Efficient Hybrid ARQ with Space-Time Coding and Low-Complexity Decoding

Mi-Kyung Oh, Young-Hyeon Kwon and Dong-Jo Park
Dept. of EECS, KAIST; 373-1 Guseong-dong, Yuseong-gu, Daejon 305-701, Republic of Korea

***Abstract*— We aim at increasing the throughput of the hybrid automatic retransmission request (HARQ) protocol in Space-Time (ST) coded multi-antenna transmission systems. By utilizing reliability information at the decoder, we obtain an improved probability of successful decoding, which enhances the overall system throughput at low-complexity. Simulations and analytical results demonstrate the performance of our scheme in AWGN and fading multi-input multi-output (MIMO) channels.**

I. Introduction

Demand in wireless connectivity has led to major advances providing reliable high data rates over wireless channels. Diversity is typically employed to mitigate channel fading effects in time, frequency, or space [1], [5]–[7]. Recently, full-diversity and full-rate (FDFR) ST codes with any number of transmit- and receive-antennas in MIMO fading channels have been proposed to achieve high performance and high data rates [5]. Although ST coding has well appreciated merits in coping with channel effects, HARQ is typically employed to improve the low throughput of ARQ-only protocols [6].

On the other hand, algebraic hard decision decoding (HDD) algorithms are popular because they lead to low complexity at the receiver. While optimal (unquantized) soft decision decoding (SDD) algorithms provide up to 3dB coding gain relative to HDD [7], HDD offers the distinct advantages of low decoding complexity and robustness to channel-induced interference such as jamming, impulsive noise, and fading; see e.g., [3] and references therein. In addition, it has been shown that by resorting to reliability information, the error resilience of HDD can be increased while retaining low complexity [8].

Many HARQ schemes with various forward error correction (FEC) algorithms have been pursued to increase the system throughput [4]. Recently, HARQ schemes combined with ST codes have been proposed [6]. Other researches have shown that ST codes can be effective in HARQ protocols [9].

In this paper, we wed the merits of HARQ, ST coding, and HDD to increase the system throughput while maintaining low-complexity. First, we combine all received copies of the same packet by using maximum ratio combining (MRC). Combining packets has the effect of increasing the number of receiver antennas with every retransmission. Therefore, with HARQ and ST coding, transmit and receive diversity is established by increasing the diversity order per retransmission. Moreover, our development of an efficient HDD algorithm provides improved error performance at low-complexity. This combination allows the receiver to perform reliable decoding and hence reduce the probability for retransmission. In this paper, we assume that acknowledgment (ACK) or negative acknowledgment (NAK) is fed back to the transmitter and the channel state information (CSI) is not available at the transmitter. However, when the channel feedback is allowed, we may further increase the system throughput in time-varying channel conditions by using adaptive modulation and coding (AMC) [4].

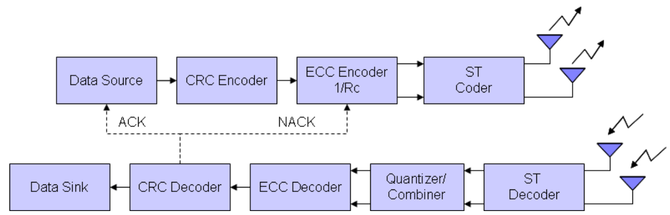

Fig. 1. The system model

II. Preliminaries

Consider HARQ system equipped with multiple antennas, as shown in Fig. 1. The cyclic redundancy check (CRC) coded packet $\boldsymbol{b} := [b_1, \ldots, b_K]$ of length K is encoded using a rate $1/R_c$ convolutional code (CC) to result in R_c encoded packets $\boldsymbol{c}^{(j)} := [c_1^{(j)}, \ldots, c_K^{(j)}]$, where $j = 1, \ldots, R_c$, $\boldsymbol{c} := \{\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(R_c)}\}$ belongs to a codeword set $\boldsymbol{C}$ with code distance d, and $c_k^{(j)}$ is an element of GF(2). For simplicity, we will use R_c transmit antennas equal to the number of encoded packets (i.e., code rate). Although we use convolutional code in this paper, we do not restrict ourselves to any particular coding system. Subsequently, encoded data packets are modulated using BPSK and then encoded using ST coding such as Alamouti [1] and FDFR codes [5]. ST coded packets are forwarded to each antenna and then transmitted over an AWGN channel with variance $\sigma^2 := N_o/2$.

