Hybrid ARQ for Non-Orthogonal Space-Time Block Codes

Rui Lin, Philippa A. Martin

Department of Electrical and Computer Engineering, University of Canterbury,
Private Bag 4800, Christchurch 8140, New Zealand
E-mail: {rli24@student, p.martin@elec}.canterbury.ac.nz

Abstract

We propose a novel hybrid automatic repeat-request (HARQ) scheme for non-orthogonal space-time block codes based on BCH and Reed Solomon codes. We use list decoding techniques to significantly reduce the decoding complexity. We obtain significant gains in terms of dropped packet rate and throughput spectral efficiency at the cost of increased complexity.

1. Introduction

Modern wireless communication networks typically transmit information in a packet by packet fashion. A packet is declared dropped if it still contains errors after the maximum allowed number of transmissions, $m_{max}$. Automatic-repeat-request (ARQ) schemes can be used to improve the dropped packet rate (DPR) at the cost of throughput spectral efficiency (TSE). In order to maintain high spectral efficiency, the average number of transmissions for each packet needs to be minimized and the code rate needs to be maximized. This requires the use of powerful high rate codes.

Space-time codes (STCs) are an effective way to combat fading. Most existing hybrid ARQ (HARQ) schemes designed for STCs use an error control code, such as a convolutional [1] or turbo [2] code. The encoded packets are then transmitted using BPSK and/or an orthogonal space-time block code (OSTBC) [3], which limits the achievable spectral efficiency.

The non-OSTBCs (NOSTBCs) of [4] allow much higher rates and obtain significant coding gains compared to OSTBCs at the cost of increased decoding complexity. In this paper, we design a HARQ scheme for these NOSTBCs and a list decoding algorithm to reduce the resulting decoding complexity. This list decoding algorithm can be applied to other HARQ schemes.

2. System Description

We consider a MIMO system with $n_t$ transmit and $n_r$ receive antennas. We assume a quasi-static channel model and so the same $n_r 	imes n_t$ channel matrix, $H$, is used for an entire frame. We assume it varies independently for each transmission/frame. The $n_r \times T_{frame}$ baseband received signal matrix corresponding to the transmission of a frame of $T_{frame}$ time slots is given by

$$Y = HX + N,$$

where $X$ is the $n_t \times T_{frame}$ transmitted matrix and $N$ is the $n_r \times T_{frame}$ additive white Gaussian noise matrix. Each element of $H$, $h_{j,i}$, is an independent identically distributed (i.i.d.) complex Gaussian random variable with zero mean and variance 0.5 per dimension, representing the channel coefficient between the $j^{th}$ receive and $i^{th}$ transmit antenna. Each element of $N$ is an i.i.d. complex Gaussian random variable with zero mean and variance $N_0/2$ per dimension.

Here we focus on the design of HARQ schemes for the NOSTBCs of [4]. The NOSTBCs are encoded using a $(n,k)$ eRS/BCH code, where $n$ is the number of encoded symbols, $k$ is the number of data symbols and the code’s symbols are defined over $GF(q)$, $q = 2, 4, 16$. In the $GF(16)$ case, the binary information bits are mapped to $GF(16)$ symbols (0, 1, …, 15) and every $k$ symbols are encoded to generate an eRS codeword.

The encoded symbols are mapped to 16-QAM as [4]

$$
egin{array}{cccc}
0 & 8 & 5 & 15 \\
14 & 1 & 9 & 4 \\
7 & 13 & 2 & 10 \\
11 & 6 & 12 & 3
\end{array}
$$

