Impact of Imperfect Channel State Information on RS Coding Aided Hybrid-ARQ in Rayleigh Fading Channels

Hoang Anh Ngo and Lajos Hanzo
School of Electronics & Computer Science, University of Southampton
http://www-mobile.ecs.soton.ac.uk
Email: {han08r, lh}@ecs.soton.ac.uk

Abstract— **Hybrid Automatic-Repeat-reQuest (ARQ) schemes are capable of maintaining reliable communications for transmission over hostile wireless channels. In coherent detection aided schemes, the achievable performance substantially depends on the accuracy of channel state information (CSI). In this paper, the impact of imperfect CSI on Reed-Solomon (RS) coded Hybrid-ARQ systems is investigated for transmission over block-Rayleigh fading channels. The corresponding goodput equation is derived and used to evaluate the system's performance. The effects of code rate and channel quality are also visualized. Furthermore, the results are utilized for the optimization of the pilot versus data symbols power allocation in order to maximize the system's goodput. Quantitatively, the proposed scheme provides a $2.5dB$ E_b/N_0 gain at a given goodput for the $(255/223)$ RS code defined over the $GF(256)$.**

Index Terms - Channel state information (CSI), Hybrid-ARQ (H-ARQ), Reed-Solomon (RS) code, goodput, error probability, power allocation.

I. INTRODUCTION

The Automatic-Repeat-reQuest (ARQ) protocol [1] provides an error-control method for data communication. It operates by using acknowledgements and time-outs to achieve reliable data transmission over the communication channels. In order to improve its efficiency, the ARQ protocol is usually combined with forward error correcting (FEC) and detecting codes. The combination is known as the Hybrid-ARQ protocol [2]. This protocol has been utilized extensively in contemporary communication systems, such as the Universal Mobile Telecommunications System (UMTS) and 3GPP Long Term Evolution (LTE) standards [3], or the IEEE 802.16 mobile WiMAX [4].

In wireless channels the transmitted signal typically experiences fading. In order to successfully recover the received signal, the channel state information (CSI), or fading coefficients have to be known by the receiver. In practice, we can only approximate the CSI by using pilot symbols.

The impact of imperfect CSI on ARQ systems was studied in [5] for transmission over Rayleigh fading channels, when using no error correction coding - only an error detecting code along with the ARQ protocol was invoked. However, practical wireless systems employ the H-ARQ protocol combined with

a strong FEC code. Therefore, in this paper, we will investigate an H-ARQ scheme, using Reed-Solomon codes for transmission over a block-fading Rayleigh channel. The proposed arrangement is is referred to as the RS/H-ARQ scheme.

First, we derive the associated goodput equation based on the achievable throughput and the accepted packet error ratio (APER). Due to the dependence of both the throughput and the APER on the CSI, the goodput also varies in accordance with the CSI. Thus, the accuracy of the CSI estimation has an important role in determining the achievable goodput. We use the mean square error (MSE) of the channel estimates as our quality metric. Moreover, it is plausible that the channel estimates' accuracy depends on the energy of pilot symbols. Thus, a power sharing between the uncoded pilot and coded data bits will be proposed in order to maximize the system's goodput. The results of Section IV will quantify the impact of CSI accuracy on the RS/H-ARQ systems' performance.

The outline of this paper is as follows. Section II describes the system model and assumptions, while Section III derives the goodput equation and a power allocation scheme for the RS/H-ARQ system communicating in Rayleigh fading channels. They are followed by numerical results in Section IV. The final section provides our conclusions.

II. SYSTEM MODEL

We will consider an H-ARQ scheme operating with the assistance of RS coding. At the transmitter, the information bits are grouped into blocks of m bits first, generating a symbol. Then a group of k information symbols is forwarded to an (n, k) RS encoder, which is defined over the finite Galois field $GF(2^m)$ and has the code rate of $R = k/n$. Subsequently, the encoded bits are concatenated to N_P pilot bits, which are known to both the transmitter and receiver. No error detection code is required as a benefit of the RS code's capability of both error detection and correction [6], provided that the code is sufficiently long [6]. As a result, each transmitted packet includes k m-bit information symbols, $(n - k)$ parity symbols and N_p pilot bits. Following modulation, the packet is transmitted over the channel. To focus our attention on the H-ARQ scheme, simple binary phase shift keying (BPSK) modulation is proposed.

