALYTICAL FRAMEWORK

This section derives the queueing model that is used to estimate the value of the following parameters for the four ARQ protocols described in Section III:

- Th : saturation throughput, defined as the maximum expected number of data frames delivered per time frame,
- T : latency, defined as the expected time elapsed between the data frame generation at S and the end of the time slot in which the data frame is correctly decoded at D ,
- N : expected buffer occupancy at S ,
- $\hat{N}$: expected buffer occupancy at S at the beginning of the time frame, and
- N_R : expected buffer occupancy at R .

In the derivation, it is assumed that data frame (job) arrivals at S constitute a continuous-time Poisson process of rate λ . At S , the transmitter buffer capacity is unbounded. Service time is slotted, i.e., service may initiate only in the scheduled time slot. The following probabilities are independent, identically distributed, and known:

- $\overline{P1}_{i,j}$: expected probability that node j does not successfully decode the data frame received from node i ,
- $\overline{P2}_{j,D}$: expected probability that D does not successfully decode the data frame (from S) jointly with the incremental redundancy frame received from node j ,
- any frame (data and ACK) is received with probability 1,
- ACK frame is decoded correctly with probability 1.

With the above assumptions timeouts are not required. These assumptions can be relaxed and timeouts can be introduced requiring only minor changes in the model. Let T_F be the duration of the time frame.

A. Non-Cooperative ARQ Protocols

Recall that in the non-cooperative protocols, S sends data frames to D , without the help of R .

1) *Type I H-ARQ*: The queueing model for type I $H-ARQ$ is shown in Fig. 3. Poisson arriving jobs are stored in the queue, which has infinite capacity. Each job represents a data frame generated and stored at S . A single server attends the arriving jobs, one at a time in first come first serve fashion. The server represents the transmission of the data frame, which takes place in the time slot assigned to S . To take TDMA access into account, the server has two states, i.e., active and rest. Cyclically, the server spends a time equal to T_F/M in the active state, and a time equal to $T_F(M-1)/M$ in the rest state. Service can be rendered only in the active state. The head of the line job goes into service when the server enters the active state. Service time is T_F/M . At the end of the service time, the job either stays in the queue for another round of service with probability $P = \overline{P1}_{S,D}$, or departs from the queue with probability $(1 - P)$.

Let random variable K represent the total number of service rounds that are required by

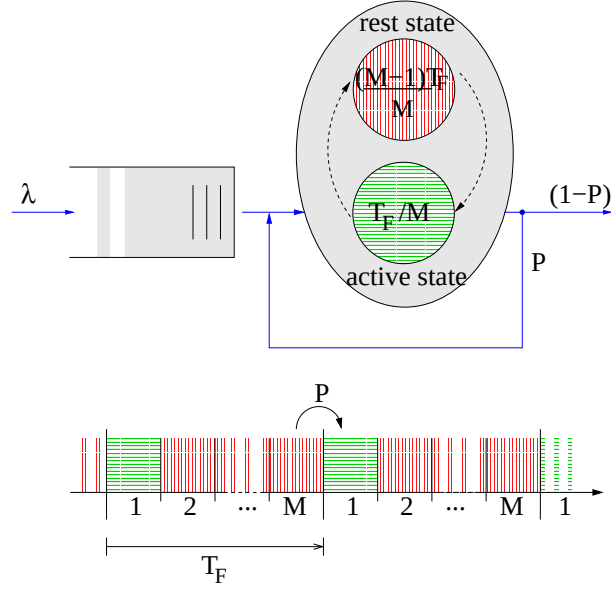


Fig. 3. Queueing model for type I $H - ARQ$

any given job. Then,

$$\bar{K} = \sum_{k=0}^{\infty} k(1-P)(P)^{k-1} = \frac{1}{1-P} \quad (1)$$

$$\bar{K^2} = \sum_{k=0}^{\infty} k^2(1-P)(P)^{k-1} = \frac{(1+P)}{(1-P)^2}. \quad (2)$$

The expected waiting time, i.e., the time elapsed between the job arrival time and the beginning of the job first service time, can be found using Pollaczek-Khinchin formula for the M/G/1 queue with vacations, i.e.,

$$W = \frac{\lambda \cdot \bar{K^2} \cdot T_F^2}{2(1-\rho)} + \frac{T_F}{2} = \left(\lambda \cdot T_F \frac{1+P}{(1-\rho)(1-P)^2} + 1 \right) \frac{T_F}{2} \quad (3)$$

with stability condition

$$\rho = \lambda \cdot \bar{K} \cdot T_F < 1. \quad (4)$$

Note that vacation intervals account for the time that arriving jobs have to wait till the beginning of the next time frame.

It is straightforward then to derive the following system parameters: the job's sojourn time in the queue

$$T = W + \left(\bar{K} - \frac{M-1}{M} \right) T_F \quad (5)$$

the time average number of jobs in the queue (including the one in service)

$$N = \lambda T \quad (6)$$

the expected number of jobs in the queue at the beginning of any time frame

$$\hat{N} = \lambda (W + \bar{K} \cdot T_F) - \frac{\lambda T_F}{2} \quad (7)$$

and the maximum departure rate

$$Th = \frac{1}{\bar{K} \cdot T_F}. \quad (8)$$

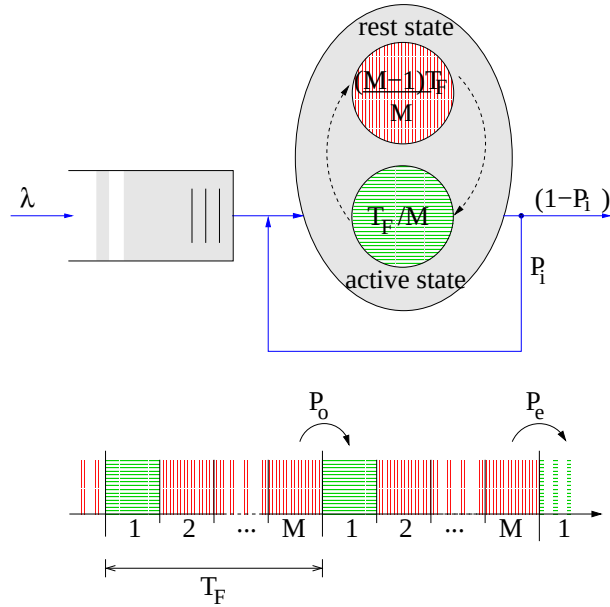


Fig. 4. Queueing model for type II $H - ARQ$

2) *Type II $H - ARQ$* : The queueing model for type II $H - ARQ$ is shown in Fig. 4. This model is the same as the type I $H - ARQ$ model described earlier in Section IV-A.1, with one difference. At the end of its i -th service round, the job either stays in the queue for another round of service with probability P_i , or departs from the queue with probability $(1 - P_i)$, where

$$P_i = \begin{cases} P_o = \overline{P1}_{S,D} & i \text{ odd} \\ P_e = \overline{P2}_{S,D} & i \text{ even.} \end{cases} \quad (9)$$

Let random variable K represent the total number of service rounds that are required by

any given job. Then,

$$\begin{aligned}\bar{K} &= \sum_{k=0}^{\infty} (2k+1)(1-P_o)(P_o \cdot P_e)^k + \sum_{k=0}^{\infty} (2k)P_o(1-P_e)(P_o \cdot P_e)^{k-1} = \\ &= 2 \sum_{k=0}^{\infty} k(P_o \cdot P_e)^{k-1} \left[(1-P_o)(P_o \cdot P_e) + P_o(1-P_e) \right] + \sum_{k=0}^{\infty} (1-P_o)(P_o \cdot P_e)^k = \frac{1+P_o}{1-P_o \cdot P_e}\end{aligned}\quad (10)$$

$$\begin{aligned}\bar{K}^2 &= \sum_{k=0}^{\infty} (2k+1)^2(1-P_o)(P_o \cdot P_e)^k + \sum_{k=0}^{\infty} (2k)^2P_o(1-P_e)(P_o \cdot P_e)^{k-1} = \\ &= \frac{1+3P_o+P_o \cdot P_e(3+P_o)}{(1-P_o \cdot P_e)^2}.\end{aligned}\quad (11)$$

Using the expressions for $\bar{K}$ and $\bar{K}^2$ in (10) and in (11), respectively, it is possible to obtain the expected waiting time from (3), the stability condition from (4), the job's sojourn time in the queue from (5), the time average number of jobs in the queue (including the one in service) from (6), the expected number of jobs in the queue at the beginning of any time frame from (7), and the maximum departure rate from (8).

