

The analysis of truncated ARQ and HARQ schemes using transition diagrams

Herald N. Kundaehi

*Department of Electronics and Telecommunications Engineering, College of Information and Communication Technologies,
University of Dar es Salaam, TANZANIA
Correspondence email: kundaehi@yahoo.com
ORCID iD: <https://orcid.org/0000-0003-2195-5861>*

Abstract

In digital communication systems, ARQ and HARQ schemes are employed in order to ensure efficient transmission of information and utilisation of the channel by retransmitting information that has not been received correctly at the receiver. In some applications, however, such as real-time services, the quality of information is degraded if too many retransmission trials are employed for each unit of erroneously received information. The retransmissions are therefore limited leading to truncated ARQ and HARQ schemes. In this study we develop a new method for deriving the expressions for the throughput efficiencies of truncated ARQ and HARQ schemes and investigate how the efficiencies vary with the number of allowed retransmissions.

Keywords: ARQ, HARQ, Throughput Efficiency, Truncated ARQ.

DOI: <http://dx.doi.org/10.4314/ijest.v12i2.5>

1. Introduction

The Error control is widely used in digital communication systems operating in unreliable transmission channels with the aim of ensuring acceptable delivery of transmitted information. At one end is Automatic Repeat reQuest (ARQ) in which there is no attempt to correct errors in the received information and the same information is therefore retransmitted if errors occur until the target information is received correctly. Since many packets have to be retransmitted to enable the reliable reception of a single packet, systems employing ARQ only exhibit good performance in low noise channels, but very poor performance in more noisy channels. At the other end is Forward Error Correction (FEC) in which the receivers attempt to correct the errors in the received transmissions and therefore there is no attempt to retransmit the target information. Such systems exhibit better performance in more noisy channels but very poor performance in low noisy channels because more error correcting but redundant information is incorporated in transmitted packets which leads to poor channel utilisation. The combination of ARQ and FEC leads to Hybrid ARQ (HARQ) which reduces the instances of retransmissions and improves the channel utilisation in systems operating in more noisy channels.

In both ARQ and HARQ based communication systems operating in more noisy channels a lot of time may be wasted in retransmitting failed packets. This is unacceptable in critical services like real-time applications, and the number of retransmitted packets has to be limited in order to avoid wasting time retransmitting outdated information at the expense of newer meaningful information. Therefore, when the number of packets being retransmitted reaches a certain value, the retransmissions are stopped and packets containing new information are transmitted instead, leading to truncated-ARQ (and truncated-HARQ) schemes. The analysis of truncated HARQ schemes employing convolutional coding in noisy channels has been undertaken by Malkamaki and Leib (2000), while Pack et al. (2008) attempted to minimize packet loss rate in truncated-ARQ schemes operating in wireless fading channels. Szczecinski et al. (2013) went on to optimise the rate in incremental redundancy HARQ in order to maximize throughput. Similar work has been done on cross-layer design that combines cooperative diversity with truncated ARQ (Dai and

Letaief, 2005; Dai and Letaief, 2008; Choi et al., 2010), and more studies have also been undertaken to investigate cross-layer combining of adaptive modulation and coding with truncated ARQ (Liu et al, 2004; Yang et. al., 2012). Attempts have also been taken to apply the cross-layer combining of truncated-ARQ and adaptive modulation to multichannel beam-forming MIMO systems (Ivanis et al, 2015), and in cooperative relay networks employing network coding (Yangi et al, 2016). The main drive in these studies was to show that truncated ARQ/HARQ can lead to improved performance in various types of systems, as well as to establish methods of analysing the truncated ARQ/HARQ systems.

Different methods have been used to model communication systems employing truncated ARQ and HARQ. For example, in the work of Malkamaki and Leib (2000) state diagrams were employed to investigate transmissions in block fading channels, while Pack et al. (2008) employed queuing models and fractional probabilistic transmissions to investigate transmissions in wireless fading channels. In the analysis of Dai and Letaief (2008), simpler models based on M-QAM, with possible differing symbol and packet error rates in the direct transmissions and retransmissions were used. The performance of conventional and truncated ARQ/HARQ schemes has also been investigated analytically by Larsson et al (2014, 2016). The approach in these reports employs the Laplace transform of the pdf representing the packet success or failure rate as a function of transmitted and retransmitted packets with the pdf varying with the type of ARQ/HARQ, channel model, and modulation scheme employed. In this report we analyse truncated-ARQ/HARQ schemes based on methods developed in Kundaeli (2010). In that study, transition diagrams were used to characterise the transmissions, with the focus being on tracing the fate of a given packet of interest as it traverses the communication network. The benefit of the approach is its simplicity in visualizing the communication process, and in deriving the performance metrics. This report therefore presents the application of the approach to give a new method for analysing truncated ARQ and HARQ schemes. The analysis of the systems is given in section 2, followed by numerical results in section 3. The discussion of the results is also given in section 3, with the conclusion given in section 4.

2. System Model and Analysis

In this analysis, we assume that the round-trip delay between the transmitter and receiver is divided into N slots, with each transmitted packet occupying 1 slot. We also assume that time is divided into frames, with a frame occupying N slots. We then note that in the SW (Stop and Wait) ARQ scheme, only one packet is transmitted in a frame, usually in the first slot, whereas in SR (Selective Repeat) based ARQ schemes, a packet is transmitted in each slot. A snapshot of the timing diagram for the SR (or SW) scheme is shown in Figure 1, with the packet of interest represented by * while the instants of decisions on the received packets are indicated by up arrows. Therefore, in Figure 1 the packet of interest is first transmitted in slot 1 of frame 0, where its failure at the ends of the frames leads to its retransmission in slot 1 of frames 1 and 2. Note that this representation fits the sliding window notation in ARQ/HARQ schemes, so that when a packet has been successfully received or discarded, the frames are shifted forward by one slot.

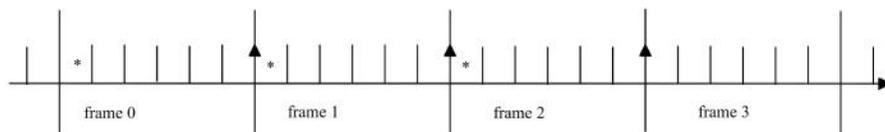


Figure 1: Timing diagram of the SR ARQ scheme with frames of N = 6.

If the duration of a slot is represented by z in the frequency domain, we obtain the transition diagram for the truncated SR scheme as shown in Figure 2.

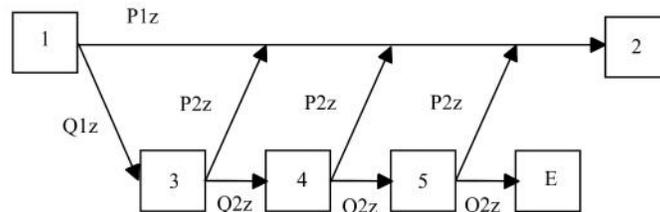


Figure 2: Transition diagram of the truncated SR ARQ scheme with M (= v) = 3.