This classic Selective Repeat ARQ (SR-ARQ) scheme employing packet buffers both at the transmitter and the receiver

is used. Again, a block-fading Rayleigh channel is assumed. Then, the received signal may be expressed as

$$y_P[k] = \sqrt{E_P} h x_P[k] + n[k], k = 1, ..., N_P, \qquad (1)$$
$$y_S[k] = \sqrt{E_S} h x_S[k] + n[k], k = N_P + 1, ..., N_P + n,$$

where k is the symbol index, $y_P[k]$, $y_S[k]$ are the received pilot and data symbols; $x_P[k]$, $x_S[k]$ are the transmitted pilot and data symbols; $n[k]$ is the zero-mean complex AWGN with power spectral density of N_0; h represents the Rayleigh fading channel coefficients of having the real part of $Re(h)$ and the imaginary part of $Im(h)$, which are independently and identically distributed (i.i.d) Gaussian random variables with a mean of zero and variance of σ^2.

At the receiver, the pilot symbols are recovered first in order to estimate the channel gains. Then, the encoded bits are demodulated with the aid of the estimated channel coefficients before being passed to the RS decoder. The decoder first evaluates the syndromes and checks for errors. If errors are detected, the error correction process is activated. Provided that all errors were successfully corrected, a positive acknowledgement (ACK) is returned to the transmitter, requesting a new packet. Otherwise, a negative ACK is sent back to request for retransmissions.

To recover the channel coefficients, the N_P pilot symbols are fed into the channel estimator, such as a Wiener filter, for generating the estimated version $\hat{h}$ of h. According to [7], the minimum mean square error (MMSE) estimate of the Channel Impulse Response (CIR) $\hat{h}$ may be obtained as

$$\hat{h} = \sum_{k=1}^{N_P} w[k] r_P[k], \qquad (2)$$

where $w[k] = \sigma^2 \sqrt{E_P} (N_P \sigma^2 E_P + N_0)^{-1}$ is the kth filter coefficient, which remains constant for all values of k due to the block-fading assumption. The resultant CIR estimation MSE is [7]

$$E[|\hat{h} - h|^2] = V^2 = \frac{\sigma^2}{1 + N_P \frac{\sigma^2 E_P}{N_0}}. \qquad (3)$$

The estimate $\hat{h}$ is a complex Gaussian random variable with zero mean and a variance of $2(\sigma^2 - V^2)$.

III. IMPACT OF IMPERFECT CSI ON RS CODING AIDED H-ARQ SYSTEMS

A. APER and Goodput

The performance of an ARQ system is typically evaluated in terms of two basic parameters, namely its reliability and throughput. With the assistance of FEC schemes, the reliability of a system may be quantified in terms of the APER, which is defined in [8] as

$$P_E = \frac{P_{ue}}{1 - P_{de}}, \qquad (4)$$

where P_{ue} is the probability of an undetected packet error and P_{de} is the probability of a detected packet error (or probability of retransmission).

The throughput η is expressed as [8]

$$\eta = R_e(1 - P_{de}) = \frac{km}{nm + N_P}(1 - P_{de}), \qquad (5)$$

where $R_e = \frac{km}{nm + N_P}$ denotes the effective rate of each packet, since the code-rate of k/n is further reduced by the pilots. Additionally, the probability of retransmission is obtained by subtracting the probability P_{ue} of an undetected packet error from the probability P_e that a received packet contains at least one symbol error. Hence, we have

$$P_{de} = P_e - P_{ue}. \qquad (6)$$

To evaluate the impact of both the APER and the achievable throughput, the so-called goodput is used, which is defined as the ratio between the expected number of information bits that were correctly received and the number of bits transmitted in a given period of time. In other words, it reflects the ratio of correctly received packets as a fraction of the total throughput, since the latter may contain undetected packet errors. Hence, the goodput η_g may be expressed as

$$\eta_g = (1 - P_E)\eta. \qquad (7)$$

B. Analysis of CSI Error on RS coding aided H-ARQ systems

Owing to its direct impact on all of the above-mentioned probabilities, the Bit Error Probability (BEP) p will be evaluated first. According to [9], if the optimal detection satisfying the minimum symbol error probability criterion is employed at the receiver the conditional BEP evaluated with the aid of imperfect estimates of the channel gain $\hat{h}$ is given by

$$p = \frac{1}{2} erfc\left(\sqrt{\frac{E_S |\hat{h}|^2 \cdot cos^2\alpha}{E_S V^2 + N_0}}\right), \qquad (8)$$

where $erfc$ is the complementary error function [10]; E_S is the mean received energy per symbol per diversity channel, $\hat{h}$ is the channel coefficient estimate and α is some angle [9], which is zero for BPSK modulation [11].