B. Cooperative ARQ Protocols

Recall that in the cooperative ARQ protocols, R helps deliver S 's data frames to D by taking part in the retransmission process.

1) *Type I C-ARQ*: The model for type I $C-ARQ$ is the network of two queues shown in Fig. 5. The two mutually exclusive servers s and r represent the data frame transmission at S and R , respectively. Both servers alternate their two states, i.e., active and rest, as already described for the queue in Section IV-A.1. The two servers' cycles have a time offset of $\Delta - 1$ time slots, i.e., the time interval between the beginning of the time slots assigned to S and the beginning of the time slot assigned to R in the same time frame is $(\Delta - 1)T_F/M$. Δ is an integer value in $[2, M]$. Upon completion of service at s , a job has three options. With probability $1 - \bar{P}1_{S,D}$, it departs from the network of queues. This option represents the case depicted in Fig. 2(a). With probability $\bar{P}1_{S,D} \cdot \bar{P}1_{S,R}$, it chooses to request another round of service from s . This option represents the case depicted in Fig. 2(c). With probability $\bar{P}1_{S,D} \cdot (1 - \bar{P}1_{S,R})$, it moves to r . This option represents the case when R makes a transmission attempt of the data frame. Upon completion of service at r , the job has two options. With probability $1 - \bar{P}1_{R,D}$, it departs from the network of queues. This option represents the case depicted in Fig. 2(b). With probability $\bar{P}1_{R,D}$, it moves back to server s . This option represents the case when R 's transmission attempt is not successfully decoded at D .

Let

$$P = \bar{P}1_{S,D} (\bar{P}1_{S,R} + (1 - \bar{P}1_{S,R})\bar{P}1_{R,D}). \quad (12)$$

be the probability that upon completion of service at s , the job will return to s (with or without passing through r). Using the expression for P given in (12) it is possible to obtain the stability condition for the network of queues from (1) and (4), and the expected waiting

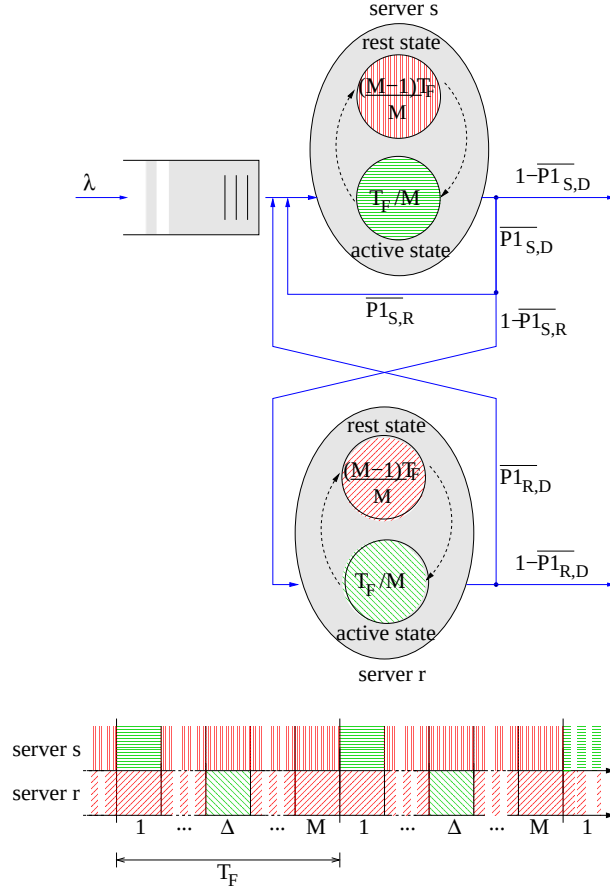


Fig. 5. Queuing model for type I $C-ARQ$

time from (2) and (3). The job's sojourn time in the network of queues is

$$T = W + \left(\bar{K} - \frac{M-1}{M} \right) T_F + \frac{\bar{P}1_{S,D}(1 - \bar{P}1_{S,R})(1 - \bar{P}1_{R,D})}{1 - P} (\Delta - 1) \frac{T_F}{M} \quad (13)$$

where $\bar{K}$ is given in (1). The term in Δ represents the additional delay incurred when the data frame is successfully delivered to D by R . The time average number of jobs in the network of queues (including the one in service) is given in (6). (Note that this quantity represents the number of data frames stored at S , which includes a copy of the data frame that is in transmission at either S or R .) The expected number of jobs in the network of queues at the beginning of a time frame is given in (7). The time average number of jobs in the queue of r is

$$N_R = \frac{\bar{P}1_{S,D}(1 - \bar{P}1_{S,R})}{1 - P} \cdot \lambda \cdot (\Delta - 1) \cdot \frac{T_F}{M}. \quad (14)$$

$\frac{N_R}{(\Delta-1)}$ represents also the utilization of r , i.e., fraction of time R is busy transmitting S 's data frames. The maximum departure rate is given in (8).

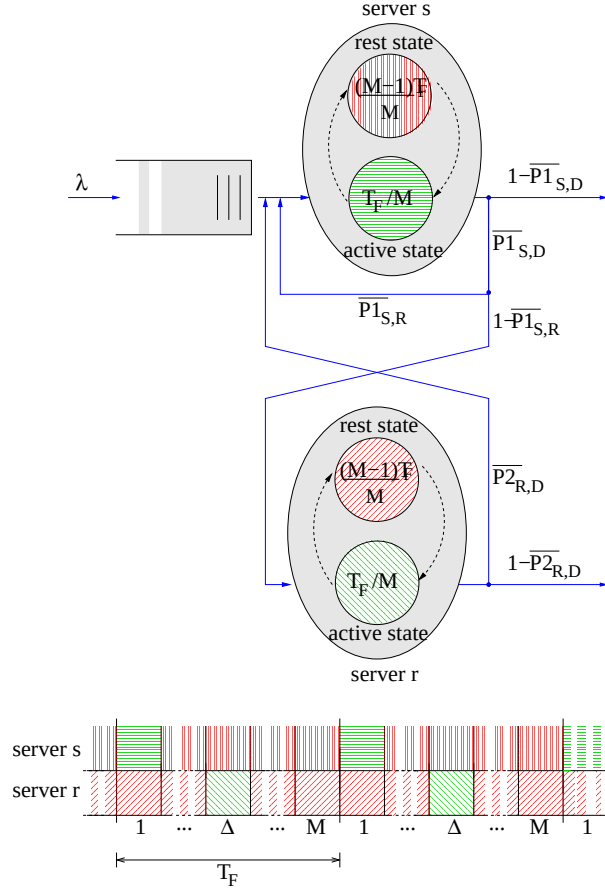


Fig. 6. Queueing model for type II $C - ARQ$

2) *Type II $C - ARQ$* : The model for type II $C - ARQ$ is the network of two queues shown in Fig. 6. This model is similar to the one described in the previous section and shown in Fig. 5, with one difference. Upon completion of service at r , the job has two options. With probability $1 - \overline{P2}_{R,D}$, it departs from the network of queues. This option represents the case depicted in Fig. 2(b). With probability $\overline{P2}_{R,D}$, it moves back to server s . This option represents the case when R 's transmission attempt is not successfully decoded a D . Recall that $\overline{P2}_{R,D}$ is the probability of unsuccessfully jointly decoding the data frame and the incremental redundancy frame at D .

