Nonbinary Type-II Hybrid ARQ in Rician Fading Channels

Lijun Zhang
School of Electronics &
Information Engineering
Beijing Jiaotong University
Beijing, 100044, CHINA
Email: zhanglijun99@tsinghua.org.cn

Victor O. K. Li
Dept. of Electrical &
Electronic Engineering
The University of Hong Kong
Hong Kong, CHINA
Email: vli@eee.hku.hk

Zhigang Cao
Dept. of Electronic Engineering
Tsinghua University
Beijing, 100084, CHINA
Email: czg-dee@tsinghua.edu.cn

***Abstract*— Type-II Hybrid ARQ (HARQ) is extended to the nonbinary case, where shortened half-rate RS code is used in the FEC subsystem instead of binary BCH code. RS code is a powerful burst-error-correction code and has the biggest minimum distance among all the codes with the same code length, thereby providing more flexible adaptation. The throughput, reliability, and delay of nonbinary type-II HARQ with M-ary PSK (MPSK) over Rician fading channel are analyzed. Binary HARQ is analyzed also for comparison. The analytical results show that nonbinary HARQ outperforms its binary counterpart with respect to throughput, reliability, and average delay, when the modulation scheme, and the error-detection and error-correction codes are selected properly.**

I. INTRODUCTION

In data communications, automatic-repeat-request (ARQ) and forward-error-correction (FEC) are the two basic schemes for error control. Hybrid ARQ (HARQ) is obtained by combing ARQ with FEC, thus inheriting the strengths and weaknesses of ARQ and FEC [1]. In an HARQ system, the FEC subsystem reduces the frequency of retransmission by correcting the error patterns which occur most frequently, thus increasing the system throughput. If an error pattern exceeds the error-correction capability of the FEC, the receiver requests a retransmission. HARQ can provide higher throughput than ARQ alone, and almost the same reliability of ARQ if proper codes are chosen. It has been approved by 3GPP [2].

The throughput and reliability of HARQ have been analyzed over binary symmetric channel (BSC) with shortened BCH codes [1], [3]. We extend the analysis to the nonbinary case, where M-ary PSK (MPSK) modulation and shortened RS codes are considered over a slow, flat, Rician fading channel. In addition, the average delay of type-II HARQ is analyzed.

II. NONBINARY TYPE-II HYBRID ARQ

In type-II HARQ, two linear codes are used. One is a high rate $(n_0,\ k_0)$ code C_0, which is designed for error detection only. The other is a half-rate *invertible* $(2k_1,\ k_1)$ code C_1, which is used for error-correction and error-detection simultaneously. We choose shortened RS codes as C_1 in our HARQ schemes. Next, we will prove the validity of the *invertible process* for RS $(2k_1,\ k_1)$, which is the key of the nonbinary type-II HARQ.

RS codes are maximum-distance-separable (MDS) codes, and have the property $d_{\min} = n-k+1$, where n, k, and $d_{\min}$ are the code length, information length, and minimum distance of the codes, respectively. When a linear code is shortened, the shortened code has at least the same $d_{\min}$ as the original code. So, RS $(2k_1,\ k_1)$ has $d_{\min} \geq k_1+1$. If all the parity-check symbols are error-free, we assume that the information symbols are all zeros. The assumption causes at most k_1 symbol errors whose locations are known, i.e., erasures. Such error pattern is correctable as long as the number of erasures is less than $d_{\min}$. Obviously, $k_1 < d_{\min}$, and the *invertible process* works.

Corresponding to the nonbinary code symbol, MPSK modulation is naturally used in the system. Additionally, MPSK is more robust against fading than other modulations [4]. Let $m = \log_2 M$, and each symbol of GF(2^{lm})-defined RS code consists of lm bits. When mapping the code symbol to channel symbol, each code symbol is represented by the concatenation of l consecutive channel symbols. Using this mapping, long and powerful RS codes can be combined with the small-constellation modulations [5]. We assume that the alphabet of C_0's code symbol has the same size as that of MPSK, then

$$k_0 = lk_1.$$

When a message D of mk_0 information bits is ready for transmission, it is encoded into a codeword $(D,\ Q)$ of mn_0 bits based on the error-detecting code C_0, where Q represents the parity-check part of $m(n_0-k_0)$ bits. At the same time, D is also encoded by code C_1 into $(D,\ P(D))$, where $P(D)$ is the redundant part of D with the same length as D.