At the receiver, we can obtain the ST decoded sequence $\tilde{\boldsymbol{c}}^{(j)}(n) := [\tilde{c}_1^{(j)}(n), \cdots, \tilde{c}_K^{(j)}(n)]$ for the nth transmission of the same packet $\boldsymbol{c}^{(j)}$, where $\tilde{c}_k^{(j)}(n) \sim \mathcal{N}(\pm\sqrt{\mathcal{E}}, N_o/(2n))$, and $\mathcal{E}$ denotes the received energy per symbol. We will explain the reduced noise variance of $N_o/(2n)$ in Section III.

Assume that N copies, $\{\tilde{\boldsymbol{c}}^{(j)}(n)\}_{n=1}^{N}$ have been received and CRC checked as unreliable. These packets are combined into a single packet by taking the average of the received values of each re-transmitted packet as follows:

$$r_k^{(j)}(N) = \frac{1}{N}\sum_{n=1}^{N} \tilde{c}_k^{(j)}(n), \tag{1}$$

where $\boldsymbol{r}^{(j)}(N) := [r_1^{(j)}(N), \ldots, r_K^{(j)}(N)]$ for $j = 1, \ldots, R_c$.

Next, the quantizer compares $r_k^{(j)}(N)$ with a decision threshold (zero when a priori probabilities are equal, i.e., $P(0) = P(1) = 1/2$) to output the hard decision sequence $\boldsymbol{z}^{(j)}(N) := [z_1^{(j)}(N), \ldots, z_K^{(j)}(N)]$ with $z_k^{(j)}(N) \in \{0, 1\}$. Letting $\boldsymbol{z} := \{\boldsymbol{z}^{(1)}(N), \ldots, \boldsymbol{z}^{(R_c)}(N)\}$, the decoder metric for $\boldsymbol{c} \in \boldsymbol{C}$, denoted by $\text{CM}_{\boldsymbol{z}}(\boldsymbol{c})$, is the Hamming distance defined as:

$$\text{CM}_{\boldsymbol{z}}(\boldsymbol{c}) = \sum_{j=1}^{R_c} \sum_{k=1}^{K} |c_k^{(j)} - z_k^{(j)}(N)|. \tag{2}$$

Then, a codeword $\boldsymbol{c}$ is more likely to be the transmitted codeword than another codeword $\boldsymbol{c}^{'}$ if and only if $\text{CM}_{\boldsymbol{z}}(\boldsymbol{c}) \leq \text{CM}_{\boldsymbol{z}}(\boldsymbol{c}^{'})$; i.e., maximum likelihood (ML) decoding always decodes $\boldsymbol{z}$ to the codeword with the smallest Hamming distance.

However, $r_k^{(j)}(N)$ close to the decision threshold are likely unreliable, which motivates us to puncture those bits belonging to the interval $-\gamma_p < r_k^{(j)}(N) < \gamma_p$, where γ_p is a suitably chosen puncturing threshold. The reliability information will be denoted as $\boldsymbol{\alpha}^{(j)} := [\alpha_1^{(j)}, \ldots, \alpha_K^{(j)}]$, with $\alpha_k^{(j)} \in \{0, 1\}$, where $\alpha_k^{(j)} = 0$ indicates that the kth bit in $\boldsymbol{z}^{(j)}(N)$ is unreliable. We call this algorithm "receive-puncturing" to differentiate it from the known transmit-puncturing used to increase FEC rates.

Packets $\{\boldsymbol{z}^{(j)}(N)\}_{j=1}^{R_c}$ with corresponding reliability information $\boldsymbol{\alpha} := [\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(R_c)}]$ are processed jointly by the Viterbi decoder to decode the original packet based on the metric

$$\text{CM}_{\boldsymbol{z}}(\boldsymbol{c}, \boldsymbol{\alpha}) = \sum_{j=1}^{R_c} \sum_{k=1}^{K} \alpha_k^{(j)} |c_k^{(j)} - z_k^{(j)}(N)|, \tag{3}$$

from which, it is shown that unreliable bits marked by $\alpha_k^{(j)} = 0$ do not contribute to the calculation of the metrics in the decoder. Thus we observe from (3) that relative to conventional HDD, the hardware complexity required by our approach entails just AND gates for implementing $\boldsymbol{\alpha}$.

If the combined packet turns out to be decoded unreliably by the CRC decoder, a retransmission is requested and the transmitter sends another copy $\{\boldsymbol{c}^{(j)}\}_{j=1}^{R_c}$. The receiver continues to request and combine packets until successful decoding.