We assume that the block-fading channel estimate $\hat{h}$ follows the Rayleigh distribution, which has the Probability Density Function (PDF) of [6]

$$f(|\hat{h}|) = \frac{|\hat{h}|}{\sigma^2 - V^2} e^{-\frac{|\hat{h}|^2}{2(\sigma^2 - V^2)}}. \qquad (9)$$

Averaging the BEP expression of Eq. (8) over the entire range of Rayleigh faded channel estimates, while taking into account their specific probability of occurance described by the PDF of (9), the average BEP may be expressed as

$$\bar{p} = \int_0^\infty p \cdot f(|\hat{h}|) \cdot d|\hat{h}| \qquad (10)$$

$$= \int_0^\infty \frac{1}{2} erfc\left(\sqrt{\frac{E_S |\hat{h}|^2}{E_S V^2 + N_0}}\right) \frac{|\hat{h}|}{\sigma^2 - V^2} \cdot e^{-\frac{|\hat{h}|^2}{2(\sigma^2 - V^2)}} d|\hat{h}|.$$

Using the Chernoff bound of Appendix I, the average BEP $\bar{p}$ is tightly upper-bounded by

$$\bar{p} \le \frac{1}{2} \cdot \frac{b}{b + c}, \qquad (11)$$

where we have

$$b = \frac{1}{2(\sigma^2 - V^2)}, \qquad c = \frac{E_S}{E_S V^2 + N_0}. \qquad (12)$$

A RS-coded symbol becomes erroneous, when one or more of its m bits is incorrectly received. Thus, the probability of an erroneous RS-coded symbol is obtained as

$$p_s = 1 - (1 - \bar{p})^m. \qquad (13)$$

Next, we will evaluate the probability of an undetectable error P_{ue}. An (n, k) RS decoder, designed to correct $t = \lfloor \frac{n-k}{2} \rfloor$ symbol errors, makes an incorrect codeword decision, when there are more than t symbol errors in a received packet. According to [6], the probability of an undetectable RS-codeword error is

$$
\begin{aligned}
P_{ue} &= \sum_{h=d}^{n} \left[\binom{n}{h} (2^m - 1) \sum_{j=0}^{h-d} (-1)^j \binom{h-1}{j} (2^m)^{h-d-j} \right] \\
&\cdot \sum_{s=0}^{\frac{d-1}{2}} \sum_{g=h-s}^{h+s} \sum_{z=z_{min}}^{z=z_{max}} \binom{h}{h-s+z} \binom{s-z}{g-h+s-2z} \\
&\cdot \binom{n-h}{z} (2^m - 2)^{g-h+s-2z} (2^m - 1)^z \\
&\cdot \frac{1}{(2^m - 1)^g} [1 - (1 - \bar{p})^m]^g [(1 - \bar{p})^m]^{n-g}, \qquad (14)
\end{aligned}
$$

where we have $d = n - m + 1$, $z_{min} = max\{0, g - h\}$ and $z_{max} = \lfloor \frac{g-h+s}{2} \rfloor$.

Furthermore, according to [12], the probability of retransmission P_{de} for a given received packet is

$$
\begin{aligned}
P_{de} &= 1 - P_{ue} - \sum_{h=0}^{\frac{n-k}{2}} \binom{n}{h} p_s^h (1 - p_s)^{n-h} \qquad (15) \\
&= 1 - P_{ue} - \sum_{h=0}^{\frac{n-k}{2}} \binom{n}{h} [1 - (1 - \bar{p})^m]^h [(1 - \bar{p})^m]^{n-h}.
\end{aligned}
$$

Substituting Eqs. (13) and (14) into Eq. (15) yields an explicit formula for the goodput as follows

$$
\begin{aligned}
\eta_g &= \frac{km}{nm + N_P} \left(1 - \frac{P_{ue}}{1 - P_{de}} \right) (1 - P_{de}) \qquad (16) \\
&= \frac{km}{nm + N_P} \sum_{h=0}^{\frac{n-k}{2}} \binom{n}{h} [1 - (1 - \bar{p})^m]^h [(1 - \bar{p})^m]^{n-h}.
\end{aligned}
$$