At first, codeword $(D,\ Q)$ is transmitted and the received word is $(\tilde{D},\ \tilde{Q})$. If no errors are detected, the word will be assumed error-free and accepted by the receiver; otherwise, the erroneous message $\tilde{D}$ is saved and a NACK is sent to the source. Upon receiving this NACK, the transmitter encodes the parity-check part $P(D)$ into a codeword $(P(D),\ Q^{(1)})$ based on C_0, and sends the parity word. The received parity word $(\tilde{P}(D),\ \tilde{Q}^{(1)})$ will be also examined according to C_0. If no errors are detected, the message D is retrieved by the *invertible* process. If errors are detected, we combine $\tilde{P}(D)$ with the previously saved $\tilde{D}$ to correct errors. If the error pattern in

$(\tilde{D},\ \tilde{P}(D))$ is correctable, the message D is recovered and accepted by the receiver; otherwise, $\tilde{D}$ is discarded and $\tilde{P}(D)$ is buffered. At the same time, the second NACK is sent back for information word retransmission. The retransmission alternates between the parity and the information words until the message is received error-free, or recovered by inversion, or by the decoding process.

III. Throughput, Reliability, and Delay

In [3], the throughput and reliability of BCH coded type-II HARQ over BSC have been analyzed. We extend the analysis to the nonbinary case in Rician fading channels. For simplicity, perfect interleaving of channel symbols and error-free feedback channel are assumed.

A. Throughput analysis

Assume that p_0 and p_1 are channel symbol error rate and code C_1's symbol error rate, respectively. We have $p_0 = p_1$ in the binary case. The probability p_c that a C_0-coded word will be transmitted error-free is

$$p_c = (1-p_0)^{n_0}. \tag{1}$$

The probability p_e that a C_0-coded word includes undetectable error patterns can be ignored if C_0 is properly selected. Let the roundtrip delay of the channel be τ and the information data rate be c, the number of C_0-coded words transmitted during time τ is simply

$$N = \frac{c\tau}{k_0}. \tag{2}$$

If the buffer capacity of the receiver is set to N, the lower bounds of the throughput for binary type-II HARQ can be written as [3]

$$\eta \geqslant \frac{k_0}{n_0} \cdot \frac{\delta_0}{\delta_0 + \delta_1 + N\delta_2}. \tag{3}$$

where

$$\begin{aligned}\delta_0 &= \frac{\gamma}{1-\beta}(1-\gamma\beta^{N-1}),\\ \delta_1 &= p_c q_1 \alpha^{N-2} + p_c(1-q_1)q_2\beta^{N-2},\\ \delta_2 &= 2 - p_c q_1 \alpha^{N-2} - \alpha\gamma\beta^{N-2},\end{aligned}$$

and

$$\begin{aligned}\alpha &= 1-(1-p_c)(1-q_1),\\ \beta &= 1-(1-p_c)(1-q_1)(1-q_2),\\ \gamma &= q_1 + q_2 - q_1 q_2.\end{aligned}$$

In above equations, q_1 and q_2 are the key terms. q_1 is the conditional probability that a message D is recovered successfully from the first received parity block $\tilde{P}(D)$ either by the inversion process or by the decoding process on the code C_1, given that errors are detected in the received word $(\tilde{D},\ \tilde{Q})$. q_2 is the conditional probability that a message D is successfully recovered from the second received information word $(\tilde{D}',\ \tilde{Q}')$ given that errors are detected in the received parity word $(\tilde{P}(D),\ \tilde{Q}^{(1)})$ and the first received information word $(\tilde{D},\ \tilde{Q})$, and that the received parity word $(\tilde{P}(D),\ \tilde{Q}^{(1)})$ fails to recover D, but detects the presence of errors in $(\tilde{D},\ \tilde{P}(D))$.

$$\begin{aligned}q_1 &= p_c + \frac{1}{1-p_c-p_e}\left\{\sum_{i=0}^{t}\binom{2k_0}{i}p_0^i(1-p_0)^{2k_0-i}\right.\\ &\quad +(1-p_0)^{2n_0} - 2(1-p_0)^{n_0}\\ &\quad \left.\cdot\sum_{i=0}^{t}\binom{k_0}{i}p_0^i(1-p_0)^{k_0-i}\right\},\end{aligned} \tag{4}$$

where t is the error-correction capacity of C_1.