After this mapping, the constellation points chosen by each codeword are written into the NOSTBC matrix column by column. Therefore, there is a one to one mapping between NOSTBC matrices and eRS/BCH codewords. The NOSTBC decoder picks the codeword with smallest squared Euclidean distance using [4]

$$
\hat{X}_k = \arg \min_{(X_c)} \sum_{t=(k-1)\tau+1}^{k\tau} \sum_{j=1}^{n_r} |y_{j,t} - \sum_{i=1}^{n_c} h_{j,i}x_{i,t}|^2,
$$

$$
k = 1, 2, \cdots, T_{frame}/\tau.
$$
where $\hat{X}_k$ is the $k^{th}$ estimated $n_t \times \tau$ NOSTBC matrix in the frame, $\tau$ is the number of time slots the NOSTBC spans, $x_{i,t}$ and $y_{j,t}$ are the transmitted signal from the $i^{th}$ transmit antenna and the received signal from the $j^{th}$ receive antenna at time $t$, respectively, and $\{X_o\}$ is the set of all possible NOSTBC matrices.

3. Proposed ARQ Schemes

We now introduce the ARQ schemes used for comparison purposes (TS-ARQ and HARQ-1) and propose a novel HARQ scheme (HARQ-2).

Truncated selective ARQ (TS-ARQ): The receiver discards the received soft information if the decoding results are in errors. Retransmission is truncated if $m_{\text{max}}$ has been reached [5].

HARQ-1: All previously received signals and corresponding channel state information (CSI) are kept to be combined with the current received signals. This has the effect of increasing the number of “virtual receive antennas” [6] with each retransmission.

HARQ-2: For a given packet, the systematically and non-systematically encoded codewords are transmitted during odd and even numbered transmissions, respectively. A decoding candidate list and the associated Euclidean distance list are generated/updated during each odd numbered transmission and are used to decode all received even numbered transmissions.

3.1. HARQ-1 Scheme

Combining the previously received signals and corresponding CSI with the current received signals has the effect of increasing the number of “virtual receive antennas” [6] with each retransmission. The receiver acts as if it has $mn_r$ receive antennas after $m$ transmissions of the same information. The combined $(m \cdot n_r) \times n_t$ channel matrix is denoted $A$ and the $(m \cdot n_r) \times \tau_{\text{frame}}$ combined received symbols matrix is given by

$$S = AX + N = \begin{bmatrix} Y^m \vdots \vdots Y^1 \end{bmatrix} \begin{bmatrix} H^m \vdots \vdots H^1 \end{bmatrix} X + \begin{bmatrix} N^m \vdots \vdots N^1 \end{bmatrix}, \quad (4)$$

where the superscripts represent transmission number.

3.2. HARQ-2 Scheme

We now propose a HARQ scheme specifically designed for the NOSTBCs of [4], HARQ-2. Its flow chart is shown in Fig. 1. Both systematic and non-systematic eRS/BCH encodings are used to generate the NOSTBCs. In [4] only systematic encoding was used in order to reduce the decoding complexity. In order to keep decoding complexity manageable for higher rate codes, we transmit the systematic codeword during the first transmission and then restrict the codewords considered by the decoder for non-systematic transmissions.

The systematically encoded codewords are transmitted during odd numbered transmissions giving a combined received signal matrix of

$$S^s = A^s X^s + N^s = \begin{bmatrix} Y^{2v-1} \vdots \vdots Y^3 \vdots Y^1 \end{bmatrix} \begin{bmatrix} H^{2v-1} \vdots \vdots H^3 \vdots H^1 \end{bmatrix} X^s + \begin{bmatrix} N^{2v-1} \vdots \vdots N^3 \vdots N^1 \end{bmatrix}, \quad (5)$$

where superscript $s$ denotes systematic transmission and $v$ is the number of systematic transmissions. The non-systematically encoded codewords are sent during
even numbered transmissions giving
\[
S^{ns} = A^{ns}X^{ns} + N^{ns}
\]
(6)

\[
\begin{bmatrix}
Y^{2w} \\
\vdots \\
Y^{4} \\
Y^{2}
\end{bmatrix}
= \begin{bmatrix}
H^{2w} \\
\vdots \\
H^{4} \\
H^{2}
\end{bmatrix} X^{ns} + \begin{bmatrix}
N^{2w} \\
\vdots \\
N^{4} \\
N^{2}
\end{bmatrix},
\]

where superscript \(ns\) denotes nonsystematic transmissions and \(w\) is the number of these transmissions.

