Jointly Adaptive Modulation and Packet Retransmission over Block Fading Channels with Robustness to Feedback Latency

Qingwen Liu, Shengli Zhou and Georgios B. Giannakis[1]
Department of Electrical and Computer Engineering
University of Minnesota
200 Union Street SE
Minneapolis, MN 55455

Abstract — **We develop an adaptive time-diversity system equipped with adaptive modulation and retransmissions over Nakagami-m block fading channels. Our scheme maximizes the average spectral efficiency, and meets the prescribed packet error rate (PER) based on both *instantaneous* signal-to-noise ratio (SNR), and *statistical* channel state information (CSI). It does not constrain the feedback latency to be less than the channel coherence time, which is a stringent requirement in most existing adaptive modulation schemes. Multiple-packet retransmissions are provided via at most one retransmission request. This minimizes round-trip-delay and buffer-size requirements, compared to multiple retransmission requests needed in auto repeat request (ARQ) protocols based on cyclic redundancy check (CRC). We obtain closed-form expressions for the average number of retransmissions per packet, and average spectral efficiency. We also express the optimal values of these parameters as functions of the average SNR, and the Nakagami parameter m.**

I. Introduction

Adaptive modulation (AM) has been widely adopted by many wireless communication standards to enhance throughput [1, 4]. AM maximizes transmission rates for a prescribed error performance over time-varying channels. However, being sensitive to feedback latency τ_f, it is impossible to implement AM when τ_f exceeds the channel coherence time τ_c [2].

Auto repeat request (ARQ) protocols on the other hand, impose no feedback latency constraint (FLC) $\tau_f < \tau_c$, and achieve higher throughput than forward error control (FEC), by providing time diversity to mitigate fading when necessary [6]. However, retransmitted packets may also suffer from fading effects, which necessitates a large number of retransmissions for certain packets. This, in turn, may lead to unacceptably large round-trip-delays and buffer sizes. So, ARQ may not meet the required quality of service, if the system can not satisfy the stringent delay or buffer-size requirements in certain real-time applications, such as video transmissions. Furthermore, CRC introduces redundancy, as well as complexity, and delay in decoding.

A natural question is whether and how one can design a transmission scheme that maximizes throughput without FLC, while adhering to prescribed performance and delay constraints. Such a scheme should also provide performance guarantees and real-time quality of service for mobile terminals, under large feedback latency conditions, e.g., "video-to-car". In this paper, we develop an adaptive time-diversity system equipped with jointly adaptive modulation and adaptive retransmission (JAMAR), which satisfies these requirements. It is simple to design, and does not require CRC, which results in less redundancy, decoding complexity, and delay.

[1]This work was supported by the NSF grant no. 0105612, and by the ARL/CTA grant no. DAAD19-01-01-2-011.

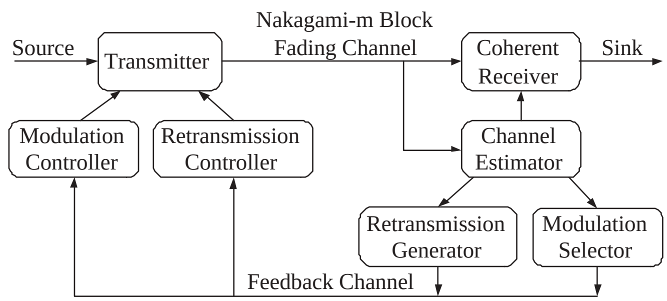

Figure 1: System and Channel Models

II. System and Channel Models

Consider the single-transmit single-receive antennae system in Fig. 1, which includes an adaptive modulation module, and an adaptive time-diversity retransmission module.

Because it captures a large class of fading channels, we adopt the Nakagami-m block flat fading channel model with additive white Gaussian noise (AWGN). The power variation of both transmitter and channel is jointly accounted for in the receive SNR γ [5], which is assumed to be constant over the duration of each packet that consists of multiple symbols. As in [3], γ is assumed to be i.i.d. from packet to packet. The Nakagami-m channel induces the following probability density function (pdf) for the receive SNR γ:

$$p_\gamma(\gamma) = \frac{m^m \gamma^{m-1}}{\bar{\gamma}^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right), \tag{1}$$

where $\bar{\gamma} := E\{\gamma\}$ is the average receive SNR, $\Gamma(m) := \int_0^\infty t^{m-1}\exp(-t)dt$ is the gamma function, and m is the Nakagami fading parameter ($m \geq 1/2$). The Nakagami-m model subsumes Rayleigh distributed channels ($m = 1$), and when $m > 1$, it approximates closely Rician channels as well [8].