All the equations in Section IV-B.1 apply to this model, by substituting $\overline{P1}_{R,D}$ with $\overline{P2}_{R,D}$.

V. SOME CONSIDERATIONS ON COOPERATION GAINS

Beside spatial diversity, performance gains when using cooperative ARQ protocols are expected for two additional reasons. First, R is contributing its bandwidth to help S deliver its data frames to D . Second, high $S-R$ and $R-D$ channel SNR helps overcome low $S-D$ channel SNR. The former reason is discussed in this section. The latter is discussed in the case study of Section VI.

To establish the performance gain originating from the fact that R is offering its allotted bandwidth to help S deliver data frames to D , two assumptions are introduced, i.e., $\overline{P1}_{S,D} =$

$\overline{P1}_{S,R} = \overline{P1}_{R,D} = p$, and $\overline{P2}_{S,D} = \overline{P2}_{R,D} = p/g$. Parameter $g \geq 1$ represents type II over type I probability gain to successfully decode S 's data frames at D . Note that in type I ARQ protocols $g = 1$. These assumptions are intended to recreate the case when S , R and D are equidistant with same channel SNR.

A. Retransmission Rate Gain

TABLE II
PROBABILITY OF RETRANSMITTING A DATA FRAME AT S

ARQ protocol	Type I	Type II
H – ARQ	p	$\frac{p(g+p)}{g(1+p)}$
C – ARQ	$p^2(2-p)$	$p^2(1 + \frac{1}{g} - \frac{p}{g})$

Table II summarizes the expected probability that S is required to retransmit its own data frame. (Recall that in type II ARQ protocols, transmission of the incremental redundancy frame may take place between two consecutive transmissions of the same data frame.) Define the data frame retransmission probability decrease (or gain) due to cooperation as

$$\begin{aligned}
 G_P(p, g) &= \frac{P(\text{type II H – ARQ})}{P(\text{type II C – ARQ})} = \frac{\frac{p(g+p)}{g(1+p)}}{p^2(1 + \frac{1}{g} - \frac{p}{g})} = \\
 &= \frac{g+p}{p(1+p)(1+g-p)}. \tag{15}
 \end{aligned}$$

Note that $G_P(p, 1) = 1/p(2-p)$ is the gain for type I. It is easy to demonstrate that $G_P(p, 1) \geq 1$ for any value of $p \in [0, 1]$, i.e., cooperation is always helpful in type I ARQ protocol. More in general, it can be demonstrated [21] that $G_P(p, g) \geq 1$ when

$$1 \leq g \leq g_z \quad \forall p \in [\frac{\sqrt{5}-1}{2}, 1] \tag{16}$$

where

$$g_z = \frac{p^3}{p^2 + p - 1}. \tag{17}$$

Note that when

$$p \leq \frac{\sqrt{5}-1}{2} \approx 0.618 \tag{18}$$

$G_P(p, g) \geq 1$ for any value of $g \geq 1$, i.e., irrespective of g value, cooperation always helps as long as $p < 0.618$. This result is explained as follows. For values of $p < 0.618$, R 's ability to successfully decode and relay the data frame received from S is sufficiently good to be always effective. For values of $p > 0.618$ and depending on the value of g , it may be more effective for S to transmit its own incremental redundancy frame (no cooperation) rather than requesting R to do so, since R is not able to help very often due to the high probability of unsuccessfully decoding the data frame received from S . High values of incremental redundancy gain can

improve the quality of the S - D channel in type II H -ARQ, and the R - D channel in type II C -ARQ. Thus the cooperation gain on the retransmission rate from S to D is less effective for high values of g .

Fig. 7 plots (15) versus p for $g = 1, 1.1, 2.5$. It is notable that the value $G_P(0.5, g) = 4/3$ is not affected by g .

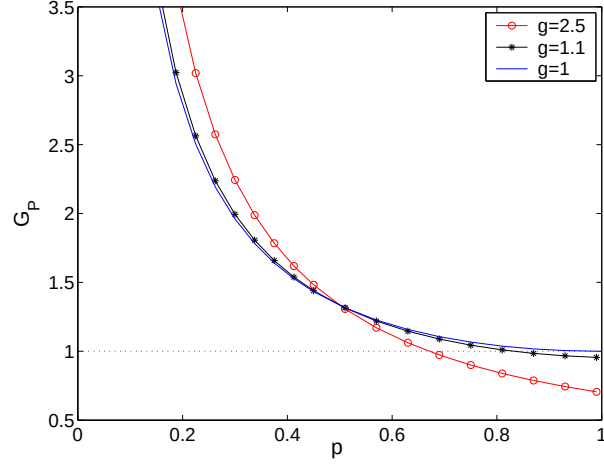


Fig. 7. $G_P(p, g)$ vs. p

B. Throughput Gain

Define the throughput increase (or gain) due to cooperation as

$$G_{Th}(p, g) = \frac{Th(\text{type II } C - ARQ)}{Th(\text{type II } H - ARQ)} = \frac{1 - p^2(1 + \frac{1}{g} - \frac{p}{g})}{1 - \frac{p(g+p)}{g(1+p)}} = \frac{(1 - p^2)(g(1 + p) - p^2)}{(g - p^2)}. \quad (19)$$

Gain $G_{Th}(p, g)$ originates from the extra bandwidth that R is making available to deliver S 's data frames to D . For type I ARQ protocol $G_{Th}(p, 1) = 1 + p(1 - p)$. It can be demonstrated that $G_{Th}(p, 1) \geq 1$ for $p \in [0, 1]$, and its maximum value is $G_{Th}(0.5, 1) = 1.25$. This result is simply explained. Lower values of p reduce the number of retransmission attempts performed by R . Higher values of p make it hard for the transmitted data frame to be received and decoded successfully at R , which is then unable to help S deliver data frames to D .

More in general, it can be demonstrated [21] that $G_{Th}(p, g)$ is monotonically decreasing in g , for any given constant value of p , i.e., $G_{Th}(p, \infty) \leq G_{Th}(p, g) \leq G_{Th}(p, 1)$ for $p \in [0, 1]$. $G_{Th}(p, g) \geq 1$ when (16) is satisfied. Once again, value g represents a break-even point between the gain originating from using R 's bandwidth to relay S 's data frames to D and the phenomenon explained earlier about high values of p affecting R 's ability to successfully decode and relay the data frame received from S .

These conclusions are confirmed by Fig. 8, which plots (19) versus p for $g = 1, 1.1, 2.5$.