We decode \(S^s\) after receiving each systematic transmission. To decode the \(k^{th}\) \((k = 1, 2, \cdots, T_{frame}/\tau)\) NOSTBC block in a frame, we calculate the squared Euclidean distance corresponding to the codewords in \(\{X_k^s\}\) using

\[
d^2[\hat{X}_k^s] = \sum_{t=(k-1)\tau+1}^{k\tau} \sum_{j=1}^{vn} \left| s^s_{j,t} - \hat{s}^s_{j,t} \right|^2,
\]

where \(\hat{X}_k^s\) is the hypothesized received codeword for the \(k^{th}\) NOSTBC codeword in \(X^s\) and \(\{X_k^s\}\) is the set of all systematic NOSTBC codewords. The \((j,t)^{th}\) and \((i,t)^{th}\) element in the matrix \(S^s, A^s\) and \(X^s_k\) are denoted \(s^s_{j,t}, a^s_{i,t}\) and \(\hat{s}^s_{j,t}\), respectively. At the same time, a decoding candidate list, \(\mathcal{L}_k\), of \(|\mathcal{L}_k| = Q\) candidate codewords is generated using

\[
\mathcal{L}_k = \arg \min_{X_k^s \in \{X_k^s\}} d^2[\hat{X}_k^s],
\]

(8)

where \(\arg \min_{X_k^s \in \{X_k^s\}} \) finds the \(Q\) candidate codewords in \(\{X_k^s\}\) with the smallest metric. It can also be generated using the list sphere decoder proposed in [4] or other reduced complexity search algorithms. The Euclidean distance corresponding to the codewords in \(\mathcal{L}_k\) is stored in a vector denoted \(d^2[\mathcal{L}_k]\). The list is updated after each received systematic transmission as the error correction capability increases with \(v\).

The lists \(\mathcal{L}_k\) and \(d^2[\mathcal{L}_k]\) are then used to decode the received nonsystematic signal matrix \(S^{ns}\) for the \(k^{th}\) \((k = 1, 2, \cdots, T_{frame}/\tau)\) NOSTBC block according to

\[
\hat{X}_k = \arg \min_{\mathcal{L}_k} \sum_{t=(k-1)\tau+1}^{k\tau} \sum_{j=1}^{vn} \left| s^s_{j,t} - \hat{s}^s_{j,t} \right|^2 + d^2[\mathcal{L}_k],
\]

(9)

where \(d^2[\mathcal{L}_k]\) is the squared Euclidean distance corresponding to the \(q^{th}\) codeword in \(\mathcal{L}_k\) and \((x^t)q\) is the \((i,t)^{th}\) element of the nonsystematic codeword matrix corresponding to the \(q^{th}\) systematic codeword in the list, \(\mathcal{L}_k^q\). The codeword with the minimum overall Euclidean distance is the output of the decoder. Alternatively, a list of codewords can be generated by replacing \(\min\) with \(\min^Q\) in (9). This would allow reliability information to be calculated for subsequent decoding stages in a concatenated scheme.

### 3.3. Choice of List Size

We now discuss how to choose the list size, \(|\mathcal{L}_k|\). The error probability when using a decoding list is

\[
P_{error} = P_{MLD} + (1 - P_{MLD})P_{list} < P_{MLD} + P_{list},
\]

(10)

where \(P_{MLD}\) is the probability of a maximum likelihood decoding (MLD) error and \(P_{list}\) is the probability of the list not containing the MLD result. Calculating the exact value of \(P_{list}\) for a MIMO fading channel is difficult and unnecessary for the problem at hand. Here, we develop a simple method to determine the list size required to obtain near MLD performance at significantly reduced complexity. The complexity savings are expected to be larger for the higher rate NOSTBCs. Our approach can be extended to other block code based ARQ schemes.

The list size must be large enough to ensure a small value of \(P_{list}\). This may be indicated by the ratio between the summed PEP contributed by codewords in the list and the summed PEP of all codewords. To this end, we consider the list generated after the first (systematic) transmission for the \(k^{th}\) NOSTBC codeword.