We denote the transition probabilities by P_1 for data + ED (Error Detection) packets whereas P_2 can be for either data (+ED) or FEC (Forward Error Correction) packets. Therefore, starting from state 1, the system will reach state 2 if the packet of interest is received error free or with correctable errors provided that the number of retransmissions does not exceed M. On the other hand, if there are M consecutive retransmission failures, the system ends in state E and the packet is taken as lost and dropped from the transmission

process. We therefore obtain the transfer functions from state 1 to 2 and E respectively for general M from Figure 2 as

$$\begin{aligned} \Phi_{12}(z) &= P_1 z + Q_1 z \sum_{m=0}^{M-1} (Q_2 z)^m P_2 z, \\ \Phi_{1E}(z) &= Q_1 z (Q_2 z)^M, \end{aligned} \tag{1}$$

and the delays as

$$\begin{aligned} L_{12} &= \left. \frac{d}{dz} (\Phi_{12}(z)) \right|_{z=1} = \frac{Q_1 + P_2 - (1 + P_2 + MP_2)Q_1 Q_2^M}{P_2}, \\ L_{1E} &= \left. \frac{d}{dz} (\Phi_{1E}(z)) \right|_{z=1} = (M + 1)Q_1 Q_2^M. \end{aligned} \tag{2}$$

We note that for the normal SR scheme where an infinite number of retransmissions is used, we obtain the delay as

$$L_\infty = \lim_{M \rightarrow \infty} L_{12} = \frac{Q_1 + P_2}{P_2}. \tag{3}$$

Since the delay is in terms of slots (packets), and this is the time in slots needed to process one packet until it is successfully received, we obtain the packet-based throughput efficiency as

$$y = \frac{1}{L_\infty} = \frac{P_2}{Q_1 + P_2}. \tag{4}$$

For the truncated SR scheme, the time needed to process one packet is

$$L_{12E} = L_{12} + L_{1E} = \frac{Q_1 + P_2 - Q_1 Q_2^M}{P_2}. \tag{5}$$

However, since in this case the packet being processed may either be correctly received or lost, the effective fraction of the packet received in state 2 is given by

$$S = 1 - Q_1 Q_2^M, \tag{6}$$

and the throughput efficiency for the truncated SR scheme is then obtained as

$$y = \frac{S}{L_{12E}}, \tag{7}$$

which evaluates to

$$y = \frac{(1 - Q_1 Q_2^M) P_2}{Q_1 + P_2 - Q_1 Q_2^M}. \tag{8}$$

We note that with infinite retransmissions this expression reduces to (4) as expected. We also note that if $P_2 = P_1$ we obtain the throughput efficiency for the conventional SR-ARQ scheme, and if $P_1 = P_2$ we obtain the throughput efficiency for the Type I SR-HARQ scheme.

The transition diagram for the truncated SR-GBN scheme is given in Figure 3 in which, starting from state 1, the system ends in state 2 provided that the number of GBN retransmissions does not exceed M . We note that the transition probabilities are P_1 for data while P_2 and P_3 can be either data or FEC packets. Therefore, with $v = 3$ and $M = 3$, the system performs SR retransmissions from state 3 to 5 and GBN retransmissions from state 6 to 8.

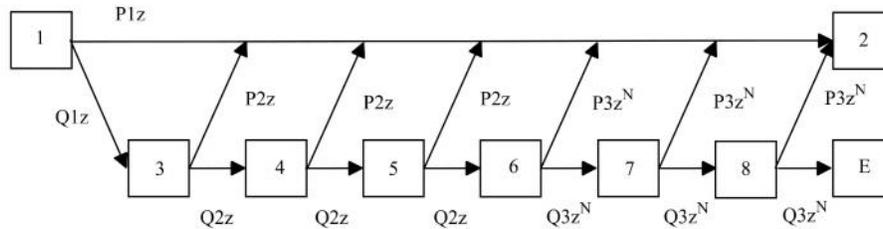


Figure 3: Transition diagram of the truncated SR-GBN scheme with $v = 3$ and $M = 3$.

The transfer functions for general v and M are then obtained as

$$\begin{aligned} \Phi_{12}(z) &= P_1 z + Q_1 z \sum_{n=0}^{v-1} (Q_2 z)^n P_2 z + Q_1 z (Q_2 z)^v \sum_{m=0}^{M-1} (Q_3 z^N)^m P_3 z^N, \\ \Phi_{1E}(z) &= Q_1 z (Q_2 z)^v (Q_3 z^N)^M, \end{aligned} \tag{9}$$

the delays as

$$\begin{aligned} L_{12} &= \frac{1}{P_2 P_3} [P_3 (Q_1 + P_2) + (NP_2 - P_3) Q_1 Q_2^v - \{NP_2 + (NM + v + 1) P_2 P_3\} Q_1 Q_2^v Q_3^M], \\ L_{1E} &= (MN + v + 1) Q_1 Q_2^v Q_3^M, \\ L_{12E} &= \frac{1}{P_2 P_3} [P_3 (Q_1 + P_2) + (NP_2 - P_3) Q_1 Q_2^v - NP_2 Q_1 Q_2^v Q_3^M], \end{aligned} \tag{10}$$

and the throughput efficiency as given by (7) with S given by

$$S = 1 - Q_1 Q_2^v Q_3^M. \tag{11}$$

Note that

$$L_v = \lim_{M \rightarrow \infty} L_{12} = \frac{1}{P_2 P_3} [P_3(Q_1 + P_2) + Q_1 Q_2^v (N P_2 - P_3)]$$

$$L_\infty = \lim_{v \rightarrow \infty} L_{12} = \frac{Q_1 + P_2}{P_2}$$
(12)

and if $P_2 = P_3 = P_1$, then the conventional SR-GBN ARQ scheme is obtained, whereas if $P_1 = P_3 = P_2$, the Type I SR-GBN HARQ scheme is obtained.

The transition diagram for the truncated SR-ST1 scheme is given in Figure 4 in which we note that when the system suffers v transmission failures, ST retransmissions start in which the failed packet is transmitted in all N slots of the frame.

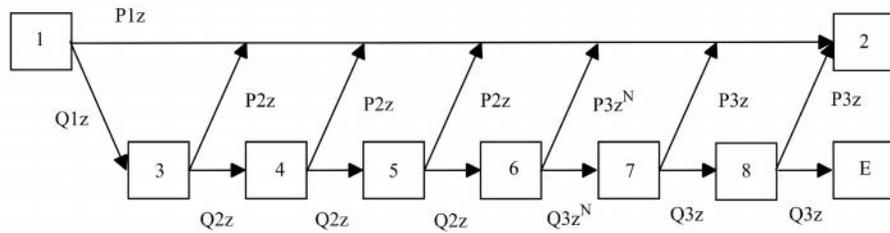


Figure 4: Transition diagram of the truncated SR-ST1 scheme with $v = M = 3$.