In the adaptive modulation module, multiple transmission modes, i.e., different constellation sizes of multi-level quadrature amplitude modulation (M-QAM), $M_n = 2^n$, $n \in S_n = \{0, 1, \ldots, 7\}$, are available. Based on statistical CSI (average SNR $\bar{\gamma}$ and Nakagami parameter m) acquired at the receiver, the modulation selector determines the modulation mode n, which is sent back to the transmitter through a feedback channel. The modulation controller then updates the mode at the transmitter.

In the adaptive time-diversity retransmission module, an adaptive retransmission protocol is implemented. If the number of retransmissions N for a certain packet is determined at the receiver, a N-retransmission request is generated by the retransmission generator, and is communicated to the retransmission controller at the transmitter via a feedback channel. Then the controller arranges retransmissions of the requested packets that are stored in the buffer, with the same modulation mode as the original one. The retransmitted packets for a certain packet are not necessarily consecutive and the intervals between retransmissions are arranged to be larger than the channel coherence time. The original and retransmitted packets are combined with maximum ratio combining (MRC) at the receiver. Coherent demodulation is used for the combined packets.

Our operating assumptions are:

A1: The channel is frequency-flat fading. Instantaneous SNR γ remains invariant per packet, but is allowed to vary from packet to packet. This corresponds to the so called block fading channel model [3].

A2: The statistical CSI is assumed to be invariant over the duration of the multiple transmissions for a certain packet, because it is affected by shadowing for instance, which varies much slower relative to fading. The possibility that multiple transmissions for a certain packet experience different statistical CSI is negligible.

A3: Perfect CSI (both instantaneous and statistical) is available at the receiver. The corresponding mode selection and the number of retransmissions N are fed back to the transmitter without error, and without imposing any FLC.

A4: The fading channel coefficients corresponding to the original and the retransmitted packets are independent and identically distributed (i.i.d.), because the time intervals between multiple transmissions for a certain packet are larger than the channel coherence time.

III. Jointly Adaptive Modulation-Retransmission

In this section, we develop the jointly adaptive modulation and adaptive retransmission scheme, and derive the number of retransmissions and average spectral efficiency in closed-form.

We impose the following performance and retransmission constraints:

C1: The average packet error rate after combining is no larger than PER_0. Thus, the corresponding average bit error rate is $P_{b0} = 1 - (1 - \mathrm{PER}_0)^{1/N_p}$, where N_p is a fixed packet length (bits/packet).

C2: At most one retransmission request per packet is generated. Thus, the delay is bounded to be less than one round-trip-delay, because the duration of multiple retransmissions per packet is usually much less than the round-trip-delay [7].

We denote the instantaneous receive SNR as γ. The BER of M-QAM over an AWGN channel can be approximated as [5]:

$$\begin{aligned} P_b(n,\gamma) &\approx 0.2\exp[-1.5/(2^n-1)\gamma] \\ &:= A\exp(-d_n\gamma), \end{aligned} \tag{2}$$

where $A := 0.2$, and $d_n := 1.5/(2^n - 1)$ for $n > 1$. When $n = 1$, we use $P_b(n = 1, \gamma) \approx 0.2\exp(-\gamma)$ for BPSK, where $d_1 = 1$. Suppose that the original packet has channel coefficient γ_i, and $N(\gamma_i)$ packets are retransmitted. Let $\{\gamma_{ij}^{(r)}\}_{j=1}^{N(\gamma_i)}$ denote the corresponding channel coefficient of these retransmitted packets. Based on (2), the BER after MRC can be obtained as:

$$\begin{aligned} &P_b(n,\gamma_i,\gamma_{i1}^{(r)},\ldots,\gamma_{iN(\gamma_i)}^{(r)}) \\ &= A\exp\left[-d_n\left(\gamma_i + \sum_{j=1}^{N(\gamma_i)}\gamma_{ij}^{(r)}\right)\right]. \end{aligned} \tag{3}$$