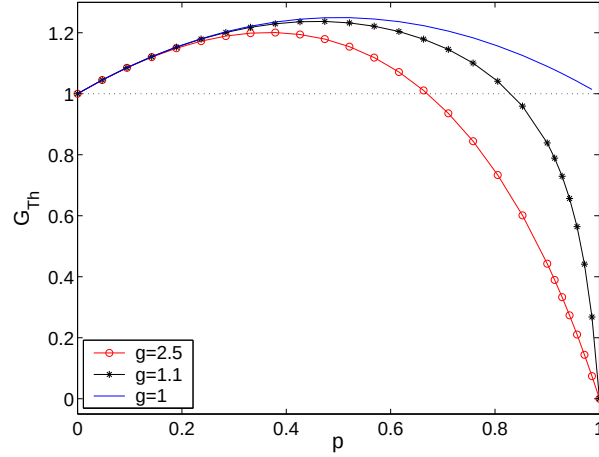


Fig. 8. $G_{Th}(p, g)$ vs. p

C. Latency Gain

The latency for type II $H - ARQ$ is

$$T(\text{type II } H-ARQ)(p, g) = \frac{T_F}{2 \left(1 - \frac{p^2}{g} - \lambda T_F(1+p)\right)(g-p^2)} \cdot \left(g(1+2p) - p^3 \left(2 + \frac{p}{g}\right) - \lambda T_F 2p^2(1-g) \right) + \frac{T_F}{M} \quad (20)$$

and for type II $C - ARQ$ is

$$T(\text{type II } C-ARQ)(p, g) = \frac{T_F(1 + p^2(1 + \frac{1}{g} - \frac{p}{g}))}{2 \left(1 - p^2(1 + \frac{1}{g} - \frac{p}{g}) - \lambda T_F\right)} + \frac{T_F}{M} + \frac{T_F(\Delta - 1)p(g - p)}{M(g(1 + p) - p^2)}. \quad (21)$$

Define the latency decrease (or gain) due to cooperation as

$$G_T(p, g) = \frac{T(\text{type II } H-ARQ)(p, g)}{T(\text{type II } C-ARQ)(p, g)}. \quad (22)$$

It can be demonstrated [21] that when $\lambda T_F \rightarrow 0$, $G_T(p, 1) \geq 1$ for $p \in [0, 1]$.

Fig. 9 plots (22) versus p for $g = 1, 1.1, 2.5$ and confirms the above conclusion. As shown in the figure, $G_T(p, 1)$ first grows, then declines slightly, to end up with a sharp increase. This behavior is the result of two factors. The first is the cooperative gain similar to the one already seen for $G_{Th}(p, 1)$ for $p < 0.5$. When this gain is about to decline ($p > 0.5$), the second factor takes over, i.e., the latency of type I $H - ARQ$ grows unbounded before it does for the type I $C - ARQ$. This is caused by the departure rate of the former protocol model that decreases and approaches the arrival rate (λ) faster than the departure rate of the latter.

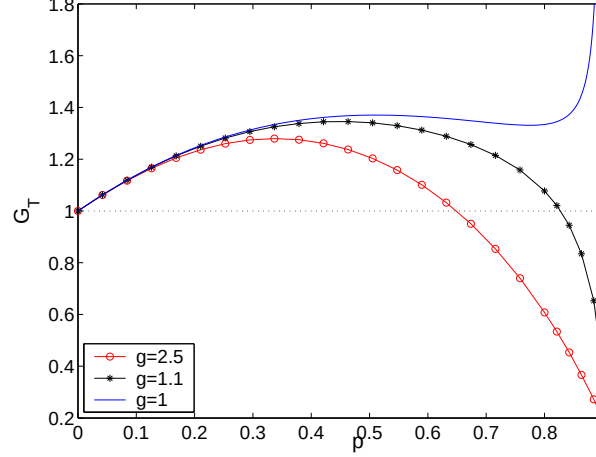


Fig. 9. $G_T(p, g)$ vs. p , $\lambda = 0.1$, $M = 8$, $\Delta = 2$

VI. A CASE STUDY

In this section a case study is presented to assess the overall performance gains when using cooperative ARQ protocols. As already mentioned, these gains originate from spatial diversity, use of R 's bandwidth to retransmit S 's data frames, and better S - R and R - D channel SNR than S - D channel SNR.

A. Assumptions on Radio Channel with Coding

The following assumptions are made on the redundancy code and the radio channel propagation properties. These assumptions are applied consistently when deriving the performance of the four ARQ protocols.

Path loss and fading affect the transmission of both data and incremental redundancy frames. Frequency-flat, block-Rayleigh fading (quasi-static) is assumed with fading level that is constant over the duration of an entire data frame transmission, i.e., time slot. The fading levels are statistically independent of the time slot, channel, and space. (These assumptions tend to favor the non-cooperative protocols, as in reality it is expected that cooperative protocols are more robust over non-cooperative protocols when fading is correlated in time, due to their spatial diversity [13].) The instantaneous SNR from node i to j is:

$$\gamma_{ij} = \frac{E_{b_i}}{N_0} \cdot K \cdot l_{ij}^{-\beta} \cdot \alpha_{ij}^2, \quad (23)$$

whereby the following definitions are used:

- E_{b_i} : transmitted energy per bit at node i ,
- N_0 : noise spectral density of the additive white Gaussian noise (AWGN) channel,
- K : path loss for an arbitrary reference distance,
- l_{ij} : distance from node i to j (normalized to the reference distance),
- β : path loss exponent,
- α_{ij} : Rayleigh distributed random variable to model the Rayleigh fading magnitude from node i to j . α_{ij}^2 has an exponential distribution with mean $E[\alpha_{ij}^2] = 1 \forall i, j$.

Each data payload is encoded into a codeword using the puncturing technique based on a rate-compatible punctured convolutional code (RCPC) [22]. The codeword is partitioned to form two frames, i.e., the data frame of N_1 bits and the incremental redundancy frame of N_2 bits. To fit the time slotted nature of the channel, $N_1 = N_2$. The N_1 bits of the data frame constitute a valid (albeit weaker) codeword. Note that before encoding, the payload is matched with a CRC code that is used at both the destination and relay to verify whether or not the received frame is decoded correctly.

The transmission error probability of a data frame sent from i to j is evaluated conservatively using the union bound technique [23], [7], i.e.,

$$P1_{i,j} \leq 1 - \left(1 - \min \left\{ 1, \sum_{d=d_f}^{\infty} a_d \cdot P(d|\gamma_{ij}) \right\} \right)^B, \quad (24)$$

whereby the following definitions are used:

- B : number of payload and CRC bits in each data frame, i.e., number of trellis branches in the codeword,
- d_f : free distance of the code [24],
- a_d : spectrum of the code [22], i.e., number of codewords of weight d ,
- $P(d|\gamma_{ij})$: probability that a wrong path at distance d is selected.

Averaging (24) over the probability density function of the instantaneous SNR, i.e., $f(\gamma_{ij})$,

$$\overline{P1}_{i,j} = \int_0^{\infty} P1_{i,j} \cdot f(\gamma_{ij}) \partial \gamma_{ij}. \quad (25)$$

Assuming that binary PSK with soft decoding is employed,

$$P(d|\gamma_{ij}) = Q\left(\sqrt{2 \cdot d \cdot \gamma_{ij}}\right), \quad (26)$$

where $Q(\cdot)$ is the Gaussian Q-function [25] and d is the weight of the codeword.