Our objective in this paper is to show that the receive-puncturing scheme employing reliability information can improve the throughput of HARQ for an ST coded system.

III. HARQ WITH RELIABILITY

In this section, we optimize the decoder so that the system throughput of HARQ can be increased. In order to attain this goal, we rely on the algorithm in [3], where the optimal puncturing threshold γ_p^* enables the codeword decision error probability (CEP) to be minimized at the output of the Viterbi decoder. For simplicity, we temporally omit the transmission index n in this section.

A. Optimal Puncturing Threshold

To derive γ_p^* that determines the entries of the reliability vector $\boldsymbol{\alpha}$, we define the *punctured code distance* as:

$$d(\boldsymbol{\alpha}) = \sum_{j=1}^{R_c} \sum_{k=1}^{K} \alpha_k^{(j)} \cdot |c_k^{(j)} - c_k^{(j)'}|, \tag{4}$$

where $\boldsymbol{c}^{(j)} \neq \boldsymbol{c}^{(j)'}$. We note that $d(\boldsymbol{\alpha})$ is a random variable corresponding to the code distance after puncturing and takes values between 0 and d (original code distance). Although the puncturing scheme decreases the code distance, the error correction capability can increase due to the enhanced signal-to-noise ratio (SNR) [3]. Specifically, the probability of having $d(\boldsymbol{\alpha}) = m$ is given by

$$Pr(d(\boldsymbol{\alpha}){=}m){=}\binom{d}{m} P_{\text{p}}^{d-m}(\gamma_p)\{P_{\text{cp}}(\gamma_p) + P_{\text{ep}}(\gamma_p)\}^m, \tag{5}$$

where $P_{\text{cp}}(\gamma_p)$, $P_{\text{ep}}(\gamma_p)$, and $P_{\text{p}}(\gamma_p)$ denote the probabilities of correct symbol, error, and puncturing, which are defined as:

$$P_{\text{q}}(\gamma_p) = \sum_{i=0}^{1} P(i) \int_{S_{\text{q}}^i} f(r|i) dr, \tag{6}$$

where $f(r|i)$ is the probability density function (PDF) of the observation r conditioned on signal class $i \in \{0, 1\}$, S_{q}^i denotes the region that yields $P_{\text{q}}(\gamma_p) \in \{P_{\text{cp}}(\gamma_p), P_{\text{ep}}(\gamma_p), P_{\text{p}}(\gamma_p)\}$, assuming that signal class $i \in \{0, 1\}$ is transmitted. These probabilities obey the relationship $P_{\text{ep}}(\gamma_p) + P_{\text{cp}}(\gamma_p) + P_{\text{p}}(\gamma_p) = 1$.

To derive the averaged CEP when receive-puncturing scheme is employed, we assume that the all-zero codeward (AZC) is transmitted and the codeword being compared with the AZC has distance d from the AZC. Then, the CEP of conventional HDD becomes

$$P_{\text{CEP}}(d) = \sum_{m=\lfloor (d+1)/2 \rfloor}^{d} \binom{d}{m} P_{\text{e}}^m P_{\text{c}}^{d-m}, \tag{7}$$

where $P_{\text{e}} = P_{\text{ep}}(0)$, and $P_{\text{c}} = 1 - P_{\text{e}}$ [7]. If a puncturing process is employed, for code distance d, the CEP averaged with respect to $d(\boldsymbol{\alpha})$ can be alternatively expressed as [3]

$$\bar{P}_{\text{CEP}}(d) = \sum_{m=0}^{d} \binom{d}{m} P_{\text{p}}^{d-m}(\gamma_p) \sum_{k=\lfloor \frac{m+1}{2} \rfloor}^{m} P_{\text{ep}}^k(\gamma_p) P_{\text{cp}}^{m-k}(\gamma_p). \tag{8}$$

We note that if $\gamma_p = 0$, i.e., $P_{\text{p}}(\gamma_p) = 0$ in (8), then the average CEP (8) becomes the CEP of conventional HDD. Since it is difficult to find the γ_p^* minimizing $\bar{P}_{\text{CEP}}(d)$ in closed form, we use Chernoff's approximation [7], based on which, we obtain the following upper bound on CEP:

$$\begin{aligned} \bar{P}_{\text{CEP}}(d) &\leq \sum_{m=0}^{d} \binom{d}{m} P_{\text{p}}^{d-m}(\gamma_p) \left\{4 P_{\text{ep}}(\gamma_p) P_{\text{cp}}(\gamma_p)\right\}^{\frac{m}{2}} \\ &= \left(P_{\text{p}}(\gamma_p) + \sqrt{4 P_{\text{ep}}(\gamma_p) P_{\text{cp}}(\gamma_p)}\right)^d. \end{aligned} \tag{9}$$