It may be readily shown that the probability P_{de} of retransmission in Eq. (15) is a monotonically decreasing function of the average BEP of $\bar{p}$, which will be further elaborated on below. When taking into account the upper bound of the average BEP in Eq. (11), the goodput may be shown to be lower bounded by

$$
\begin{aligned}
\eta_g &\geq \frac{km}{nm + N_P} \sum_{h=0}^{\frac{n-k}{2}} \binom{n}{h} \qquad (17) \\
&\cdot \left[1 - \left(1 - \frac{b}{2(b+c)} \right)^m \right]^h \left[\left(1 - \frac{b}{2(b+c)} \right)^m \right]^{n-h}.
\end{aligned}
$$

C. Power Allocation

Let us assume that the amount of total transmit energy E_T is constant, regardless of the number of pilot symbols transmitted as well as of the number of transmit and receive antennas employed. This assumption implies that the higher the pilot symbols' energy, the lower the data symbols' energy and vice versa. In order to characterize this relationship, let us define a power-allocation factor ε, which quantifies the percentage of pilot energy in the total transmit energy. Given this assumption, the amount of energy allocated to pilot symbols becomes $\varepsilon E_T = N_P E_P$, while the remaining energy of $(1-\varepsilon)E_T = nE_S$ is assigned to the data and parity symbols.

Naturally, assigning a large fraction of the total energy to the pilot symbols, which is equal to the product of N_P and E_P, is expected to reduce the MSE of channel estimation. However, this automatically reduces the energy of the useful payload data, hence reducing the average SNR. Consequently, the BEP will increase. Therefore, a specific power allocation scheme is required here, which optimizes the attainable system performance.

As mentioned above, the goodput of the RS/H-ARQ system is a monotonically decreasing function of the BEP $\bar{p}$. However, it may be observed from Eq. (11) that the BEP $\bar{p}$ is reduced, when the ratio c/b increases. This implies that the goodput is a monotonically increasing function of the variable c/b. In order to maximize the goodput of the RS/H-ARQ system, it is necessary to find the maximum value of c/b in Eq. (12). This problem may be simplified to finding the best power-allocation factor of ε by solving the optimization problem of

$$\varepsilon_{opt} = \arg \max_{0 < \varepsilon < 1} \{\eta_g\} = \arg \max_{0 < \varepsilon < 1} \left\{ \frac{c}{b} \right\}. \qquad (18)$$

This problem was solved in [5] for an uncoded scenario. Using the results of [5], we have

$$\varepsilon_{opt} = \frac{n + \gamma - \sqrt{n^2 + n\gamma + n^2\gamma + n\gamma^2}}{-(n\gamma - \gamma)}, \qquad (19)$$

where $\gamma = E_T \sigma^2 / N_0$ represents the ratio of the total transmit power to the noise experienced at the receiver.

Also following from [5], the value of ε_{opt} is limited as

$$\frac{1 - \sqrt{n}}{1 - n} \leq \varepsilon_{opt} \leq 0.5. \qquad (20)$$

IV. NUMERICAL RESULTS

As mentioned in Section III, the goodput reflects the combined impact of both the APER and the throughput. Thus, in this section, we will focus our attention on the goodput results to evaluate the performance of the proposed RS/H-ARQ system.

Fig. 1 shows the dependence of the goodput on the normalized CSI MSE and the energy per bit to noise power spectral density ratio (E_b/N_0) for three different code rates, namely for $R = \{0.92, 0.88, 0.84\}$, while fixing the RS codeword length to 255 symbols ($n = 255$), each of which has 8 bits per symbol ($m = 8$). Fig. 1 indicates that the goodput degrades rapidly, when the normalized CSI MSE is above the critical value of 10^{-2} for our simulation. This CSI MSE value is achieved at a

received E_b/N_0 value of around $10dB$, when an equal pilot-data power per symbol is allocated.

In Fig. 2, the goodput of a range of code rates is quantified with the aid of the thin lines for the range of $R = \{0.22, ..., 0.98\}$ in steps of 0.02. The maximum achievable goodput at a specific E_b/N_0 value and a given value of R is presented by the bold curve marked with filled circles. A more detailed picture emerges by observing the three-dimensional graph of Fig. 3. Clearly, the attainable goodput depends on both the code rate and the E_b/N_0. More explicitly, when considering the maximum goodput curve marked by the bold line. Taking a point on the curve and mapping it to the x and y axes, we could identify the best code rate, which provides the maximum goodput at a given value of E_b/N_0. It is interesting to note that the goodput does not improve further, when the code rate is reduced below 0.45. Therefore, the code-rate range of $0.45 \div 1$ may be deemed appropriate for the RS/H-ARQ system associated with $m = 8$ and $n = 255$.