$$q_2 = p_c + \frac{1}{(1-p_c)(1-q_1)}\cdot\sum_{i=1}^{t-1}\Delta_i S_{t-i}(1-\Delta_0 - S_{t-i}), \tag{5}$$

where

$$\Delta_i \triangleq \binom{k_0}{i}p_0^i(1-p_0)^{k_0-i}, \tag{6}$$

and

$$S_j = \sum_{l=1}^{j}\Delta_l. \tag{7}$$

Now we extend the above expressions to the nonbinary case. As mentioned in the previous section, C_0 is an $(n_0,\ k_0)$ linear code, C_1 is a shortened RS $(2k_1,\ k_1)$ code with error correction capacity of t, and $p_0 \neq p_1$. The main modifications are in (4) and (6). For clarity, recall that

$$q_1 = p_c + \frac{\Phi_1+\Phi_2+\Phi_3+\Phi_4}{1-p_c-p_e}, \tag{8}$$

where

$$\Phi_1 = \Pr\left\{\begin{array}{l}\text{at least one error in } \tilde{D},\ \text{at least one}\\ \text{error in } \tilde{P}(D),\ \text{and the total number of}\\ \text{errors in } \tilde{D} \text{ and } \tilde{P}(D) \text{ does not exceed } t\end{array}\right\}$$

$$\Phi_2 = \Pr\left\{\begin{array}{l}\text{no errors in } \tilde{D},\ \text{at least one error in } \tilde{Q},\\ \text{and the number of errors in } \tilde{P}(D) \text{ is}\\ \text{at least one, but no more than } t\end{array}\right\}$$

$$\Phi_3 = \Pr\left\{\begin{array}{l}\text{no errors in } \tilde{P}(D),\ \text{at least one error in}\\ \tilde{Q}^{(1)},\ \text{and the number of errors in } \tilde{D}\\ \text{is at least one, but no more than } t\end{array}\right\}$$

$$\Phi_4 = \Pr\left\{\begin{array}{l}\text{no errors in } \tilde{D},\ \text{no errors in } \tilde{P}(D),\\ \text{at least one error in } \tilde{Q},\ \text{and at least}\\ \text{one error in } \tilde{Q}^{(1)}\end{array}\right\}.$$

For nonbinary HARQ, we obtain

$$\begin{aligned}\Phi_1 &= \sum_{i=2}^{t}\binom{2k_1}{i}p_1^i(1-p_1)^{2k_1-i} - 2(1-p_1)^{k_1}\\ &\quad \cdot\sum_{i=2}^{t}\binom{k_1}{i}p_1^i(1-p_1)^{k_1-i},\end{aligned} \tag{9}$$

$$\begin{aligned}\Phi_2 &= \Phi_3 = (1-p_1)^{k_1}[1-(1-p_0)^{n_0-k_0}]\\ &\quad \cdot\sum_{i=1}^{t}\binom{k_1}{i}p_1^i(1-p_1)^{k_1-i},\end{aligned} \tag{10}$$

and

$$\Phi_4 = (1-p_1)^{2k_1}[1-(1-p_0)^{n_0-k_0}]^2. \quad (11)$$

Substituting (9), (10) and (11) into (8) and considering that $p_1 = 1-(1-p_0)^l$ and $k_0 = lk_1$, we get

$$\begin{aligned} q_1 &= p_c + \frac{1}{1-p_c-p_e}\left\{\sum_{i=0}^{t}\binom{2k_1}{i}p_1^i(1-p_1)^{2k_1-i}\right. \\ &\quad +(1-p_0)^{2n_0} - 2(1-p_0)^{n_0} \\ &\quad \left.\cdot\sum_{i=0}^{t}\binom{k_1}{i}p_1^i(1-p_1)^{k_1-i}\right\}, \end{aligned} \quad (12)$$

Similarly, (6) becomes

$$\Delta_i \triangleq \binom{k_1}{i}p_1^i(1-p_1)^{k_1-i}. \quad (13)$$

Substituting (13) into (7) and (5) yields q_2.