The transfer functions are then obtained as

$$\Phi_{12}(z) = \begin{cases} P_1z + Q_1z \sum_{n=0}^{v-1} (Q_2z)^n P_2z, & M = 0 \\ P_1z + Q_1z \sum_{n=0}^{v-1} (Q_2z)^n P_2z + Q_1z (Q_2z)^v \sum_{m=0}^{M-1} (Q_3z)^m P_3z^N, & M > 0 \end{cases}$$

$$\Phi_{1E}(z) = \begin{cases} Q_1z (Q_2z)^v, & M = 0 \\ Q_1z (Q_2z)^v z^{N-1} (Q_3z)^M, & M > 0 \end{cases}$$
(13)

the delays as

$$L_{12} = \begin{cases} \frac{1}{P_2} [(Q_1 + P_2) - (1 + P_2 + vP_2)Q_1Q_2^v], & M = 0 \\ \frac{1}{P_2 P_3} \left[P_3(Q_1 + P_2) + \{P_2 - P_3 + (N - 1)P_2P_3\}Q_1Q_2^v \right], & M > 0 \end{cases}$$

$$L_{1E} = \begin{cases} (v + 1)Q_1Q_2^v, & M = 0 \\ (M + N + v)Q_1Q_2^vQ_3^M, & M > 0 \end{cases}$$
(14)

$$L_{12E} = \begin{cases} \frac{1}{P_2} [(Q_1 + P_2) - Q_1 Q_2^v] & M = 0 \\ \frac{1}{P_2 P_3} [P_3 (Q_1 + P_2) + \{P_2 - P_3 + (N - 1) P_2 P_3\} Q_1 Q_2^v - P_2 Q_1 Q_2^v Q_3^M] & M > 0 \end{cases} \quad (15)$$

and the throughput efficiency as given as in (7) with S given by

$$S = 1 - Q_1 Q_2^v Q_3^M. \quad (16)$$

Note that

$$L_v = \lim_{M \rightarrow \infty} L_{12} = \frac{1}{P_2 P_3} [P_3 (Q_1 + P_2) + \{P_2 - P_3 + (N - 1) P_2 P_3\} Q_1 Q_2^v] \quad (17)$$

$$L_\infty = \lim_{v \rightarrow \infty} L_{12} = \frac{Q_1 + P_2}{P_2}.$$

The transition diagram for the truncated SR-ADP (SR Alternate Data/Parity) HARQ scheme is given in Figure 5 with $v = 3$ and $M = 2$. While we use P_2 for block FEC encoding, rate 2,1 systematic convolutional encoding yields data (P_1) and parity (P_3) code words. When the system first suffers a transmission error it attempts v SR retransmissions to state 6. The parity word is then sent and if it is error free, it is decoded by inversion (P_3), otherwise it is combined with the data word received in state 5 and decoded in state 7. If the decoding is negative (Q_4), the data word is again transmitted in state 8. If the data word is received in error (Q_5), the system transits to state 9 where the erroneous data word is combined with the parity word received in state 6 and decoded. If the decoding is successful (P_6), the system transits to state 2 otherwise it transits to state E with the packet dropped and a new data word taken for transmission. Note that if the decoding in state 9 is unsuccessful, and M exceeds 2, the process performed in states 6 to 9 can be repeated.

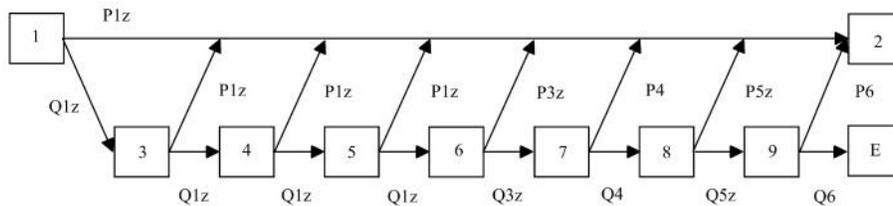


Figure 5: Transition diagram of the truncated SR-ADP HARQ scheme with $v = 3$ and $M = 2$.

The transfer functions are then given by

$$\Phi_{12}(z) = \sum_{n=0}^v (Q_1 z)^n P_1 z + (Q_1 z)^{v+1} \begin{cases} 0, & M = 0 \\ \sum_{m=0}^{M_3} (\Pi_Q z^2)^m (P_3 z + Q_3 P_4 z + Q_3 Q_4 P_5 z^2 + Q_3 Q_4 Q_5 P_6 z^2) \\ + (\Pi_Q z^2)^{M_1} (P_3 z + Q_3 P_4 z), & M = odd \\ \sum_{m=0}^{M_2} (\Pi_Q z^2)^m (P_3 z + Q_3 P_4 z + Q_3 Q_4 P_5 z^2 + Q_3 Q_4 Q_5 P_6 z^2), & M = even \end{cases} \quad (18)$$

$$\Phi_{1E}(z) = (Q_1 z)^{v+1} \begin{cases} 1, & M = 0 \\ (\Pi_Q z^2)^{M_1} Q_3 Q_4 z, & M = odd \\ (\Pi_Q z^2)^{M_2} \Pi_Q z^2, & M = even \end{cases} \quad (19)$$

where $\Pi_Q = Q_3 Q_4 Q_5 Q_6$, $M_i = \text{floor}((M-i)/2)$ and $\text{floor}(x)$ is the largest integer that does not exceed x .

The delays are then obtained as

$$L_{12} = \frac{1 - (1 + P_1 + vP_1)Q_1^{v+1}}{P_1} + Q_1^{v+1} \begin{cases} 0, & M = 0 \\ \left\{ v + 1 + \frac{1 + Q_3 Q_4}{1 - \Pi_Q} \right\} - \left\{ (2M_3 + v + 3) + \frac{1 + Q_3 Q_4}{1 - \Pi_Q} \right\} \Pi_Q^{M_3+1} \\ + (2M_1 + v + 2)(1 - Q_3 Q_4) \Pi_Q^{M_1}, & M = odd \\ \left\{ v + 1 + \frac{1 + Q_3 Q_4}{1 - \Pi_Q} \right\} - \left\{ (2M_2 + v + 3) + \frac{1 + Q_3 Q_4}{1 - \Pi_Q} \right\} \Pi_Q^{M_2+1}, & M = even \end{cases} \quad (20)$$

$$L_{1E} = Q^{v+1} \begin{cases} (v + 1), & M = 0 \\ (2M_1 + v + 2)(\Pi_Q)^{M_1} Q_3 Q_4, & M = odd \\ (2M_2 + v + 3)(\Pi_Q)^{M_2} \Pi_Q, & M = even. \end{cases} \quad (21)$$

Noting that $M_3 + 1 = M_1$ and $M_2 + 1 = M_0$ we obtain

$$L_{12E} = \frac{1 - (1 + P_1 + vP_1)Q_1^{v+1}}{P_1} + Q_1^{v+1} \times \begin{cases} (v+1), & M = 0 \\ \left\{ v+1 + \frac{1 + Q_3Q_4}{1 - \Pi_\varrho} \right\} + \left\{ 1 - \frac{1 + Q_3Q_4}{1 - \Pi_\varrho} \right\} \Pi_\varrho^{M_1}, & M = odd \\ \left\{ v+1 + \frac{1 + Q_3Q_4}{1 - \Pi_\varrho} \right\} - \left\{ \frac{1 + Q_3Q_4}{1 - \Pi_\varrho} \right\} \Pi_\varrho^{M_0}, & M = even \end{cases} \quad (22)$$

with

$$S = 1 - Q_1^{v+1} \begin{cases} 1, & M = odd \\ (\Pi_\varrho)^{M_1} Q_3Q_4, & M = odd \\ (\Pi_\varrho)^{M_0}, & M = even. \end{cases} \quad (23)$$

The Type II HARQ Chase Combining scheme in which all packets transmitted and retransmitted for a certain data word are similar is given in Figures 6. We note that with the system as given in Figure 6, the same FEC code word is transmitted in the odd states (3, 5, 7, ..). If there are non-correctable errors, the received word is combined with all previously received code words using appropriate weighting before decoding in the next state. As usual, if the system reaches state E the packet is discarded and a new one picked for transmission.