Recall that the joint decoding of both data and incremental redundancy frames takes place only when the data frame alone cannot be successfully decoded at D . The probability of not being able to decode the payload successfully after receiving the incremental redundancy frame from node $j \in \{S, R\}$ is upper bounded by:

$$P2_{j,D} \leq 1 - \left(1 - \min \left\{ 1, \sum_{d_S=d_f}^{\infty} \sum_{d_j=d_{f2}-d_S}^{\infty} a_{d_S,d_j} \cdot P(d_S + d_j|\gamma_{SD}, \gamma_{jD}) \right\} \right)^B \quad (27)$$

$$\overline{P2}_{j,D} = \int_0^{\infty} \int_0^{\infty} P2_{j,D} \cdot f(\gamma_{SD}) f(\gamma_{jD}) \partial \gamma_{SD} \partial \gamma_{jD} \quad (28)$$

whereby the following definitions are used:

- d_{f2} : free distance of the parent code [24],
- a_{d_S,d_j} : spectrum of the code, i.e., number of codewords of weight d_S in the first N_1 bits, and weight d_j in the other N_2 bits,
- $P(d_S + d_j|\gamma_{SD}, \gamma_{jD})$: probability that a wrong path at distance $d_S + d_j$ is selected, i.e.,

$$P(d_S + d_j|\gamma_{SD}, \gamma_{jD}) = Q\left(\sqrt{2d_S \cdot \gamma_{SD} + 2d_j \cdot \gamma_{jD}}\right). \quad (29)$$

B. Results

The system parameters are set as follows: $T_F = 1$, $K = 60$ dB, $\beta = 4$, $l_{S,D} = 1$, and $M = 8$. Payload and CRC comprise 128 bits that are encoded into 256 bit codewords using a rate-compatible punctured convolutional code (RCPC) with rate 1/2, parent code rate of 1/4, puncturing period of 8, memory of 4, and generator polynomials G(23,35,27,33)(octal) [22].

Unless otherwise indicated, R is at half distance between S and D , i.e., a good location for successful cooperation.

Simulation results have confidence interval values of 10% or better, at 95% confidence level. In the simulation, frame error probabilities are given by (24) and (27), using the instantaneous value of Rayleigh fading. For the analytical model, Monte Carlo integration is used to estimate the expected data frame error probabilities.

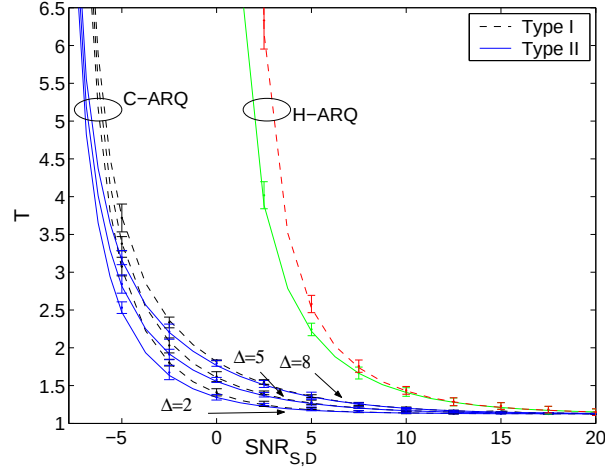


Fig. 10. T vs. $SNR_{S,D}$ (dB), $\lambda = 0.5$, $\Delta = 2, 5, 8$

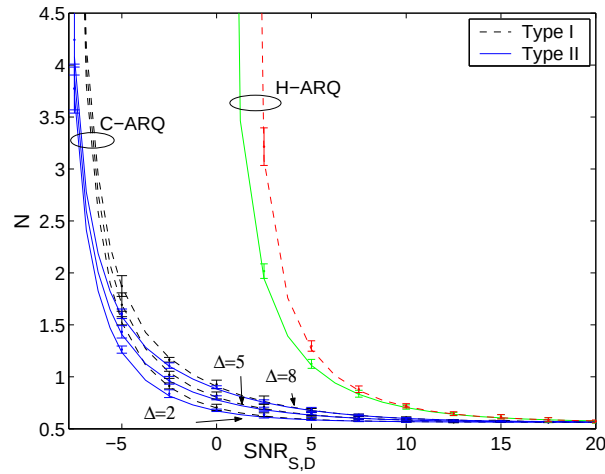


Fig. 11. N vs. $SNR_{S,D}$ (dB), $\lambda = 0.5$, $\Delta = 2, 5, 8$

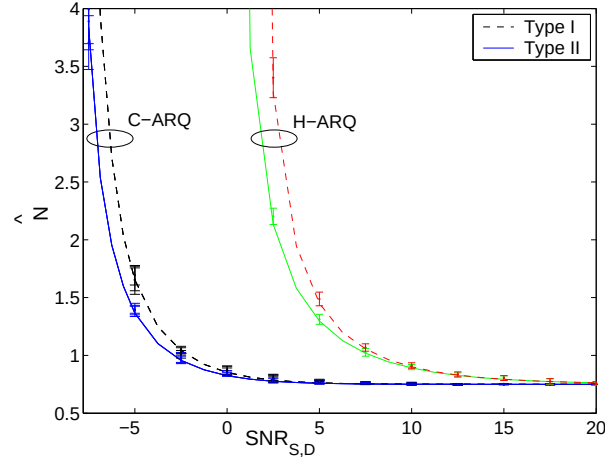


Fig. 12. $\hat{N}$ vs. $SNR_{S,D}$ (dB), $\lambda = 0.5$, $\Delta = 2, 5, 8$

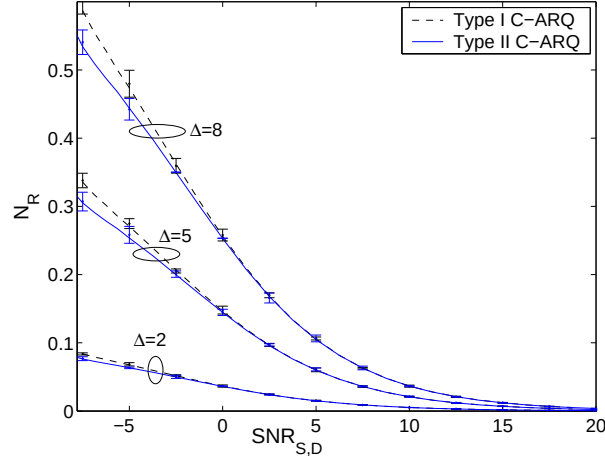


Fig. 13. N_R vs. $SNR_{S,D}$ (dB), $\lambda = 0.5$, $\Delta = 2, 5, 8$

1) *Performance Comparison of the ARQ Protocols:* Figs. 10-13 plot the expected frame latency, i.e., T , the expected buffer occupancy at S , i.e., N and $\hat{N}$, and at R , i.e., N_R , respectively, for various SNR values of the S - D channel, i.e., $SNR_{S,D}$. The arrival rate for the four ARQ protocols is $\lambda = 0.5$. For the two C -ARQ protocols, $\Delta = 2, 5$, and 8 .

Fig. 14 plots T versus λ , for the four ARQ protocols, when $SNR_{S,D} = 3$ dB.

Fig. 15 plots the saturation throughput, i.e., Th , versus $SNR_{S,D}$ for the four protocols.