B. Reliability analysis

In HARQ, if code C_0 is properly chosen, p_e can be made very small and upper bounded as follows [3]:

$$p_e \leqslant (1-p_c)\cdot 2^{-m(n_0-k_0)}. \quad (14)$$

Assume the shortened RS $(2k_1,\ k_1)$ code has the capacity of detecting at most d $(d>t)$ errors, the probability that the number of errors in a span of $2k_1$ code symbols exceeds the error-detection capacity is simply

$$\sigma \triangleq \sum_{l>d}^{2k_1}\binom{2k_1}{l}p_1^l(1-p_1)^{2k_1-l}. \quad (15)$$

The probability of error $\Pr(\xi)$ for the HARQ scheme is bounded by [3]:

$$\frac{p_e}{p_e+p_c} \leqslant \Pr(\xi) \leqslant \frac{p_e+\sigma}{p_e+p_c}. \quad (16)$$

The leftmost term is just the probability of error for ideal selective-repeat (SR) ARQ scheme. We observe that the reliability performance of type-II HARQ is very close to that of SR HARQ when σ is small.

C. Average delay

In type-II HARQ, the probability that an information word is transmitted successfully for the first time is p_c. Once errors occur in the received word, the transmission of parity check word will be triggered, and we call this the first retransmission. The probability that information is correctly obtained after the first retransmission is $(1-p_c)q_1$, and the probability that information is successfully retrieved after the second retransmission is no more than $(1-p_c)^2q_1$. Similarly, we can obtain the upper bounds of the probabilities that information

TABLE I

PARAMETERS OF TYPE-II HARQS.

Scheme	Invertible codes	Buffer size	Modulation
I	BCH(1000,500)	123	BPSK
II	RS(60,30)	342	QPSK
III	RS(60,30)	342	8PSK
IV	RS(400,200)	35	8PSK

is recovered after the third, fourth, etc. retransmissions, and the upper bound of the average delay can be expressed as:

$$\begin{aligned} \upsilon &\leq \frac{\tau}{2}p_c + \frac{3\tau}{2}(1-p_c)q_1 + \frac{5\tau}{2}(1-p_c)^2q_1 \\ &\quad + \frac{7\tau}{2}(1-p_c)^3q_1 + \cdots \\ &= \frac{\tau}{2}p_c + \frac{\tau}{2}q_1\sum_{i=1}^{\infty}(1+2i)(1-p_c)^i \\ &= \frac{\tau}{2}\left(p_c + q_1\frac{2-p_c-p_c^2}{p_c^2}\right). \end{aligned} \quad (17)$$

Divide υ by τ to get the normalized delay

$$\hat{\upsilon} \leq \frac{1}{2}\left(p_c + q_1\frac{2-p_c-p_c^2}{p_c^2}\right). \quad (18)$$

When $p_c \to 1$, $\hat{\upsilon} \to \frac{1}{2}$.

D. Coherent MPSK in Rician fading channels

The throughput, reliability, and delay are all related to the channel symbol error rate p_0. The performance of MPSK in Rician fading channels is very difficult to obtain, and one has to resort to approximations. Fortunately, Sun and Reed found closed-form solutions for the average error rate of coherent MPSK modulation scheme over slow, flat, Rician fading channel with Rician parameter K [6]

$$p_0 = \frac{1}{\pi}\left(\frac{1+K}{\gamma_s}\right) \cdot\int_{-(\pi/2)}^{(\pi/2)} \frac{\exp\left(-\dfrac{K\sin^2\frac{\pi}{M}\sec^2\theta}{\frac{1+K}{\gamma_s}+\sin^2\frac{\pi}{M}\sec^2\theta}\right)}{\frac{1+K}{\gamma_s}+\sin^2\frac{\pi}{M}\sec^2\theta}\mathrm{d}\theta, \quad (19)$$

where γ_s denotes the average signal to noise ratio (SNR) of received signals.

IV. NUMERICAL RESULTS

We assume the roundtrip delay τ and data rate c to be 30 ms and 2 Mbps, respectively [7]. The other parameters are listed in Table I. Each scheme uses a shortened Hamming code with 24-bit parity check to detect errors. Scheme I is just the binary type-II HARQ, proposed by Lin and Yu [3]. Schemes II through IV are nonbinary HARQs, and each receiver buffer size N is derived by (2). The Rician parameter K equals 5 dB.