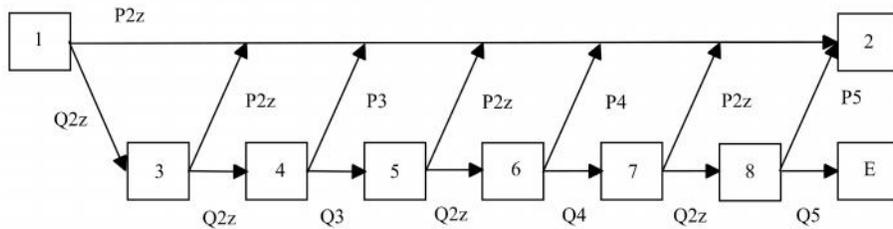


Figure 6: Transition diagram of the Chase Combining HARQ scheme with M = 3 retransmissions.

We obtain the transfer functions from Figure 6 as

$$\Phi_{12}(z) = P_2z + \sum_{m=0}^{M-1} (Q_2)^m \prod_{i=2}^{m+2} Q_i P_2 z^{m+2} + \sum_{m=0}^{M-1} (Q_2)^{m+1} \prod_{i=2}^{m+2} Q_i P_{m+3} z^{m+2} \quad (24)$$

$$\Phi_{1E}(z) = (Q_2z)^{M+1} \prod_{m=1}^M Q_{2+m},$$

the delays as

$$L_{12} = P_2 + \sum_{m=0}^{M-1} (m+2)(Q_2)^m \prod_{i=2}^{m+2} Q_i (P_2 + Q_2 P_{m+3}) \tag{25}$$

$$L_{1E} = (M+1)(Q_2)^{M+1} \prod_{m=1}^M Q_{2+m},$$

with

$$S = 1 - (Q_2)^{M+1} \prod_{m=1}^M Q_{2+m}. \tag{26}$$

Although the Type II Incremental Redundancy (IR) schemes can be represented by similar transition diagrams, the fact that the transmitted bits are different in each transmission does not produce results that compare well with the schemes presented above. Analysis of the IR scheme is therefore not included in this report.

3. Results and Discussions

The results of this study are given in Figures 7 to 17. In all these results, the parameters have been based on BCH codes fixed at packet length = FL = 255 bits with data bits = k = 215 bits and error correcting capability of t = 5 bits, and where packet combining is used, FL = 511 bits with error correcting capability of t = 9 bits, and maximum number of retransmissions = M = 4. In Figure 7 we compare the analytical and simulation results for the SR-ARQ scheme. As can be seen, both the analytical and simulation computations give the same throughput efficiencies, and these do not differ much from those of the ideal SR-ARQ scheme. The results show that the simulation results agree with the analytical ones and this was found to be the case for the rest of the results. Therefore, the simulation results have been omitted in the rest of this report.

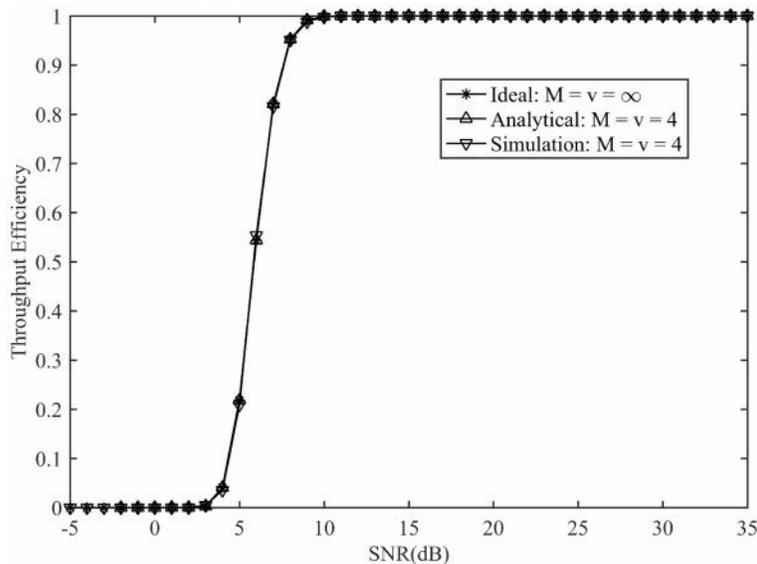


Figure 7: Throughput efficiency of the Ideal, Analytical and Simulated SR-ARQ schemes at FL = 255 bits and M = v = 4.

Figure 8 shows the results when Type I HARQ packets are used for the SR scheme. Note that the first transmission employs only error detection, while the retransmissions employ error correction. As expected, the throughput efficiency increases with the number of retransmissions (M = v) towards the ideal SR scheme.

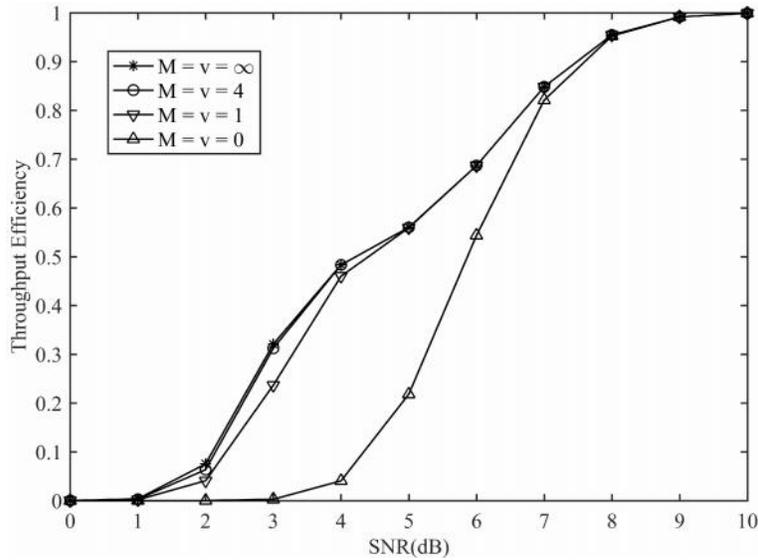


Figure 8: Throughput efficiency of the SR-HARQ scheme for various values of retransmissions M at $FL = 255$ bits and $t = 5$.

Figure 9 presents the results for the SR-GBN scheme in comparison with the ideal SR case. Two observations can be made in these results. First, at low values of SNR the throughput efficiency decreases with increase in M , the number of GBN retransmissions. This is expected since for each packet retransmitted in the GBN mode, N packet slots are used compared to only one packet slot in the SR scheme. Secondly, at higher values of SNR, the throughput efficiency becomes the same independent of the value of M . The implication here is that as the SNR increases, it becomes harder for the system to reach GBN retransmissions, resulting in improved performance.