Next, we will investigate the effect of the power allocation regime suggested in Section III-C. It can be observed in Fig. 4 that the goodput monotonically decreases with the BEP, which has confirmed our predictions outlined in the previous section. The optimal power-allocation fraction ε_{opt} is shown in Fig. 5. Three different RS codes were investigated, namely the $\{255/223; 97/85; 32/28\}$ codes, which have the same code rate of 0.875 but different codeword lengths. As expected, a larger fraction of the power is assigned to the pilot symbols in the low E_b/N_0 region. In addition, when increasing the codeword length, the system has to assign a larger amount of transmit energy, but relies on the same number of pilot symbols. Hence, this leads to the relative reduction of the power assigned to pilot symbols. This may be observed in Fig. 5.

Based on using the value of ε_{opt} characterized in Fig. 5, the overall goodput performance of RS/H-ARQ systems is examined in Fig. 6. Clearly, the optimized pilot-data power allocation scheme assists the system in achieving an improved goodput. It is also interesting to note that the longer the RS codeword, the higher the goodput. When the proposed optimum power allocation is employed, there is an approximately $1.75dB$ improvement for the $(32/28)$ RS code, while this value is about $2.5dB$ for the $(255/223)$ RS code. Additionally, the longer codewords achieve the maximum attainable goodput at lower E_b/N_0 values than the short ones.

V. CONCLUSION

In this paper, we have studied the effects of imperfect CSI on the achievable goodput of the RS/H-ARQ systems. The results of Fig. 6 demonstrated that the CSI estimation plays an important role in recovering the received signal. Furthermore, the power allocation between pilot and data symbols has been optimized for achieving the maximum attainable goodput. More particularly, a power-allocation factor ε_{opt}, which ranges from $\frac{1-\sqrt{n}}{1-n}$ to half of the total transmit power, has been defined and optimized in order to maximize the system's goodput. It was found that the system achieved a $2.5dB$ E_b/N_0 gain at a given goodput for the $255/223$ RS code, when the proposed power allocation scheme was utilized.

APPENDIX I

By using the Chernoff bound of $erfc(x) < e^{-x^2}$ [5], the upper bound of the average BEP can be expressed as

$$\bar{p} \leq \int_0^\infty \frac{1}{2} e^{-c|\hat{h}|^2} \cdot 2b|\hat{h}| \cdot e^{-b|\hat{h}|^2} d|\hat{h}|. \qquad (21)$$

Substituting $z = |\hat{h}|^2$ into the above inequality, we have

$$\bar{p} \leq \frac{b}{2} \int_0^\infty e^{-(b+c)z} dz = \frac{-b}{2(b+c)} e^{-(b+c)z} \Big|_0^\infty = \frac{b}{2(b+c)}. \qquad (22)$$