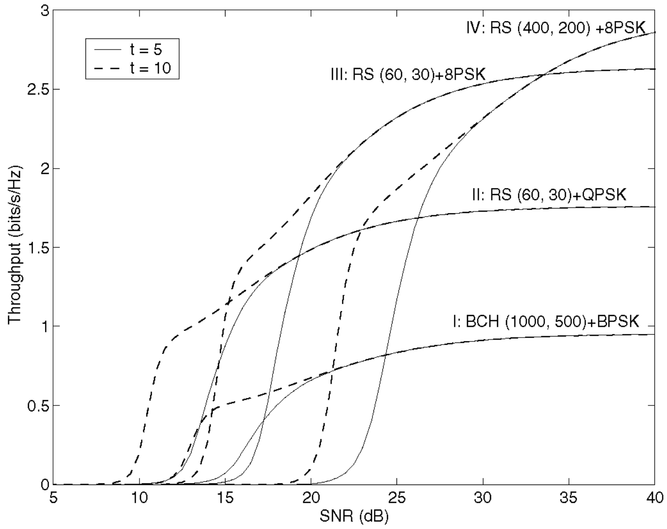

Fig. 1. Throughput comparison of Schemes I through IV.

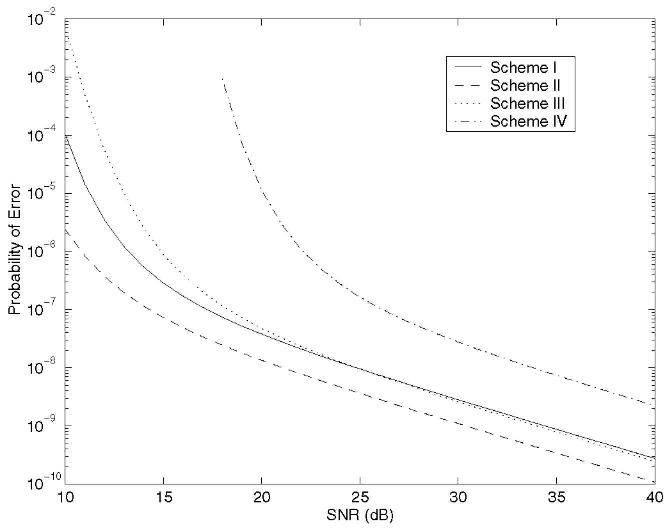

Fig. 2. Reliability of Schemes I through IV with $t = 5$.

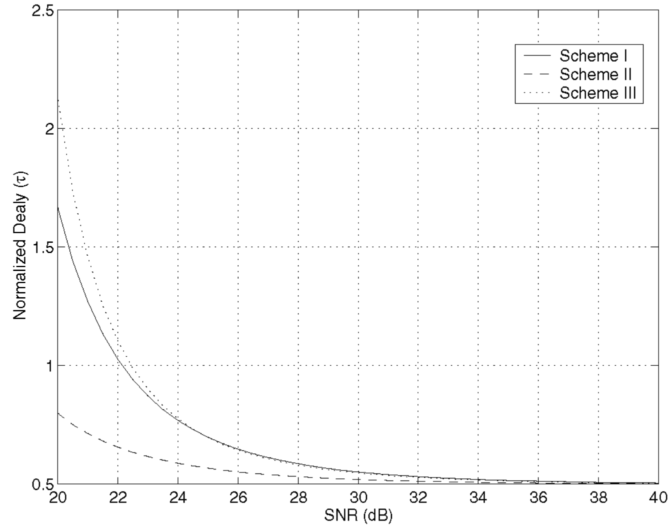

Fig. 3. Upper bounds of normalized delays of Schemes I, II, and III with $t = 5$.

The throughputs of Schemes I through IV with $t = 5,\ 10$, are illustrated in Fig. 1. Obviously, the nonbinary HARQ schemes greatly outperform the binary schemes in throughput. Scheme II shows superior throughput performance over the binary case in the whole range of SNR considered. There exists a crossover between the throughput curves of Schemes II and III. Before the crossover point, the throughput of Scheme III is lower than that of II, and after the crossover point, the throughput of Scheme III exceeds that of II by a large margin. The two nonbinary HARQs can compensate for each other. As for Scheme IV, poor throughput can be observed at low SNR, say, 20 dB. The advantage of Scheme IV, which uses long RS code, becomes obvious with the increase of SNR. The throughput of Scheme IV outperforms all other schemes as SNR increases beyond 33.6 dB.