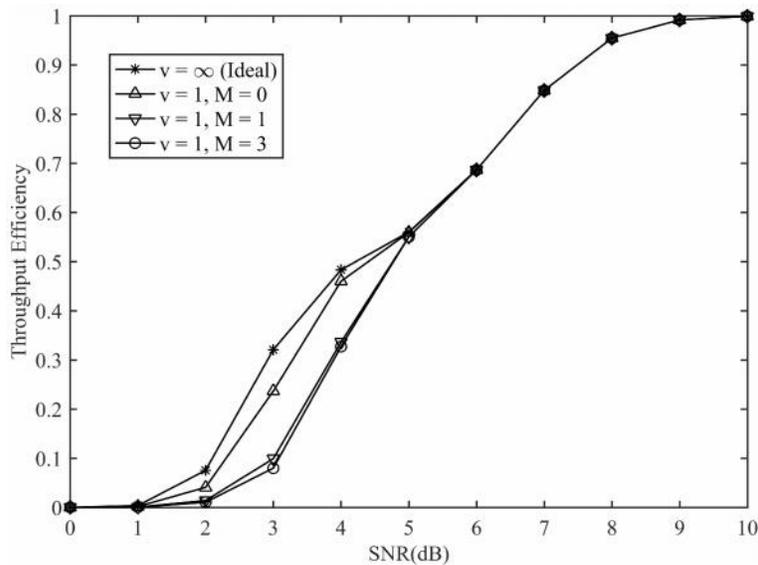


Figure 9: Throughput efficiency of the SR-GBN-HARQ scheme for different values of M at $v = 1$, $FL = 255$ bits and $t = 5$.

Figure 10 shows further how the throughput efficiency varies with the GBN retransmissions for the truncated SR-GBN scheme in comparison with the conventional SR-GBN scheme. As can be seen, the efficiency decreases with the retransmissions M , attaining worst case status in the conventional SR-GBN scheme where M is infinite. The efficiency of the truncated scheme is therefore expected to increase with decreasing M , or decrease as it increases. This variation can also be derived from (7), (10) and (11) as

$$\frac{dy}{dM} \approx -\frac{P_1 Q_1^{M+v+1} [(N-1)(1-Q_1^{v+1})]}{L_{12E}^2} \tag{27}$$

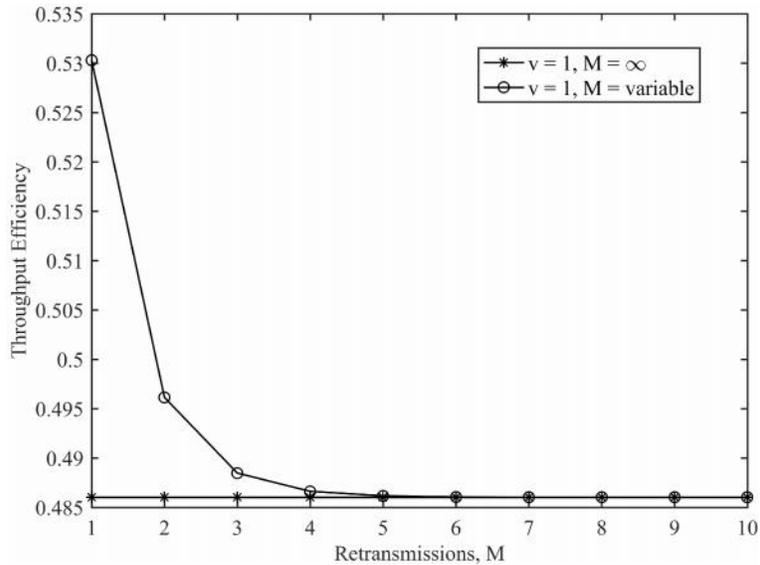
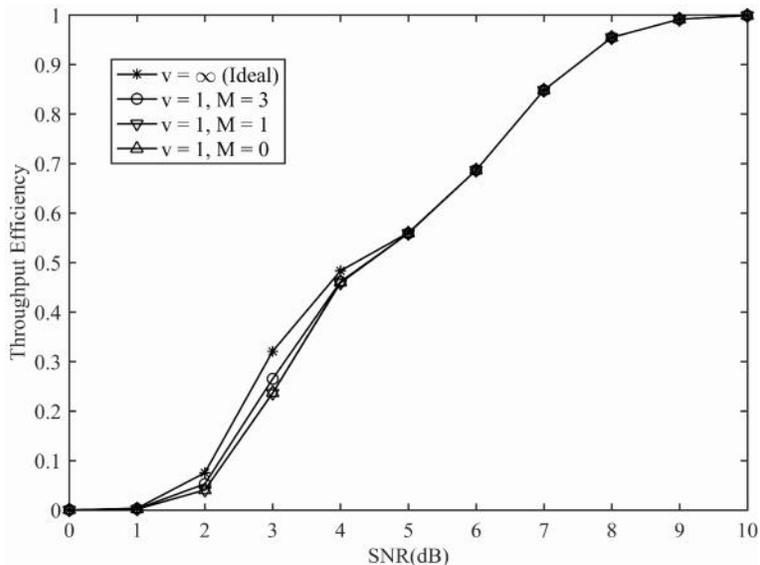


Figure 10: Throughput efficiency of the conventional and truncated SR-GBN schemes at $v = 1$, $SNR = 6.7$ dB and $FL = 255$ bits.

Figure 11 shows that the performance of the SR-ST1 scheme improves with the number of retransmissions unlike in the case of the SR-GBN scheme. Note that in the SR-ST1 scheme the initial retransmissions involves the wastage of N packet slots because N copies of the failed packet are transmitted. In subsequent retransmissions, however, only one packet slot is used. Therefore after the initial retransmissions, only one extra slot is needed each time to prevent the packet from being discarded, instead of N packet slots each time as in the SR-GBN scheme, and this leads to improved performance.



Figures 11: Throughput efficiency of the SR-ST1-HARQ scheme for different values of M at $v = 1$, $FL = 255$ bits and $t = 5$.

As in the case of the SR-GBN scheme, the conventional and truncated SR-ST1 schemes can be used to illustrate in detail how the throughput efficiency varies with the retransmissions. As seen in Figure 12, the throughput efficiency of the truncated SR-ST1

scheme increases with ST retransmissions, attaining best performance status in the standard SR-ST1 scheme where the number of retransmissions is infinite. This variation of the efficiency can be derived from (7), (15) and (16) as

$$\frac{dy}{dM} \approx \frac{P_1^2 Q_1^{M+v+1} [Q_1 + (N-1)P_1 Q_1^{v+1}]}{L_{12E}^2} \tag{28}$$

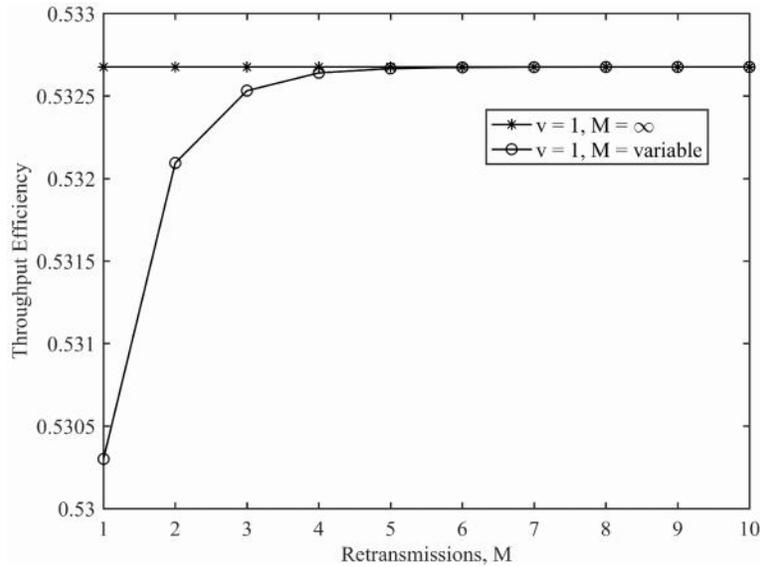


Figure 12: Throughput efficiency of the conventional and truncated SR-ST1 schemes at $v = 1$, $SNR = 6.7$ dB and $FL = 255$ bits.