By differentiating the bounded $\bar{P}_{\text{CEP}}(d)$ in (9) with respect to γ_p, and setting it equal to zero, the optimal threshold γ_p^* is found to obey

$$\frac{P_{\text{ep}}(\gamma_p^*)}{P_{\text{cp}}(\gamma_p^*)} = \left\{ \frac{P'_{\text{ep}}(\gamma_p^*)}{P'_{\text{cp}}(\gamma_p^*)} \right\}^2, \tag{10}$$

where prime denotes differentiation. Choosing the threshold to satisfy (10) is shown in [3] to optimize the decoder performance. The solution of (10) is unique and valid for the exact CEP in (8) [3].

B. Throughput Analysis

To derive the throughput of HARQ, we assume an ideal retransmission protocol with no delay between packet retransmissions, and the error/delay-free feedback channel. The throughput Γ is then a function of the average number of transmissions N_r by the HARQ protocol.

Let $P(D_u(n))$, $P(D_d(n))$, and $P(D_c(n))$ be the probabilities of a combined received sequence after n transmissions that contain undetected errors, detected errors, and no errors, respectively. Note that $P(D_d(n))$ is equivalent to the event of a retransmission request. These obey the relationship:

$$P(D_u(n)) + P(D_d(n)) + P(D_c(n)) = 1. \tag{11}$$

Assuming that $P(D_u(n))$ is negligible for most CRC codes, N_r can be expressed as

$$\begin{aligned} N_r =& 1 + P(D_d(1)) + P(D_d(1), D_d(2)) + \\ & \cdots + P(D_d(1), D_d(2), \ldots, D_d(n)) + \cdots . \end{aligned} \tag{12}$$

The joint probability can be lower and upper bounded by [2]

$$1 + \sum_{n=1}^{\infty} \prod_{i=1}^{n} P(D_d(i)) \leq N_r \leq 1 + \sum_{n=1}^{\infty} P(D_d(n)). \tag{13}$$

$P(D_d(n))$ can also be upper bounded as:

$$P(D_d(n)) \simeq 1 - P(D_c(n)) \leq 1 - (1 - P(E(n)))^K, \tag{14}$$

where $P(E(n))$ is the probability of a decoding error event during Viterbi decoding on the nth transmission. Thus, the upper bound of N_r in (13) is given by

$$N_r \leq 1 + \sum_{n=1}^{\infty} \left\{ 1 - (1 - P(E(n)))^K \right\}. \tag{15}$$

Given the free distance d_{free} of the convolutional code, $P(E)$ $(= P(E(1)))$ can also be bounded by [7]

$$P(E) \leq \sum_{d=d_{\text{free}}}^{\infty} a_d P_{\text{CEP}}(d), \tag{16}$$

where the distance spectra a_d represents the number of paths with distance d. Notice that the upper bounded CEP of the conventional HDD is given by $P_{\text{CEP}}(d) \leq [4P_{\text{e}} \cdot P_{\text{c}}]^{d/2}$, and the average CEP of our algorithm with optimal puncturing threshold γ_p^* is

$$\bar{P}_{\text{CEP}}(d) \leq \left(P_{\text{p}}(\gamma_p^*) + \sqrt{4P_{\text{ep}}(\gamma_p^*) \cdot P_{\text{cp}}(\gamma_p^*)} \right)^d. \tag{17}$$

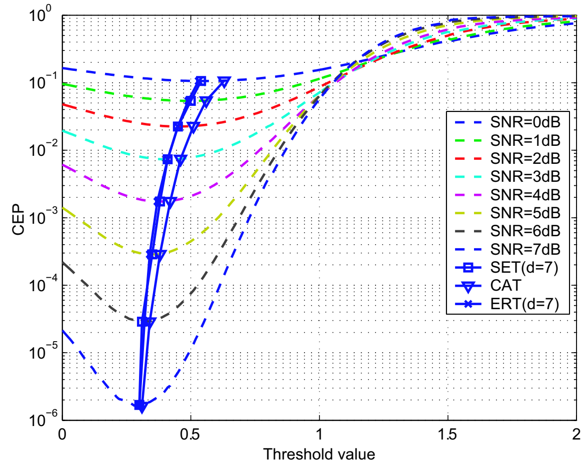

Fig. 2. Puncturing threshold vs. CEP

Let us now describe how $P(E(n))$ is determined. To this end, we will use the variance of the received symbol $c_k^{(j)}(n)$ and the effective SNR after n packets are combined:

$$\begin{aligned} \text{var}[c_k^{(j)}(n)] &= \frac{1}{n} \cdot \text{var}[c_k^{(j)}(1)], \\ \text{SNR}(n) &= n \cdot \text{SNR}, \end{aligned} \tag{18}$$

where SNR $:= \mathcal{E}/\sigma_\omega^2$ represents SNR without combining. Consequently, the MRC of n packets increases the effective received SNR by a factor n.

Let $P_{\text{CEP}}(d;n)$ and $\bar{P}_{\text{CEP}}(d;n)$ be the CEP of conventional HDD and the CEP of our scheme with n combined packets. $P_{\text{CEP}}(d;n)$ can be calculated by plugging P_{e} with the reduced noise variance by a factor n into $P_{\text{CEP}}(d)$, while $\bar{P}_{\text{CEP}}(d;n)$ can be obtained by choosing the optimal puncturing threshold corresponding to the increased SNR. Recalling that our optimal puncturing thresholds minimize the CEP, we deduce that

$$\bar{P}_{\text{CEP}}(d;n) \leq P_{\text{CEP}}(d;n), \tag{19}$$

which leads to a lower $P(E(n))$ for our algorithm.

Based on this derivation of the average number of transmissions, we can evaluate a lower bound of our throughput performance. Because we utilized a rate 1 ST code (e.g., Alamouti code [1]), the throughput can be expressed as $\Gamma := 1/N_r$. Therefore, our assertion that receive-puncturing enhances system throughput has been confirmed.

IV. Simulation Results

To verify performance, we conduct simulations for both AWGN and Rayleigh fading channels. In all experiments, we use Alamouti ST code with 2 transmit and 1 receive antenna and encode each data block ($K = 20000$) using a rate $R_c = 1/2$ convolutional code (133,171) with a constraint length 7, and minimum code distance 10.

In Fig. 2, we illustrate the puncturing threshold versus CEP. It is observed that there exist optimal thresholds that minimize the CEP for each SNR. Noting that $P_{\text{CEP}}(d)$ of conventional HDD corresponds to $\bar{P}_{\text{CEP}}(0)$, we deduce the improvement

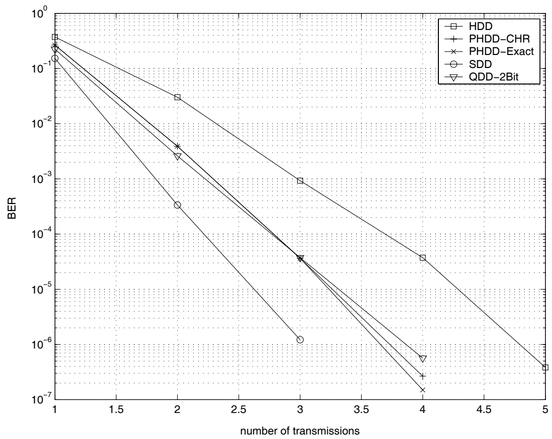

Fig. 3. BER vs. the number of transmissions in AWGN (SNR=0dB)

of our scheme. We also plot the minimum points found by simulation (referred to as SET) and the points corresponding to the puncturing thresholds of (8) and (10) (referred to as ERT and CAT), for which the CEP difference between them is quite small.

BER performance versus the average number of transmissions for SNR=0dB is shown in Fig. 3. Because the number of bits in our decoding algorithm (referred to as punctured (P)HDD) is 2-bit (1-bit for HDD + 1-bit reliability information), we compare with a 2-bit quantized SDD. It is observed that our BER is similar to a 2-bit quantized SDD, or even better when the number of transmissions increases. However, one should recall that the complexity of our algorithm is lower than that of a 2-bit quantized SDD. We also note that the gaps between the performance using the optimal thresholds calculated by Chernoff bound of CEP as in (9) and those by the exact CEP as in (8) are relatively small.

Figure 4 depicts the average number of transmissions versus SNR. As expected from Fig. 3, the number of retransmissions is similar to 2-bit quantized SDD. It is seen that at high SNR, where on average only one transmission is needed, the error performances of the 4 schemes tend to coincide.