In the figures, analytical and simulations results are shown with continuous curves and confidence interval bars, respectively. The good match between the analytical results and the simulation results under various conditions supports the correctness of the derived model. The plots confirm the superiority of the cooperative over non-cooperative ARQ protocols, especially in the low SNR region and for high arrival rates. Also, the known advantage of type II protocols over type I protocols is clearly documented in the plots. Finally, the plots indicate that for the cooperative protocols the best latency performance is achieved when R

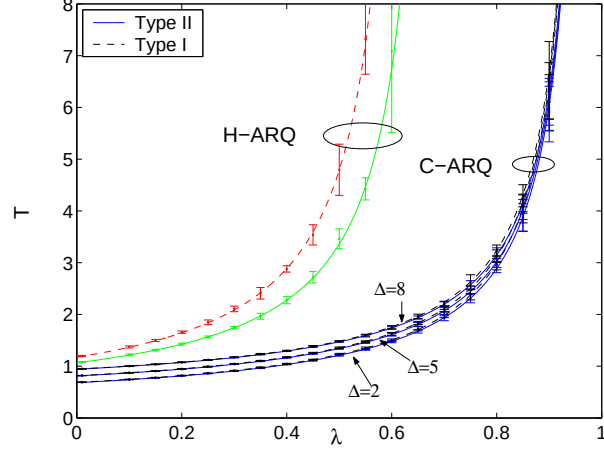


Fig. 14. T vs. λ , $SNR_{S,D} = 3$ dB, $\Delta = 2, 5, 8$

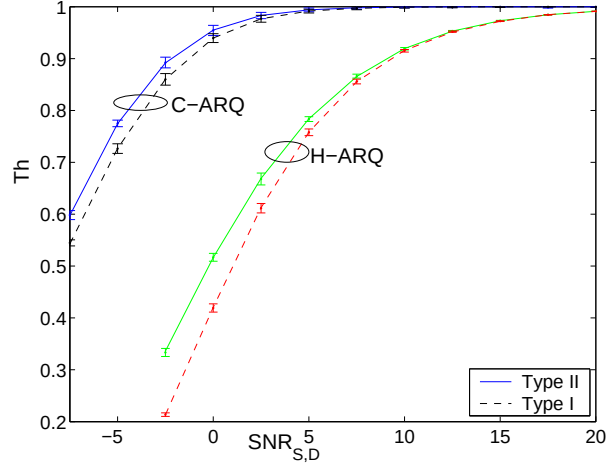


Fig. 15. Th vs. received $SNR_{S,D}$ (dB)

is scheduled to transmit in the time slot next to the one assigned to S , i.e., $\Delta = 2$.

2) *Additional Simulation Results:* Fig. 16 plots the expected data frame loss rate versus λ , for the four ARQ protocols, when the buffer capacity at S , i.e., B_S , is limited to 2 and 3 data frames. Arriving data frames are dropped when the buffer is full. $SNR_{S,D} = 3$ dB and $\Delta = 2$. Results are obtained through simulation only.

Fig. 17 plots T versus λ , when $SNR_{S,D} = 3$ dB and $\Delta = 2$, for three distinct data frame arrival processes, i.e., Poisson, compound Poisson, and deterministic. In the compound Poisson process each Poisson arrival represents one batch of data frames. The number of data frames in each batch is a random variable uniformly distributed between 1 and 3. The deterministic arrival process generates arrivals at regular intervals, the first arrival of each simulation run occurring at the time frame beginning.

The plots in both Figs. 16 and 17 confirm the superior performance of the cooperative ARQ protocols in various scenarios.

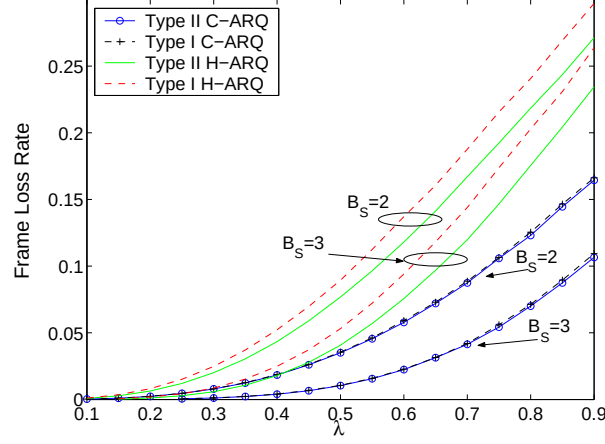


Fig. 16. Data frame loss rate vs. λ , $SNR_{S,D} = 3$ dB, $\Delta = 2$, $B_S = 2, 3$

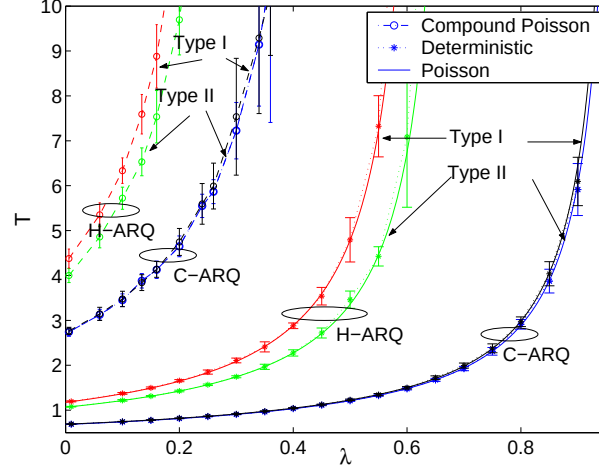


Fig. 17. T vs. λ , $SNR_{S,D} = 3$ dB, $\Delta = 2$

3) *Throughput Gain as a Function of R Position:* Fig. 18 plots the throughput gain of type II $C - ARQ$ over type II $H - ARQ$, as a function of R 's location relative to S and D . Results are obtained from the analytical model. The cooperative ARQ yields up to 72% throughput gain when compared to the non-cooperative ARQ. This gain may be even higher in the presence of lower S - D channel SNR values (in the figure $SNR_{S,D} = 0$ dB).

VII. CONCLUSION

The first delay model for single-source and single-relay cooperative stop and wait ARQ protocols for TDMA radio networks was derived in this paper. The analytical model provides results that indicate under what conditions the cooperative ARQ protocols are superior to their non-cooperative counterparts. Both frame latency and saturation throughput may be improved by using cooperative ARQ protocols. Equivalently, cooperative ARQ protocols may reduce the SNR that is required to meet the desired throughput and frame latency. These results and

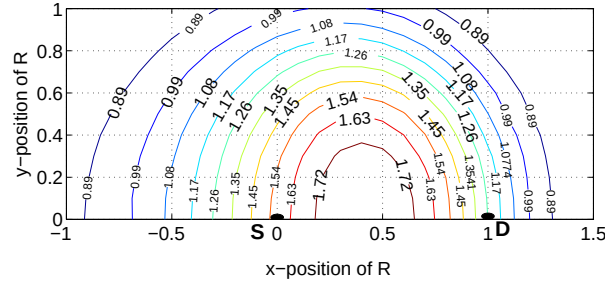


Fig. 18. Throughput gain, G_{Th} , of type II $C-ARQ$ over type II $H-ARQ$ vs. position of node R , $SNR_{S,D} = 0$ dB

the derived analytical framework may provide useful insights into the use of cooperative ARQ protocols in a number of wireless TDMA solutions. These include ETSI HIPERLAN/2 [18], IEEE 802.16 (WiMax) [19], and TDMA protocols over IEEE 802.11 (WiFi) [20]. Applications may also be found in some types of wireless sensor networks [26], where the signal power and SNR levels are limited.