Figure 13 shows the performance of the SR-ADP scheme. Two observations emanate from these results. First, since in the SR-ADP scheme some of the retransmitted packets employ error correction, the performance is better than that of the SR-ARQ scheme depicted by Figure 7 in which all packets employ only error detection.

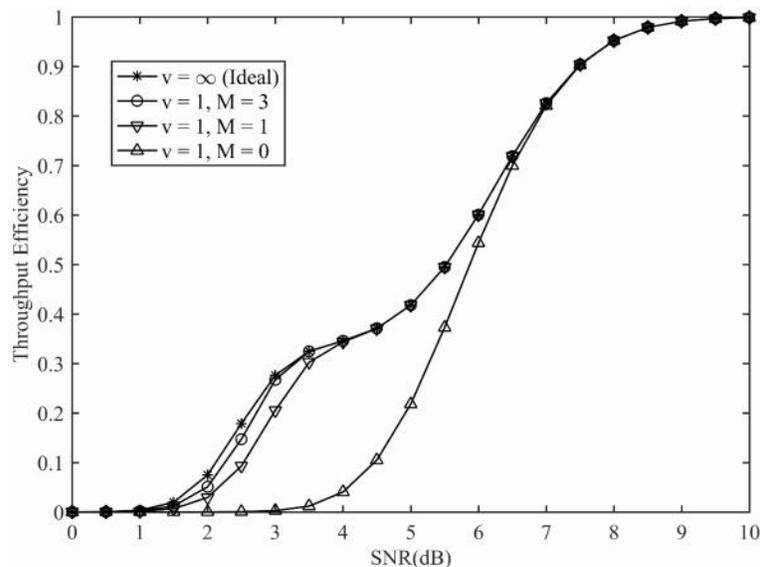


Figure 13: Throughput efficiency of the SR-ADP-HARQ scheme for different values of M at $v = 1$, $FL = 255$ bits and $t = 5$.

Secondly, the performance could have been worse than that of the SR-HARQ scheme depicted in Figure 8 in which all retransmitted packets employ error correction. However, in the SR-ADP scheme a failed retransmitted packet is combined with the

previous failed packet, and retransmission is employed only when the combined packet also fails. For this reason the performance is better than it would be otherwise. Note, however, that as the SNR increases, the chances of retransmissions decreases and the performance approaches that of the ideal SR schemes.

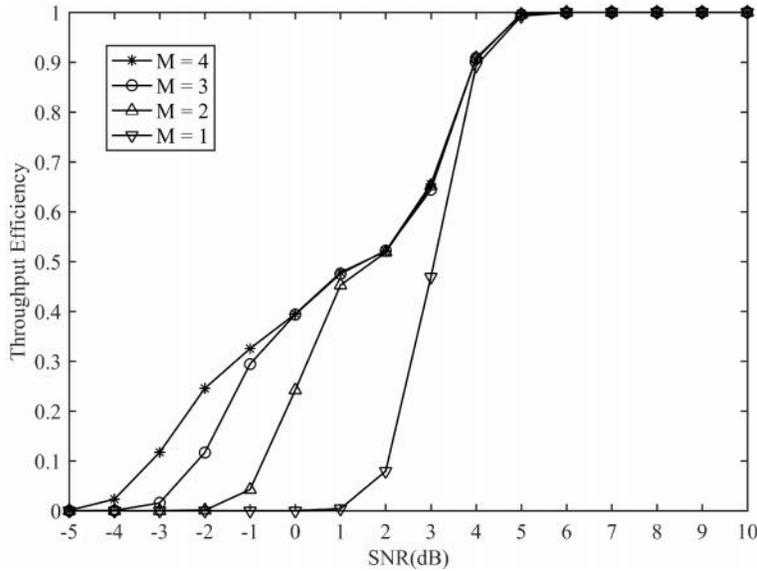


Figure 14: Throughput efficiency of the CC-HARQ scheme for various values of M at FL = 255 bits and t = 5.

The performance of the CC scheme is given in Figure 14 and in this study it has been implemented without fading in order to maintain compatibility with the other schemes. As seen, the performance increases with the number of retransmissions, and the increase is higher with each retransmission than in the other schemes. This is expected because each received packet is soft combined with all previously received packets which increases the effective SNR.

Comparison of the analysed schemes is given in Figure 15. As seen, the SR-ARQ scheme has the lowest performance because it only employs error detection while the other schemes employ error correction in the retransmitted packets. Note that all SR-based schemes lie below the SR-HARQ scheme in the same order as in the conventional SR-based schemes. The CC scheme, however, turns out to have the highest performance, and this is attributed to the soft-combining employed.

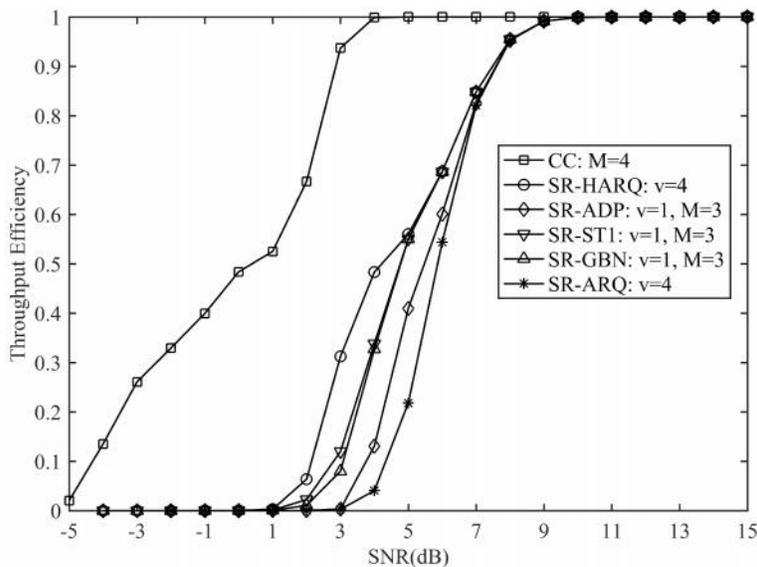


Figure 15: Comparison of the throughput efficiencies of the SR-ARQ, SR-HARQ, SR-GBN, SR-ST1, SR-ADP and CC schemes with $v + M = 4$ at FL = 255 bits and t = 5.

The buffer storage space required at the receivers in conventional ARQ/HARQ schemes depends on the types of scheme employed and the channel conditions (Kundaali, 2013). In order to determine the required storage space at the receiver, we assume that the packets are numbered consecutively, and they have to be delivered to the destination in that order. Therefore, referring to Figure 1 for the SR scheme and assuming the worst case scenario in which only packet 1 in frame 0 fails continuously, then in each frame $N-1$ packets will be received correctly, and these have to be stored. In the ideal SR scheme the stored packets will only be delivered to the destination after packet 1 is received correctly, while in the SR (v) scheme they will be delivered if any retransmission succeeds, but must be delivered if the allowed v consecutive retransmissions fail. Figure 16 shows the required buffer space for 3 levels of allowed retransmissions. The important observation here is that at high SNR (low probability of bit error) packet failure decreases to zero and no storage space is required. As SNR decreases however, packet failure rate increases and more storage space is required, but that levels off to a constant value because of the limited retransmissions. The results in Figure 17 show the comparison between the ideal SR ARQ scheme needing infinite storage space and truncated schemes where the allowed retransmissions limit the amount of required storage space.