When deciding how many retransmissions are needed, the receiver treats $\{\gamma_{ij}^{(r)}\}_{j=1}^{N(\gamma_i)}$ as random variables, and evaluates the average BER as:

$$\begin{aligned} \overline{P_b}(n,\gamma_i) &= \int_0^\infty \cdots \int_0^\infty P_b(n,\gamma_i,\gamma_{i1}^{(r)},\ldots,\gamma_{iN(\gamma_i)}^{(r)}) \\ &\cdot p_{\gamma_{i1}^{(r)},\ldots,\gamma_{iN(\gamma_i)}^{(r)}}\left(\gamma_{i1}^{(r)},\ldots,\gamma_{iN(\gamma_i)}^{(r)}\right) \\ &d\gamma_{i1}^{(r)},\ldots,d\gamma_{iN(\gamma_i)}^{(r)} \\ &= A\exp(-d_n\gamma_i)\prod_{j=1}^{N(\gamma_i)}\int_0^\infty \exp\left(-d_n\gamma_{ij}^{(r)}\right) \\ &\cdot p_{\gamma_{ij}^{(r)}}\left(\gamma_{ij}^{(r)}\right)d\gamma_{ij}^{(r)} \\ &= A\exp(-d_n\gamma_i)\prod_{j=1}^{N(\gamma_i)}\left(1+\frac{d_n}{m}\bar{\gamma}\right)^{-m} \\ &= A\exp(-d_n\gamma_i)\left(1+\frac{d_n}{m}\bar{\gamma}\right)^{-mN(\gamma_i)}. \end{aligned} \tag{4}$$

Due to performance constraint C1, the average BER should satisfy:

$$\overline{P_b}(n,\gamma_i) \le P_{b0}. \tag{5}$$

Then, from (4) and (5), we find $N(\gamma_i)$ as:

$$\begin{aligned} N(\gamma_i) &= \frac{d_n/m}{\ln(1+d_n\bar{\gamma}/m)}\left[\frac{1}{d_n}\ln\left(\frac{A}{P_{b0}}\right)-\gamma_i\right] \\ &:= B(\gamma_{th}-\gamma_i), \end{aligned} \tag{6}$$

where

$$\begin{aligned} B &:= \frac{d_n}{m\ln(1+d_n\bar{\gamma}/m)}, \\ \gamma_{th} &:= \frac{\ln(A/P_{b0})}{d_n}. \end{aligned} \tag{7}$$

The instantaneous SNR threshold γ_{th} determines whether a packet should be retransmitted or not. If $\gamma_i < \gamma_{th}$, then $N(\gamma_i) > 0$, and retransmissions are initiated; otherwise, no retransmission is needed. We notice that the value of $N(\gamma_i)$ is not always an integer. Approximating $N(\gamma_i)$ with the nearest integer greater than or equal to $N(\gamma_i)$, $\lceil N(\gamma_i)\rceil$, introduces extra redundancy. This leads to better average BER performance, and constitutes a future topic we would like to investigate further. The following analysis benchmarks the performance of our proposed scheme.

If $N^*(\gamma_i)$ denotes the number of retransmitted packets, which is requested through the feedback channel, then