Further work is required in this field to consolidate and generalize these initial findings. For example, how is double-source cooperative ARQ going to work? What happens if source and relay share the same channel and collisions may occur? An initial attempt to address some of these open questions can be found in [14]. However, much more work is required in this field.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Aria Nosratinia, Todd Hunter, and Harsh Shah for their valuable technical input on coded cooperation.

REFERENCES

- [1] P. Smyth, *Mobile and Wireless Communications: Key Technologies and Future Applications*. IEE, 2004.
- [2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Elsevier Computer Networks*, vol. 38, no. 4, pp. 393–442, March 2002.
- [3] D. Bertsekas and R. Gallager, *Data Networks (2nd ed.)*. Prentice-Hall, Inc, 1992.
- [4] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74 – 80, Oct. 2004.
- [5] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, 1979.
- [6] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," in *Proc. IEEE ISIT*, Washington, 2001, p. 294.
- [7] M. Janani, A. Hedyat, T. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 362–371, Feb. 2004.
- [8] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity–Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, 2003.
- [9] —, "User cooperation diversity–Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, 2003.
- [10] E. Zimmermann, P. Herhold, and G. Fettweis, "The impact of cooperation on diversity-exploiting protocols," in *Proc. of 59th IEEE Vehicular Technology Conference (VTC Spring)*, 2004.
- [11] T. E. Hunter and A. Nosratinia, "Cooperative diversity through coding," in *Proc. IEEE ISIT*, Lausanne, Switzerland, 2002, p. 220.
- [12] E. Zimmermann, P. Herhold, and G. Fettweis, "On the performance of cooperative relaying protocols in wireless networks," *Eur. Trans. on Telecom. (ETT)*, vol. 16, no. 1, pp. 17–35, 2005.

- [13] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 7 – 18, Jan. 2005.
- [14] P. Gupta, I. Cerutti, and A. Fumagalli, "Three transmission scheduling policies for a cooperative ARQ protocol in radio networks," in *Proc. WNCG conference*, October 2004.
- [15] S. Lin, D. Costello, and M. Miller, "Automatic-repeat-request error-control schemes," *IEEE Communications Mag.*, vol. 22, no. 12, 1984.
- [16] S. Lin and P. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels," *IEEE Trans. on Comm.*, vol. 30, no. 7, pp. 1701–1719, 1982.
- [17] Y.-M. Wang and S. Lin, "A modified selective-repeat type-II hybrid ARQ system and its performance analysis," *IEEE Trans. on Comm.*, vol. 31, no. 5, pp. 593–608, 1983.
- [18] F. M. Delicado, P. Cuenca, and L. Orozco-Barbosa, "Multiservice communications over TDMA/TDD wireless LANs," in *Proc. Wired/Wireless Internet Communications*, vol. 3510. Springer, 2005.
- [19] *IEEE Standard for Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE, 2004.
- [20] Wu-chi Feng and Wu-chang Feng, "Implementing a low power TDMA protocol over 802.11," in *Proc. Wireless Communications and Networking Conference*, March 2005, pp. 75 – 80.
- [21] I. Cerutti, A. Fumagalli, and P. Gupta, "Single-source single-relay cooperative ARQ protocols in TDM radio networks," The University of Texas at Dallas, <http://opnear.utdallas.edu/publications/reports/UTD-EE-01-2006.pdf>, Tech. Rep. UTD/EE-01/2006, 2006.
- [22] J. Hagenauer, "Rate-compatible punctured convolutional codes (rpsc codes) and their applications," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 389–400, 1988.
- [23] E. Malkamaki and H. Leib, "Evaluating the performance of convolutional codes over block fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1643–1646, 1999.
- [24] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice-Hall, 1995.
- [25] J. G. Proakis, *Digital Communications (Fourth Edition)*. McGraw-Hill International Edition, 2001.
- [26] M. Tacca, P. Monti, and A. Fumagalli, "Cooperative and non-cooperative ARQ protocols for microwave recharged sensor nodes," in *Proc. EWSN*, 2005.

APPENDIX

A. Retransmission Rate Gain

The condition

$$G_P(p, g) = \frac{g + p}{p(1 + p)(1 + g - p)} > 1 \quad (30)$$

holds when

$$p^4 - p^3g - p^2g + pg > 0. \quad (31)$$

For $p \neq 0$, the roots of the above polynomial in p , i.e., $p^3 - p^2g - pg + g$, can be found as follows. Let

$$\begin{aligned} Q &= \frac{(-2g^3 - 9g^2 + 27g)^2}{27^2 4} + \frac{(-3g - g^2)^3}{27^2} = \\ &= \frac{g^2}{27^2} \left[-\frac{135}{4}g^2 - \frac{297}{2}g + 27^2 \right]. \end{aligned} \quad (32)$$

For $g \leq -\frac{27}{5}$, $Q < 0$ and there are three real roots in:

$$p_1 = 2\rho^{1/3} \cos(\theta/3) + \frac{g}{3} \quad (33)$$

$$p_2 = -\rho^{1/3}(\cos(\theta/3) + \sqrt{3}\sin(\theta/3)) + \frac{g}{3} \quad (34)$$

$$p_3 = -\rho^{1/3}(\cos(\theta/3) - \sqrt{3}\sin(\theta/3)) + \frac{g}{3} \quad (35)$$

where ρ and θ are the magnitude and phase angle of

$$\rho e^{i\theta} = \frac{g}{2} + i\sqrt{-Q}. \quad (36)$$

The three roots, i.e., p_1 , p_2 , and p_3 , are graphically plotted in Fig. 19 for $g \geq 1$. It is possible to verify that condition (31) holds when $p > p_1$ and $p_2 < p < p_3$. Since $p_2 < 0$ and $p_1 > 1$, for the range of values of $p \in [0, 1]$ condition (31) imposes $p < p_z$ where $p_z \equiv p_3$. Fig. 20 plots the value of p_z versus g .

An alternative way to compute p_z is by iteration. Polynomial (31) can be expressed as

$$\left(1 - \frac{p}{g}\right)p^2 + p - 1. \quad (37)$$

Considering for a moment $(1 - \frac{p}{g})$ as a constant, the positive root of the quadratic expression is

$$p^+ = f(p^-) = \frac{\sqrt{\frac{5g+4p^-}{g}} - 1}{2}. \quad (38)$$

Given the desired precision, ϵ , solve (38) iteratively, i.e.,

- 1) $p^- := 1$,
- 2) $p^+ := f(p^-)$,
- 3) if $p^- - p^+ < \epsilon$ then stop,
- 4) else $p^+ := p^-$, then go to 2.

p^+ converges to p_z from above. Note that $p_z \leq f(1)$ is an upper bound that is simple to compute.