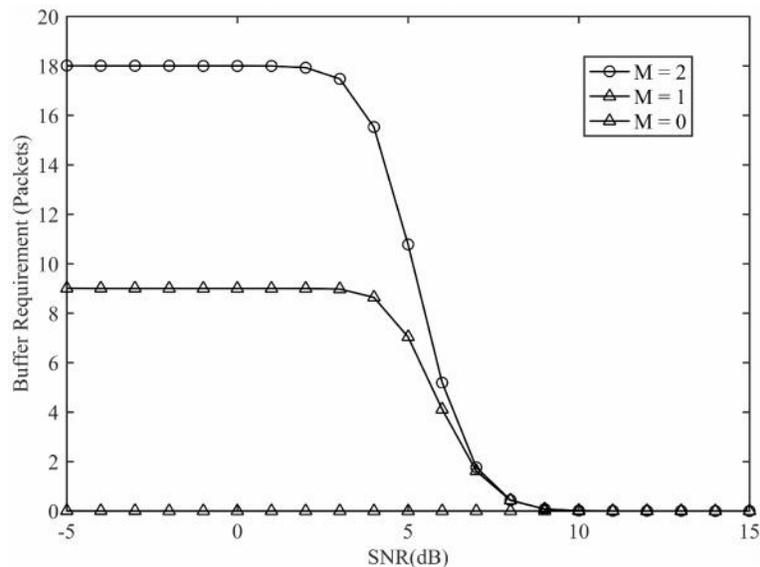


Figure 16: Receiver buffer requirements for the SR-ARQ scheme for 3 values of allowed retransmissions at $FL = 255$ bits.

The GBN scheme does not require storage space because once a packet fails, all packets that had been transmitted after the failure are retransmitted. Therefore the behaviour of the SR-GBN scheme is governed more by the SR phase such that the storage requirements for the SR-GBN (v, M) scheme are the same as those of the SR (v) scheme. The behaviour of the scheme can therefore be represented by Figures 16 and 17. In the SR-ST1 scheme, we first note that its behaviour is governed more by the SR phase like in the SR-GBN scheme. We note, however, that when a packet fails in the ST scheme, that packet is retransmitted continuously while packets transmitted after its failure are received and stored. The required storage space for the SR-ST1 (v, M) scheme is therefore $(v + 1) * (N - 1)$ packets and its behaviour resembles the one given in Figures 16 and 17 save for the minimum $N - 1$ packets. In the SR-ADP (v, M) scheme, the v SR retransmissions call for storage space of $v * (N - 1)$ packets. The M ADP retransmissions also call for storage space of $M * (N - 1)$ packets. Since in the ADP phase 1 packet is used for combining with an incoming packet, however, the required storage space for the SR-ADP (v, M) scheme is $(v + M) * (N - 1) + 1$ packets. In the CC (M) scheme, any failed packet is combined with new incoming packets and must therefore be stored. Therefore, a CC scheme operating according to Figure 1 and utilising M retransmissions needs storage space for $M * N$ packets. The maximum buffer storage requirements for the analysed ARQ/HARQ schemes are given in Table 1. We note that the required storage space for the SR-GBN (v, M) and SR-ST1 (v, M) schemes do not depend on M .

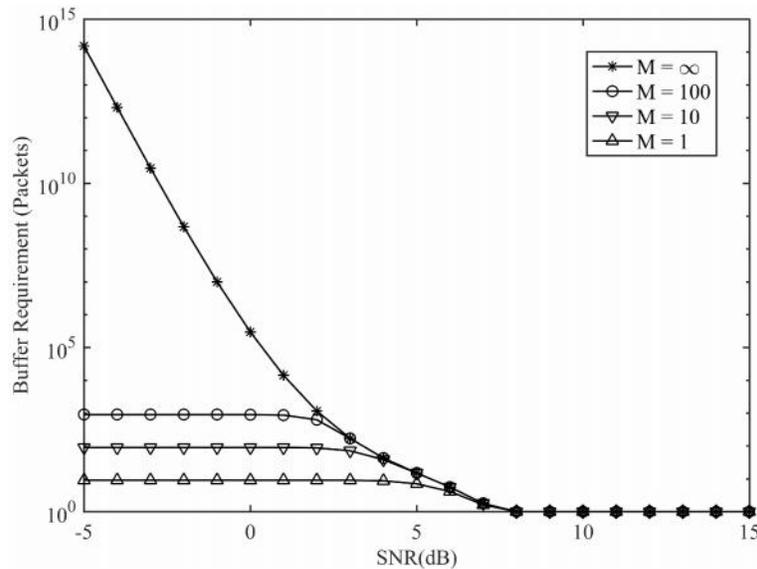


Figure 17: Comparison of receiver buffer requirements for ideal and truncated SR-ARQ schemes at FL = 255 bits.

Table 1: Storage Requirements for some ARQ/HARQ Schemes

Scheme	v	M	space
Ideal SR ARQ	infinity	-	infinity
Truncated SR-ARQ	v	-	$v*(N-1)$
SR-GBN	v	M, infinity	$v*(N-1)$
SR-ST1	v	M, infinity	$(v+1)*(N-1)$
SR-ADP	v	M	$(v + M)*(N-1) + 1$
CC	-	M	$M*N$

The motivation for limiting the retransmissions in ARQ schemes arose from the desire to have systems whose performance approach that of the ideal SR scheme but whose requirements for buffer storage at the receiver are far less than the infinite values needed for the ideal SR scheme. This led to the development of the SR-GBN, SR-ST and other mixed mode schemes. The challenge then turned to how to simplify the complex circuits associated with handling decision making and buffer storage in mixed mode ARQ schemes at the receiver. The need to truncate the retransmissions is imposed by systems that demand real-time delivery of transmitted information, especially for critical applications. This demand puts restrictions on the choice of both ARQ and HARQ schemes. For example, in the SR-GBN and SR-ST1 schemes, the value of M does not affect the required storage space, and it is only used to meet the real-time delivery requirements of the transmitted information. Likewise, once the maximum allowed number of retransmissions has been chosen to suit other design requirements of the communication systems, it becomes necessary to search for ARQ/HARQ schemes which yield maximum throughput efficiency. Note also that HARQ schemes employing Chase Combining and Incremental Redundancy are truncated systems by default (Chelli and Alouini, 2013). However, these differ from the other systems analysed in this report because the retransmitted packets may not be the same, and also they use soft combining at the receiver. In addition, the packet combining makes the systems more complex.

Apart from throughput, throughput efficiency and storage space requirements the other important performance parameters associated with ARQ and HARQ schemes are packet delay and packet drop rate (Malhotra et al., 2011). In conventional systems, the packet delay is determined by the type of ARQ/HARQ scheme employed, and the channel conditions. In truncated schemes however, the maximum delay is fixed by the allowed number of retransmissions, and the delay performance can therefore be chosen at will. For example in the SR-GBN and SR-ST1 schemes, the value of M can be used to fine tune the delay requirements of the packets being delivered to the destination. The packet delay also increases as the SNR decreases, and the delay can be chosen to match the SNR. We note, however, that the delay used for computing the throughput in the results above is not the actual delay of the packets. In the SR-ARQ scheme for example, each packet delay in the results above translates into a frame duration in real-time applications. The packet drop rate also depends on the maximum allowed retransmissions. When infinite retransmissions are used in real-time applications, more packets are dropped. Intuitively, however, the packet drop rate is higher at low SNR and lower at high SNR. The allowed number of retransmissions can therefore be chosen to match the acceptable packet drop rate depending on the SNR.