$$N^*(\gamma_i) := [N(\gamma_i)]_+ = \max[N(\gamma_i), 0]. \tag{8}$$

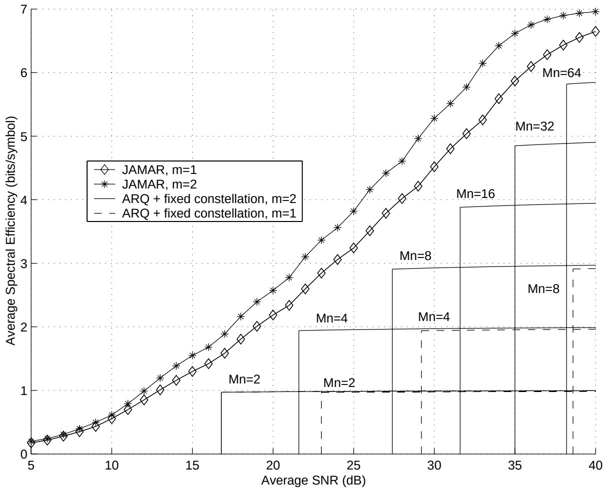

Figure 2: Average spectral efficiency vs. $\bar{\gamma}$

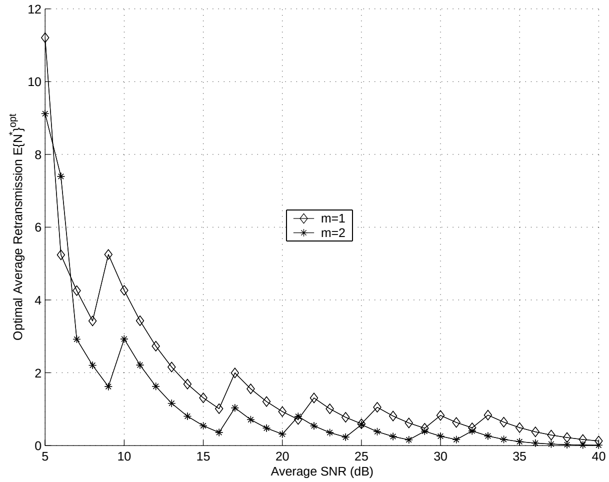

Figure 3: Average retransmission number vs. $\bar{\gamma}$

We denote its average value as $\overline{N^*} := E_{\gamma_i}\{N^*(\gamma_i)\}$. Thus, a closed-form expression for $\overline{N^*}$ can be obtained as:

$$\begin{aligned}\overline{N^*} &= \int_0^{\gamma_{th}} B(\gamma_{th}-\gamma_i)p(\gamma_i)d\gamma_i \\ &= B\gamma_{th}\left[1-\frac{\Gamma(m, m\gamma_{th}/\bar{\gamma})}{\Gamma(m)}\right] \\ &- B(\bar{\gamma}/m)\left[m-\frac{\Gamma(m+1, m\gamma_{th}/\bar{\gamma})}{\Gamma(m)}\right],\end{aligned} \tag{9}$$

where $\Gamma(m,x) := \int_x^\infty t^{m-1}\exp(-t)dt$ is the complementary incomplete Gamma function.

From (6), the average spectral efficiency (throughput) can be computed as:

$$\begin{aligned}\mathrm{S_e} &= \int_0^\infty \frac{\log_2(M_n)}{1+N^*(\gamma_i)}\ p_{\gamma_i}(\gamma_i)d\gamma_i \\ &= \log_2(M_n)\Big[\int_0^{\gamma_{th}} \frac{1}{1+N^*(\gamma_i)}\ p_{\gamma_i}(\gamma_i)d\gamma_i \\ &+ \int_{\gamma_{th}}^\infty p_{\gamma_i}(\gamma_i)d\gamma_i\Big] \\ &= \log_2(M_n)\left[\int_0^{\gamma_{th}} \frac{1}{1+B(\gamma_{th}-\gamma_i)}\cdot\frac{m^m\gamma_i^{m-1}}{\bar{\gamma}^m\Gamma(m)}\right. \\ &\left.\cdot\exp\left(-\frac{m\gamma_i}{\bar{\gamma}}\right)d\gamma_i + \frac{\Gamma(m, m\gamma_{th}/\bar{\gamma})}{\Gamma(m)}\right].\end{aligned} \tag{10}$$

Since $\overline{N^*}$ and $\mathrm{S_e}$ depend on the transmission mode $n \in S_n$, the average SNR $\bar{\gamma}$, and the Nakagami parameter m, we propose the following algorithm to maximize $\mathrm{S_e}$:

Step 1: For the mode $n \in S_n$, average SNR $\bar{\gamma}$, and Nakagami parameter m, calculate $\overline{N^*}$ and $\mathrm{S_e}$ using (9) and (10), respectively:

$$\overline{N^*}(n, \bar{\gamma}, m) \quad \text{and} \quad \mathrm{S_e}(n, \bar{\gamma}, m). \tag{11}$$

Step 2: Among the values in (11), select the optimal spectral efficiency, modulation mode, and corresponding average retransmission number at each $(\bar{\gamma}, m)$ as:

$$\begin{aligned}\mathrm{S_e^{opt}}(\bar{\gamma}, m) &= \max_{n\in S_n}[\mathrm{S_e}(n, \bar{\gamma}, m)], \\ n^{\mathrm{opt}}(\bar{\gamma}, m) &= \arg\max_{n\in S_n}[\mathrm{S_e}(n, \bar{\gamma}, m)], \\ \overline{N^*}^{\mathrm{opt}}(\bar{\gamma}, m) &= \overline{N^*}(n^{\mathrm{opt}}(\bar{\gamma}, m), \bar{\gamma}, m).\end{aligned} \tag{12}$$

Notice that $\mathrm{S_e^{opt}}(\bar{\gamma}, m)$, $n^{\mathrm{opt}}(\bar{\gamma}, m)$, and $\overline{N^*}^{\mathrm{opt}}(\bar{\gamma}, m)$ as functions of $(\bar{\gamma}, m)$, can be obtained off-line.