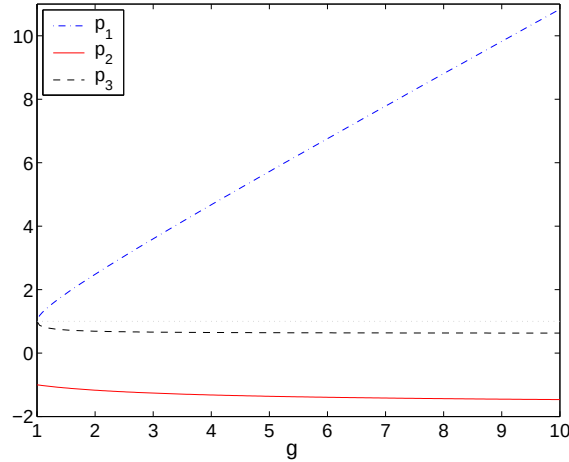


Fig. 19. p_1, p_2, p_3 vs. g

Consider the first partial derivative

$$\frac{\partial G_P(p, g)}{\partial p} = \frac{2gp^2 + 2p^3 - g^2 - g - 2g^2p}{p^2(1+p)^2(-g-1+p)^2} \leq 0 \quad \forall g \geq 1. \quad (39)$$

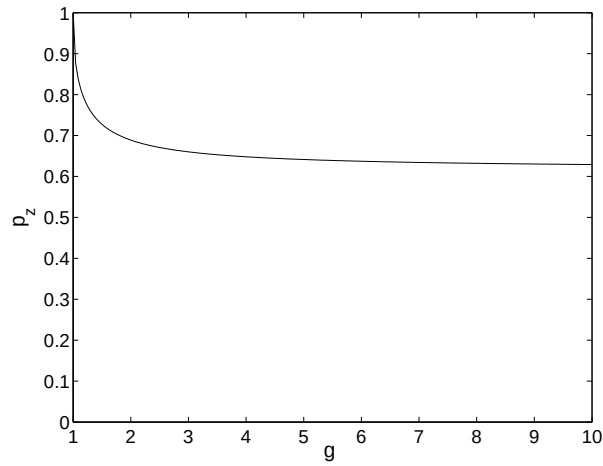


Fig. 20. p_z vs. g

Based on (39), which is always true, $G_P(p, g)$ is monotonically decreasing in p , which yields $G_P(1, g) \leq G_P(p, g) \leq G_P(0, g)$ for $g \in [1, \infty)$.

Condition (31) can also be rewritten as

$$g(-p^2 - p + 1) + p^3 > 0. \quad (40)$$

Define

$$g_z = \frac{p^3}{p^2 + p - 1}. \quad (41)$$

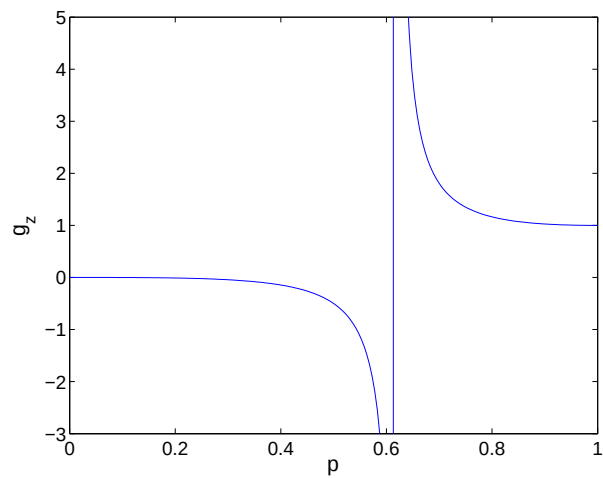


Fig. 21. g_z vs. p

Fig. 21 plots g_z as a function of $p \in [0, 1]$. As shown in the figure, g_z is positive when

$p^2 + p - 1 > 0$, i.e., for

$$p > \frac{\sqrt{5} - 1}{2}. \quad (42)$$

Thus, when $p \in [\frac{\sqrt{5}-1}{2}, 1]$, conditions (30) and (40) hold if $g < g_z$.

g_z is negative when $p \in [0, \frac{\sqrt{5}-1}{2}]$. In this interval, condition (30) and (40) hold if $g > g_z$. Since on this interval $g_z < 0$, the condition becomes $g > 0$ and always holds, as $g \geq 1$.

The first partial derivative

$$\frac{\partial G_P(p, g)}{\partial g} = \frac{2p - 1}{p(1 + p)(-g - 1 + p)^2} \quad (43)$$

is positive when the numerator is positive, i.e.,

$$p > \frac{1}{2}. \quad (44)$$

Thus, for $p \in [0, 0.5]$, $G_P(p, g)$ is monotonically decreasing in g , i.e., $G_P(p, \infty) \leq G_P(p, g) \leq G_P(p, 1)$. For $p \in (0.5, 1]$, $G_P(p, g)$ is monotonically increasing in g for $p \in [0.5, 1]$, i.e., $G_P(p, 1) \leq G_P(p, g) \leq G_P(p, \infty)$. For $p = 0.5$, $G_P(0.5, g) = \frac{4}{3}$ independently of the value of g .

B. Throughput Gain

Condition

$$G_{Th}(p, g) = \frac{(1 - p^2)(g(1 + p) - p^2)}{(g - p^2)} > 1 \quad (45)$$

holds when

$$p^4 - p^3g - p^2g + pg > 0. \quad (46)$$

Thus, condition (46) is same as (31), and the previous section considerations apply here too.

Consider the first partial derivative

$$\frac{\partial G_{Th}(p, g)}{\partial p} = \frac{-2g^2p + 4p^3g - 3g^2p^2 + 4p^4g + p^2g - 2p^5 + g^2}{(g - p^2)^2}. \quad (47)$$

It is possible to show that (47) has a maximum in

$$g_{\max} = \frac{1}{2}(2p + 3p^2 - 1) \cdot \left(1 + p^2 + 4p + \sqrt{1 + 2p^2 + 16p + p^4 - 16p^3}\right) \quad (48)$$

and that $g_{\max} \geq 1$ when p is approximately $p \in [0.3, 0.5]$. In other words, the partial derivative is monotonically increasing in p for $p \in [0, 0.3]$ (i.e., $G_{Th}(0, g) \leq G_{Th}(p, g) \leq G_{Th}(0.3, g)$), it reaches a maximum, and then it is monotonically decreasing in p , for $p \in [0.5, 1]$ (i.e., $G_{Th}(1, g) \leq G_{Th}(p, g) \leq G_{Th}(0.5, g)$).

Consider the first partial derivative

$$\frac{\partial G_{Th}(p, g)}{\partial g} = \frac{(1 - p^2)p^3}{(g - p^2)^2} \leq 0 \quad \forall g \neq p^2. \quad (49)$$

Based on (49), which is always true except when $g = p = 1$, $G_{Th}(p, g)$ is monotonically decreasing in g , which yields $G_{Th}(p, \infty) \leq G_{Th}(p, g) \leq G_{Th}(p, 1)$ for $p \in [0, 1]$.

C. Latency Gain

In the case of type I ARQ protocols the latency gain $G_T(p, 1) \geq 1$ when

$$T(\text{type II } C - \text{ARQ})(p, g) \geq T(\text{type II } H - \text{ARQ})(p, g) \quad (50)$$

$$\frac{T_F(1 + p)}{2(1 - p - \lambda T_F)} + \frac{T_F}{M} \geq \frac{T_F(1 + p^2(2 - p))}{2(1 - p^2(2 - p) - \lambda T_F)} + \frac{T_F}{M} + \frac{T_F(\Delta - 1)p(1 + p)}{M(1 + p - p^2)}. \quad (51)$$

This holds if

$$\frac{(p - \lambda T_F p)}{(1 + p - p^2)} \geq \frac{(\Delta - 1)p(1 + p)}{M(1 + p - p^2)}. \quad (52)$$

When $\lambda T_F \rightarrow 0$, the condition is satisfied when

$$p \geq 1 - \frac{M}{\Delta - 1}. \quad (53)$$

For $\Delta \in [2, M]$, the right hand side of the condition is bounded by

$$1 - \frac{M}{M - 1} \leq \left[1 - \frac{M}{\Delta - 1} \right] \leq -M + 1. \quad (54)$$

Thus, condition (53) always holds when $p \in [0, 1]$.