For Rayleigh fading channel described by Jakes' model with Doppler frequency 100Hz for 11MHz symbol transmission on a 2.4GHz carrier frequency, we assume that the channel information is perfectly known to the receiver and zero forcing (ZF) equalization is used. As shown in Fig. 5, our scheme performs well compared with SDD, conventional HDD, and 2-bit quantized SDD.

V. Conclusions

We have derived an efficient HARQ protocol equipped with ST coding to enable transmit and receive diversity. Relying on reliability information, we relied on receive-puncturing to further increase the overall system throughput at low decoding complexity thanks to HDD. Simulations illustrated the merits of our algorithm for both AWGN and fading channels.

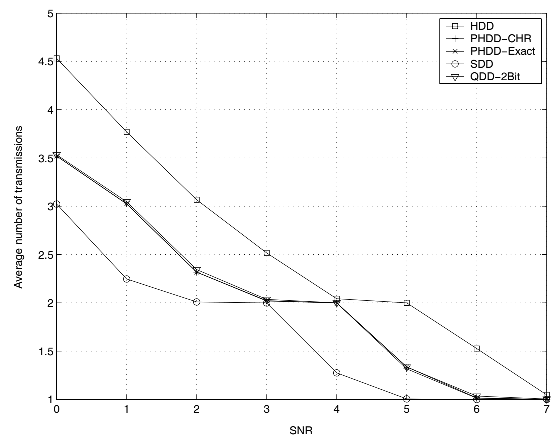

Fig. 4. number of transmissions vs. SNR in AWGN

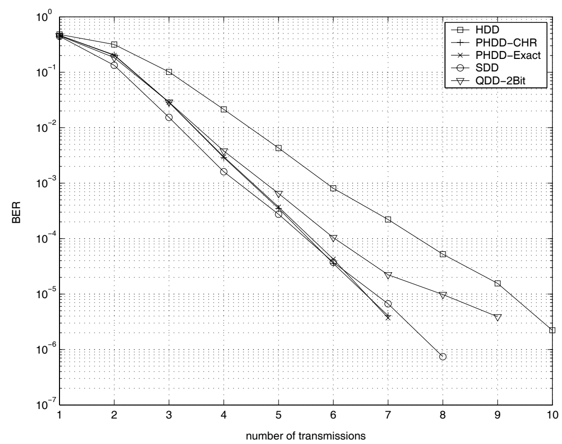

Fig. 5. number of transmissions vs. BER in fading channel (SNR=6dB)

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communication," *IEEE Jounral on Selec. Areas in Commun.* vol. 16, no. 8, pp. 1451-1458, Oct. 1998.

[2] S. Kallel and D. Haccoun, "Generalized type II Hybrid ARQ scheme using punctured convolutional coding," *IEEE Tran. on Commun.,* vol. 38, no. 11, pp. 1938-1946, Nov. 1990.

[3] Y.-H. Kwon, M.-K. Oh, D.-J. Park, and G. B. Giannakis, "Efficient hard decoding algorithm with 1-bit reliability," *IEEE Commun. Letters*, submitted Oct. 2003.

[4] Q. Liu, S. Zhou, and G. B. Giannakis, "Combining adaptive modulation and coding with truncated ARQ enhences throughput," *Proc. of Signal Proc. Workshop on Advances in Wireless Commun.*, pp. 587-591, Rome, Italy, June 2003.

[5] X. Ma, and G. B. Giannakis, "Full-diversity full-rate complex-field space-time coding," *IEEE Tran. on Signal Proc.,* vol. 51, no. 11, pp. 2917-2930, Nov. 2003.

[6] A. V. Nguyen and M. A. Ingram, "Hybrid ARQ protocols using space-time codes," *IEEE Proc. of the 54th Veh. Tech. Conf.,* vol. 4, pp. 2364-2368, 7-11 Oct. 2001.

[7] J. G. Proakis, *Digital communications,* McGraw-Hill, 4th edition, 2000.

[8] L.-L. Yang and L. Hanzo, "Performance analysis of coded M-ary orthogonal signaling using errors-and-erasures decoding over frequency-seletive fading channels," *IEEE Jounral on Selec. Areas in Commun.* vol. 19, no. 2, pp. 211-221, Feb. 2001.

[9] H. Zheng, "The performance of BLAST with Hybrid ARQ in Ricean fading channels," *Proc. of the 54th Veh. Tech. Conf,* vol. 2, pp. 901-904, 7-11 Oct. 2001.