In summary, given C1 and C2, our novel scheme has two operational stages:

Stage 1: Based on the estimated average SNR $\bar{\gamma}$, and the Nakagami parameter m, switch the modulation mode to $n^{\mathrm{opt}}(\bar{\gamma}, m)$ during the given modulation switching period τ_s ($\tau_s \gg \tau_c$ usually) via feedback without FLC.

Stage 2: Based on the estimated instantaneous SNR γ_i, Nakagami parameter m, average SNR $\bar{\gamma}$, and modulation mode $n^{\mathrm{opt}}(\bar{\gamma}, m)$, determine $N(\gamma_i)$ via (6). Then, retransmit $N^*(\gamma_i)$ packets via retransmission request feedback without FLC.

IV. Numerical Results

Given $\mathrm{PER}_0 = 0.001$ and $N_p = 1000$, we have $P_{b0} \approx 10^{-6}$. The optimal spectral efficiency $\mathrm{S_e^{opt}}(\bar{\gamma}, m)$, and the average retransmission number $\overline{N^*}^{\mathrm{opt}}(\bar{\gamma}, m)$, obtained as in Step 2, are shown in Figs. 2 and 3 respectively, for $m = 1, 2$. The average number of retransmissions is less than 2, when the average SNR exceeds 14dB for $m = 1$, which can be afforded in practice.

In order to illustrate the advantage of our scheme, we compare it with the fixed-constellation M-QAM combined with the truncated ARQ. To guarantee the stringent delay and buffer-size constraints, the truncated ARQ allows at most one-retransmission (one-truncated ARQ). The average spectral efficiency is set to zero, if the average PER is larger than PER_0, in order to satisfy the error performance requirement. So, this scheme also satisfies performance and delay constraints without FLC. Comparison based on average spectral efficiency is shown in Fig. 2, for $m = 1, 2$, respectively. Evidently, our scheme improves average spectral efficiency.

V. Conclusions

We proposed a jointly adaptive modulation and adaptive retransmission scheme over Nakagami block fading channels. It maximizes the spectral efficiency for a prescribed average PER, without imposing feedback latency to be less than the wireless channel's coherence time. It has at most one retransmission loop, which reduces round-trip-delay and buffer sizes relative to existing multi-loop retransmissions. We obtain the average number of retransmissions and spectral efficiency in closed-form. Furthermore, we propose an algorithm to maximize spectral efficiency.

References

[1] 3GPP TR 25.848 V4.0.0, *Physical layer aspects of UTRA high speed downlink packet access (release 4)*, 2001.

[2] M. -S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Kluwer Journal on Wireless Communications*, vol. 13, no. 1–2, pp. 119–143, May 2000.

[3] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1273–1289, May 2001.

[4] A. Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, J. McGeehan, and P. Karlsson, "A comparison of the HIPERLAN/2 and IEEE 802.11a wireless LAN standards," *IEEE Communications Magazine*, vol. 40, no. 5, pp. 172–180, May 2002.

[5] A. J. Goldsmith and S. -G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1230, October 1997.

[6] E. Malkamaki and H. Leib, "Performance of truncated type-II hybrid ARQ schemes with noisy feedback over block fading channels," *IEEE Transactions on Communications*, vol. 48, pp. 1477–1487, September 2000.

[7] E. Malkamaki, D. Mathew, and S. Hamalainen, "Performance of hybrid ARQ techniques for WCDMA high data rates," in *Proc. of Vehicular Tech. Conf.*, vol. 4, pp. 2720–2724, 2001 Spring.

[8] G. L. Stüber, *Principles of Mobile Communication.* Massachusetts: Kluwer Academic Publishers, 2nd Edition, 2001.