REFERENCES

[1] J. M. Wozencraft and M. Horstein, "Coding for two-way channels," tech. rep., Research Laboratory of Electronics, M.I.T, 1961.
[2] S. Lin and P. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels," *IEEE Transactions on Communications*, vol. 30, pp. 1701–1719, Jul 1982.
[3] L. Hanzo, J. S. Blogh, and S. Ni, *3G, HSPA and FDD versus TDD Networking: Smart Antennas and Adaptive Modulation*. Wiley, 2nd ed., Feb. 2008.
[4] IEEE, www.ieee.org, *IEEE Standard for Local and metropolitan area networks. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems*, ieee std 802.16e ed., 2005.
[5] L. Cao, P.-Y. Kam, and M. Tao, "Impact of imperfect channel state information on ARQ schemes over Rayleigh fading channels," in *Proceedings of IEEE International Conference on Communications (ICC) '09*, pp. 1–5, June 14–18, 2009.
[6] R. Steele and L. Hanzo, *Mobile Radio Communications: Second and Third-generation Cellular and WATM Systems*. John Wiley - IEEE Press, 2nd ed., May 1999.
[7] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 4th ed., 2002.
[8] S. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, Oct 1994.
[9] P. Y. Kam, "Optimal detection of digital data over the nonselective Rayleigh fading channel with diversity reception," *IEEE Transactions on Communications*, vol. 39, pp. 214–219, Feb. 1991.
[10] J. G. Proakis, *Digital communications*. McGraw-Hill, 4th ed., 2001.
[11] Y. Zhu, P.-Y. Kam, and Y. Xin, "A decision-feedback channel estimation receiver for independent nonidentical Rayleigh fading channels," in *Proceedings of IEEE Vehicular Technology Conference (VTC) Spring 2008*, pp. 376–379, May 11–14, 2008.
[12] S. B. Wicker, "Reed-solomon error control coding for Rayleigh fading channels with feedback," *IEEE Transactions on Vehicular Technology*, vol. 41, pp. 124–133, May 1992.
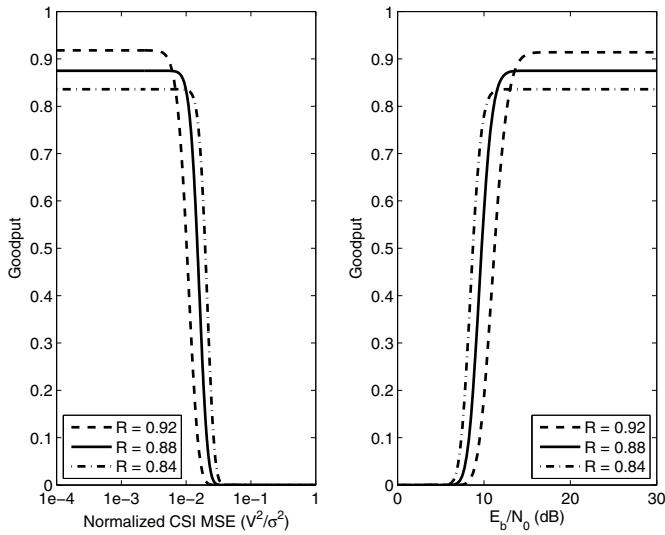
Fig. 1. Lower bound of the goodput for the RS/H-ARQ transmission over a block-fading, independent Rayleigh channel: $m = 8$, $n = 255$, $N_P = 8$
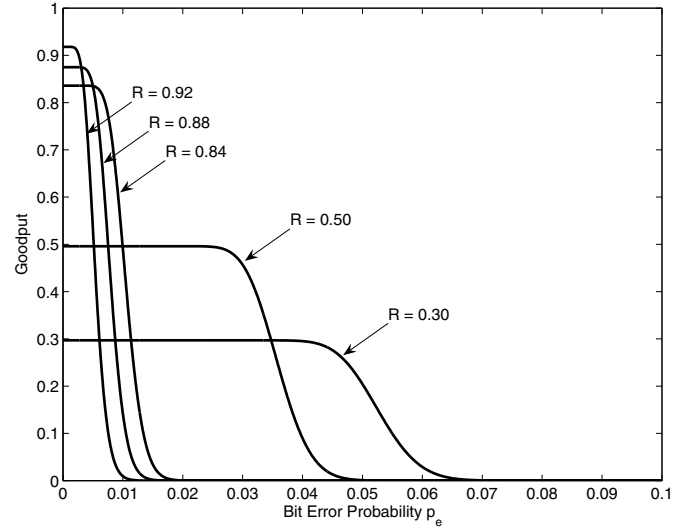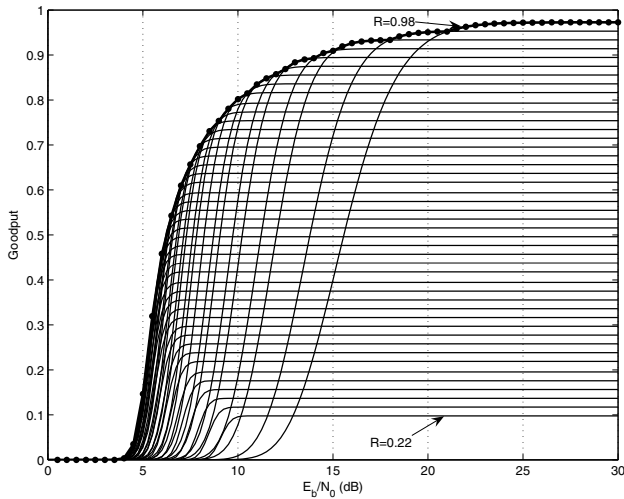