We also note that packet retransmission can be used to mitigate variations in the channel conditions that vary with time. For example, truncated HARQ is added on top of AMC (adaptive modulation and coding) where the type of modulation and modes of

HARQ are changed to adapt to changes in the channel conditions. Studies have shown that the combination of AMC and truncated ARQ makes the systems more robust and also applicable to time limited applications (Liu et al., 2004; Kang et al, 2010; Yang et al, 2012; Ivanis et al, 2015). While AMC and truncated HARQ can separately be employed to deal with varying channel conditions, some studies have been carried out to compare the effects of AMC and truncated HARQ, showing that AMC yields improved performance in noisy channels (Sassioui *et al.*, 2017). We also note that if restrictions are applied to the results of some truncated ARQ/HARQ based systems employed in different environments or analysed using different approaches, results similar to the ones presented in this report are obtained. For example, the analysis reported by Dai and Letaief (2008) applies to M-QAM transmissions, with possible differing symbol and packet error rates in the direct transmissions and retransmissions. If the expressions in that analysis are subjected to BPSK modelling and non-variable error rates in the direct and retransmission modes, however, we obtain results similar to the ones presented in this report.

4. Conclusion

In this report, systems employing truncated ARQ/HARQ schemes have been analysed using transition diagrams, leading to a new method of analysing such systems. The analysis has shown how the throughput efficiency varies with the different parameters of the communication systems. If specific performance requirements are prescribed for a given communication system operating in a given channel, the number of allowed retransmissions can be chosen to suit the requirements. Apart from the throughput efficiency (and throughput), the packet delay and buffer space requirements are the other main performance parameters in such systems. The number of allowed retransmissions in the systems can also be chosen to fix the required packet delay and buffer storage space. Although analytical results for truncated ARQ/HARQ systems have been obtained in other studies, they can be shown to yield similar results to the ones obtained in this report. Generally, the results in this report contribute to the search for systems providing limited delivery time solutions for improved performance. Future applications of the methods developed in this study include variable length packet ARQ schemes, and error control in 4G/5G systems where parallel ARQ sessions are employed.

Acknowledgements

The authors would like to acknowledge the reviewers for their valuable comments.

References

- Chelli, A. and Alouini, M.-S. 2013. On the performance of hybrid-ARQ with incremental redundancy and with code combining over relay channels. *IEEE Transactions on Wireless Communications*, Vol. 12, No. 8, pp. 3860–3871.
- Choi, Y., Oh, S. and Choi, S. 2010. S-ARQ: A new truncated ARQ for IP-based wireless network. *Journal of Communication Networks*, Vol. 12, No. 2, pp. 174-180.
- Dai, L. and Letaief, K.B. 2005. Cross-layer design for combining cooperative diversity with truncated ARQ in ad-hoc wireless networks. *IEEE Global Telecommunications Conference*, 28 Nov.-2 Dec. 2005, St. Louis, MO, USA.
- Dai, L. and Letaief, K. B. 2008. Throughput maximization of ad-hoc wireless networks using adaptive cooperative diversity and truncated ARQ. *IEEE Transactions Communications*, Vol. 56, No. 11, pp. 1907-1918.
- Ivanis, P., Drajić, D. and Brkić, S. 2015. Cross-layer combining of adaptive modulation and truncated ARQ in multichannel beamforming MIMO systems. *Radioengineering*, Vol. 24, No. 4, pp. 1050-1059.
- Kang, C. G., Park, S. H. and Kim, J. W. 2010. Design of adaptive modulation and coding scheme for truncated hybrid ARQ. *Wireless Personal Communications*, Vol. 53, No. 2, pp. 269–280.
- Kundaeli, H. N. 2010. Throughput analysis of ARQ schemes using state transition diagrams. *Journal of Science and Technology*, Vol. 30, No. 2, pp. 165-176.
- Kundaeli, H. N. 2013. Throughput-delay analysis of the SR-ST-GBN ARQ scheme. *Mediterranean Journal of Electronics and Communications*, Vol. 9, No. 1, pp. 503-513.
- Larsson, L., Rasmussen, L. K. and Skoglund, M. 2014. Throughput analysis of ARQ schemes in gaussian block fading channels. *IEEE Transactions on Communications*, Vol. 62, No. 7, pp. 2569-2588.
- Larsson, L., Rasmussen, L. K. and Skoglund, M. 2016. Throughput analysis of hybrid-ARQ —A matrix exponential distribution approach. *IEEE Transactions on Communications*, Vol. 64, No. 1, 416-428.
- Liu, Q., Zhou, S. and Giannakis, G. B. 2004. Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links. *IEEE Transactions on Communications*, Vol. 52, No. 5, pp. 1746-1755.
- Malhotra, R., Gupta, V. and Bansal, D. 2011. Simulation & performance analysis of wired and wireless computer networks. *Global Journal of Computer Science and Technology*, Available at: <<https://computerresearch.org/index.php/computer/article/view/676>>. Date accessed: 17 May 2019.
- Malkamäki, E. and Leib, H. 2000. Performance of truncated Type-II Hybrid ARQ schemes with noisy feedback over block fading channels. *IEEE Transactions on Communications*, Vol. 48, No. 9, pp.1477-1487.

- Pack, S., Shen, X., Mark, J. W. 2008. Optimizing truncated ARQ scheme over wireless fading channels. *IEEE Transactions on Vehicular Technology*, Vol. 57, No. 2, pp. 1302-1305.
- Sassioui, R., Jabi, M., Leszek Szczecinski, L., Long Bao Le, L. B., Benjillali, M. and Pelletier, B. 2017. HARQ and AMC: Friends or foes? *IEEE Transactions on Communications*, Vol. 65, No. 2, pp. 635-650.
- Szczecinski, L., Khosravirad, S. R., Duhamel, P. and Rahman, M. M. 2013. Rate Allocation and adaptation for incremental redundancy truncated HARQ. *IEEE Transactions Communications*, Vol. 61, No. 6, pp. 2580-2590.
- Yangi, Y., Chen, W., Li, O., Liu, Q. and Hanzo, L. 2016. Truncated-ARQ aided adaptive network coding for cooperative two-way relaying networks: Cross-layer design and analysis. *IEEE Access*, Vol.4, pp. 9361-9376.
- Yang, Y., Ma, H. and Aïssa, S. 2012. Cross-Layer combining of adaptive modulation and truncated ARQ under cognitive radio resource requirements. *IEEE Transactions on Vehicular Technology*, Vol. 61, No. 9, pp. 4020-4030.

Biographical notes

Herald N. Kundaeli received his B. Sc., M.Sc. and PhD degrees in 1979, 1982 and 1996 respectively. He has conducted teaching and supervision of undergraduate and postgraduate students, research and consultancies in electronics and telecommunications at the University of Dar es Salaam since 1979. He is currently an Associate Professor in the Department of Electronics and Telecommunication Engineering at the University of Dar es Salaam, Tanzania. His expertise and research interests are in electronics, wireless communication and optical fibre communication; with specialities in synchronisation and error control, signal processing, and the applications of mathematical models to interdisciplinary research.

Received September 2019

Accepted April 2020

Final acceptance in revised form April 2020