The PEP between a pair of codeword matrices is [7]

\[
P(X_k, \hat{X}_k) \leq \frac{1}{4} \exp \left( \frac{1}{2} \left( \frac{E_s}{4N_0} \right)^2 n_r \sum_{i=1}^{r} \lambda_i \right) \cdot Q \left( \sqrt{n_r} \left( \frac{E_s}{4N_0} \sum_{i=1}^{r} \lambda_i - \sum_{i=1}^{r} \lambda_i \right) \right),
\]

(11)

where \(E_s\) is the energy per symbol at each transmit antenna, \(Q(x)\) is the complementary error function, \(\lambda_i, i = 1, \cdots, r\) are the non-zero eigenvalues of the codeword distance matrix \(A(X_k, \hat{X}_k) = (X_k - \hat{X}_k)(X_k - \hat{X}_k)^H\) and \((-)^H\) denotes the Hermitian transpose operator. The ratio of PEPs for a list of size \(|\mathcal{L}_k|\) is

\[
\mathcal{R}(|\mathcal{L}_k|) = \frac{\mathbb{E}[X_k] \left( \sum_{\hat{X}_k \in \mathcal{L}_k, \hat{X}_k \neq X_k} P(X_k, \hat{X}_k) \right)}{\mathbb{E}[X_k] \left( \sum_{\hat{X}_k \in \{X_k^s\}, \hat{X}_k \neq X_k} P(X_k, \hat{X}_k) \right) X_k \in \{X_k^s\},
\]

(12)
Equation (10) shows that it is meaningful to discuss the list size: list size \( |L_k| \) only when \( \mathcal{P}_{\text{MLD}} \) is sufficiently small, therefore, a sufficiently large SNR must be considered.

We propose the following method for choosing the list size:

1. Calculate \( \mathcal{R}(|L_k|) \) for various \( |L_k| \).
2. Find the SNR where the first transmission starts to have a DPR < 1.
3. Choose a list size that obtains \( \mathcal{R}(|L_k|) \approx 1 \) or \( \mathcal{R}(|L_k|) > 0.9 \) at the SNR from step 2, for a negligible or small loss in performance, respectively.

4. Simulation Results and Discussion

We use the NOSTBC defined using the rate \( R_{\text{sec}} = 3/16 \) (16, 3) eRS code and 16-QAM (\( E_b = 1 \)). Since it only has 4096 codewords, we use a brute force decoder. However, the proposed schemes can easily be used with higher rate NOSTBCs and the list decoding algorithm of [4]. We use a rate \( R_{\text{CRC}} = 0.985 \) CRC-16 (USA) code [8] of length 1080 for error detection. The CRC encoded bits are fed into the NOSTBC encoder, which has rate \( R_{\text{sec}} = n_t \times R_{\text{sec}} = 3/4 \) and spans \( \tau = 4 \) time slots. The resulting packet spans \( n_t = 4 \) antennas and 90 NOSTBC blocks (\( T_{\text{frame}} = 360 \)).

Performance is measured using DPR and TSE. Only finite delay is allowed in most realistic communication systems. Therefore, a packet is declared dropped if it still contains errors after \( m_{\text{max}} \) transmissions. We set \( m_{\text{max}} = 4 \) as in [5] [9]. The DPR is computed as

\[
\text{DPR} = \frac{p_{\text{drop}}}{p_{\text{total}}} \quad (13)
\]

where \( p_{\text{drop}} \) is the number of dropped packets and \( p_{\text{total}} \) is the total number of packets transmitted (excluding retransmissions). For M-QAM, the TSE is computed by

\[
\text{TSE} = \frac{\log_2(M) \times R_{\text{sec}} \times R_{\text{CRC}} \times (1 - P_{\text{Drop}})}{m_{\text{sec}}}, \quad (14)
\]

where \( m_{\text{sec}} \) is the average number of transmissions of a packet. All results are plotted against the data bit energy, \( E_b \), divided by

\[
N_0 = \frac{n_t E_b}{\log_2(M) R_{\text{sec}} 10^{b_1 SNR} R_{\text{CRC}}} \quad (15)
\]
about 0.1 dB loss. A list of 100 codewords causes about 0.5 dB loss while a list of 10 codewords causes a significant performance loss compared to MLD.