We also see the floors of all the schemes at low SNR's, especially for Scheme IV. The phenomenon can be explained as follows. On the one hand, the modulation used in Scheme IV is 8PSK, which has relatively high symbol error rate compared to QPSK and BPSK. On the other hand, RS (400, 200) code is defined over GF(2^9), and each code symbol is represented by the concatenation of three consecutive 8PSK symbols, potentially incurring more errors than in other schemes. At low SNR's, 8PSK modulation gives high channel symbol error rate, and hence high word error rate. If we still set $t = 5$ or $t = 10$, error patterns exceeding the error-correction capacity will occur at high frequency, which incurs many retransmissions before a word is accepted by the receiver, thus degrading the throughput to very near zero (the floor). The floor effect can be mitigated by increasing the error-correction capacity t of RS (400, 200).

All the scenarios use a shortened Hamming code with 24-bit parity check to detect errors. Fig. 2 depicts the reliability of all the schemes. The nonbinary HARQs, except Scheme IV, have very similar, or even higher reliability than binary HARQ, Scheme I. The probability of error for Scheme IV is unsatisfactory, especially at low SNR's. We can enhance the reliability of the scheme by appending more bits to detect errors, such as 32-bit parity check instead of 24-bit parity check.

The upper bounds of normalized delay of Schemes I, II, and III are illustrated in Fig. 3. The error correction capacity of all the HARQs is set to five, corresponding to the throughput curves in Fig. 1. Notice that Scheme IV is excluded due to its severe floor behavior at low SNR's, which will cause intolerable delay and render it useless in some cases. Similar to the reliability comparison, nonbinary schemes outperform the binary counterparts with respect to delay. For example, normalized delays will not exceed 1.67, 0.80, and 2.12 τ for Schemes I, II, and III, respectively, when SNR is around 20 dB. As SNR increases, the delays of the three schemes gradually approach the limit of 0.5.

The above numerical results help us choose suitable scheme according to different requirements. If bandwidth utilization is the primary goal, Schemes III or IV is recommended. If reliability is the most important factor, Scheme II will be the best choice. If the application has strict time-constraint, Scheme II will be selected.

V. Conclusions

We extend Lin and Yu's work to nonbinary type-II HARQ by combining shortened RS codes with MPSK modulation. The throughput, reliability, and delay are analyzed.

HARQ Schemes I through IV are considered over Rican fading channels. The numerical results show that nonbinary schemes (Schemes II, III, and IV) outperform the binary case (Scheme I) by a large margin, and provide more flexible choice according to different requirements. In particular, Schemes III and IV have high bandwidth utilization, while Scheme II caters to the need of strict reliability and time-constraint.

Invertible code-based HARQ has very simple structure, but it can achieve very good throughput and reliability. Furthermore, this kind of HARQ can also be embedded in MIMO systems, and turbo processing technology can be used to enhance the decoding of RS codes.

Acknowledgment

This research is supported in part by the National Natural Science Foundation of China (No. 6001161943)/Research Grants Council of Hong Kong (No. N_HKU 709/00) Joint Research Scheme, and Research Foundation of Beijing Jiaotong University under contract number 2003RC042.

References

[1] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*. New York: Prentice Hall, 1983.

[2] 3GPP. (2001) TS 25.301: Radio interface protocol architecture (release 5). [Online]. Available: http://www.cwts.org

[3] S. Lin and P. S. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels," *IEEE Trans. Commun.*, vol. 30, pp. 1701–1719, July 1982.

[4] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: John Wiely & Sons, 1965.

[5] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. Boston: Kluwer Academic publishers, 1994.

[6] J. Sun and I. S. Reed, "Performance of MDPSK, MPSK, and noncoherent MFSK in wireless Rician fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 813–816, June 1999.

[7] S. Aramvith, C. W. Lin, S. Roy, and M. T. Sun, "Wireless video transport using conditional retransmission and low-delay interleaving," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 558–565, June 2002.