Fig. 2. Goodput at different RS code rates for transmission over a block-fading, independent Rayleigh channel: $m = 8$, $n = 255$, $N_P = 8$, the code rate was varied between $R = 0.22$ and $R = 0.98$ in steps of 0.02
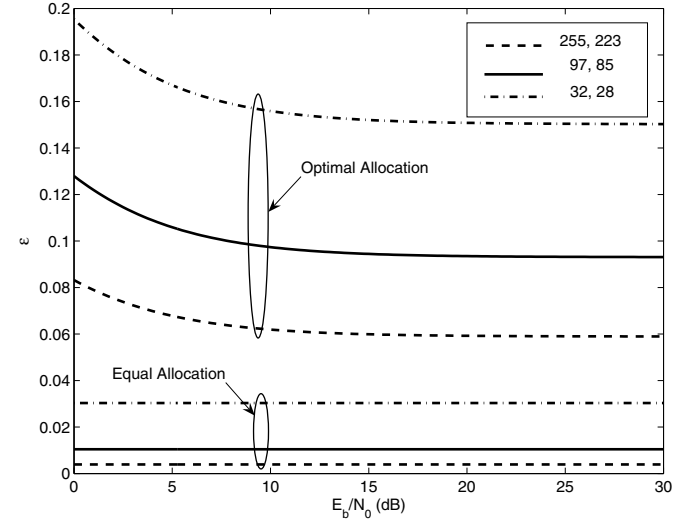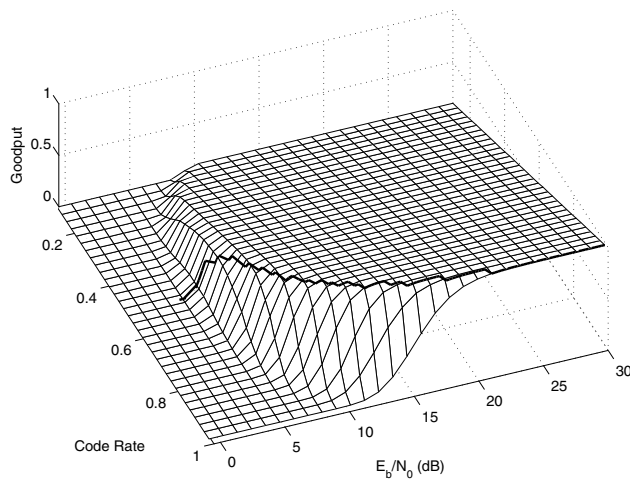

Fig. 3. Goodput at different RS code rates for transmission over a block-fading, independent Rayleigh channel: $m = 8$, $n = 255$, $N_P = 8$, the code rate was varied between $R = 0.22$ and $R = 0.98$ in steps of 0.02

Fig. 4. The dependence of goodput on the bit error probability for transmission over a block-fading, independent Rayleigh channel: $m = 8$, $n = 255$, $N_P = 8$

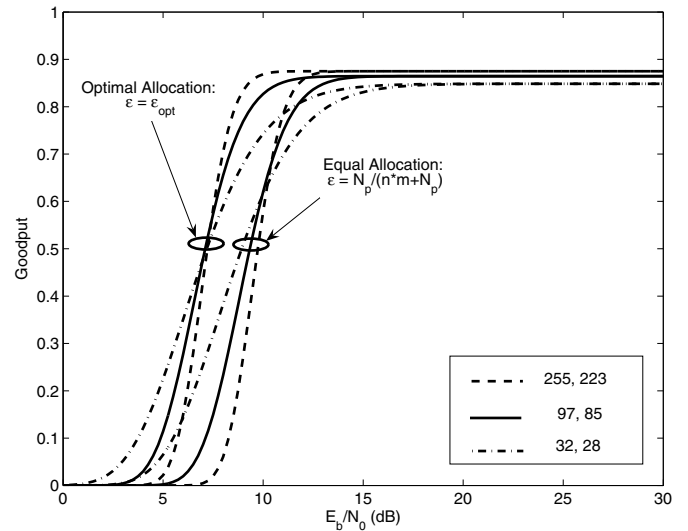Fig. 5. Power Allocation: $m = 8$, $N_P = 8$

Fig. 6. Lower bound of goodput based on both equal and optimized power allocation schemes for transmission over a block-fading, independent Rayleigh channel: $m = 8$, $N_P = 8$