4.2. Simulation Results

Unless otherwise stated simulation results presented for HARQ-2 use $|\mathcal{L}_k| = 500$. Fig. 4 shows the DPR curves and Fig. 5 shows the TSE curves, which tend to $\log_2(M) \times R_{\text{ostbc}} \times R_{\text{CRC}} = 2.95 \text{ bits/s/Hz}$. This is a much higher value than that offered in [1] and [2]. Note that if we use the eRS(16,9) NOSTBC in [4] the maximum achievable TSE increases to 8.85 bits/s/Hz.

Fig. 4 and Fig. 5 clearly show that HARQ-1 and HARQ-2 with the eRS(16,3) NOSTBC offer significant improvements in DPR (approximately 5 dB and 6 dB at a DPR of $10^{-3}$, respectively) and TSE compared to the TS-ARQ scheme. They also show that HARQ-2 has the best performance by approximately 1 dB.

We now consider a time varying Rayleigh fading channel\(^2\) with normalized Doppler frequencies of $f_d T = 0.0001, 0.001$ and 0.01. The DPRs are shown in Fig. 6. The proposed HARQ schemes are more robust over these channels. All curves are obtained assuming retransmissions are separated by 5 packet lengths. The performance of the TS-ARQ scheme gets worse as the channel variation gets faster. The performance of HARQ-1 and HARQ-2 gets worse for $f_d T = 0.0001$ as the channel is highly correlated for retransmissions, reducing the benefit of retransmissions. The performance for $f_d T = 0.0001$ can be improved by transmitt-

---

\(^2\)Note the channel matrix for the last time slot of each NOSTBC block is used for all $\tau = 4$ time slots.

4.3. Discussion

HARQ design for MIMO systems can be viewed from a STC design perspective because each retransmission has the effect of increasing the number of virtual receive antennas [6]. Therefore, if a packet is transmitted $m$ times, the diversity order is at least $r m n$, assuming independent CSI for each transmission. The design criteria for MIMO systems with diversity order of four or more is to maximize the minimum Euclidean distance between codewords [7]. Therefore, the component codes should have as large a minimum Euclidean distance as possible and the ARQ scheme should be designed to maximize this. HARQ-1 achieves this by combining all soft information from previous transmissions and HARQ-2 further increases the minimum distance by transmitting and then combining systematic and non-systematic codewords.

The HARQ-2 scheme maps information bits to codewords in two different error control codes to increase the total Euclidean distance. This can achieve a large coding gain as long as the codes are chosen properly. For example, the minimum squared Euclidean distance for both systematic and non-systematic (16,3) eRS codewords is 6.4. The combined minimum squared Eu-
Figure 5: TSE of an OSTBC and NOSTBC using the ARQ schemes and $m_{max} = 4$.

cclidean distance is $2 \times 6.4 = 12.8$ for HARQ-1 with 1 retransmission. For the HARQ-2 scheme, the combined minimum squared Euclidean distance after 1 retransmission is 20.6 due to the use of both systematic and non-systematic encodings. This means information vectors that are mapped close together by a systematic encoding, are mapped further apart by the non-systematic encoding.

5. Conclusions

In this paper, we proposed a list decoding algorithm and a novel HARQ scheme using the NOSTBCs of [4]. The data packets are CRC and NOSTBC encoded, but no other error control code is used. This is done to simultaneously achieve high rates and good performance. Simulation results show that the proposed list algorithm significantly reduced the decoding complexity without degrading performance. HARQ-2, which is designed for the NOSTBCs of [4], achieves significantly better performance than TS-ARQ and HARQ-1.

References


