

**Performance of Hybrid ARQ Using Trellis Coded Modulation over Rayleigh Fading Channel**

Qian Huang, Sammy Chan, Li Ping and King-Tim Ko

**Abstract**—Trellis coded modulation has been shown as an attractive coding scheme for broadband wireless networks due to its bandwidth efficiency. Various hybrid ARQ schemes using trellis-coded modulation (TCM) as underlying forward error control (FEC) have been proposed in the literature. In this paper, we study the performance of the hybrid ARQ with concatenated two-state trellis-coded modulation (CT-TCM) over Rayleigh fading channels. The throughput of the CT-TCM hybrid ARQ scheme over a quasi-static Rayleigh fading channel is evaluated by both analysis and simulation. Besides, the performance of the CT-TCM hybrid ARQ over fully interleaved Rayleigh fading channels is evaluated by simulation and compared with other existing TCM-based hybrid ARQ schemes. The numerical results demonstrate that the CT-TCM hybrid ARQ achieves improved coding gain and throughput.

**Index Terms**—Hybrid ARQ, trellis-coded modulation (TCM).

I. INTRODUCTION

Coded modulation techniques, such as trellis coded modulation (TCM) [1]–[4], have been studied for the physical layer design to improve bandwidth efficiency in broadband wireless networks. As an important application of TCM codes for error control, various hybrid ARQs using TCM as the underlying FEC scheme have been investigated [5]–[12]. A simple approach is to use the copies of a TCM code in all further retransmissions, as the schemes in [5]–[7], [9]. However, it results in a repetition coding and is not efficient as far as coding gain is concerned. The schemes of [10]–[12] improve the coding gain by combining different TCM codes used for retransmissions. In the scheme of [10], two different TCM codes are alternatively used in further retransmissions. However, this approach becomes a repetition coding after two transmission attempts. An alternative solution is to employ more component codes and retransmit a different component code at a time. However, the design of the component codes used in such a TCM-based scheme is still an open problem.

In [12], we have proposed a novel hybrid ARQ scheme that uses the concatenated two-state TCM (CT-TCM) codes presented in [4]. The notable feature of the CT-TCM codes is that they offer a low-cost but comparable performance to some existing Turbo-type coded-modulation schemes [2], [3]. The CT-TCM encoder comprises multiple component encoders concatenated in parallel by symbol-based interleavers as shown in Fig. 1, where the number of component encoder is $M$. The symbol-based interleaver for component encoder $m$ ($0 \leq m < M$) is specified as $z^{(m)} = \{z_k^{(m)}\}$, of which the $k$th element ($k \geq 0$), denoted as $z_k^{(m)}$, is an integer randomly generated but follows the rule: $z_k^{(m)} \mod M = k \mod M$. With the interleaver, the $k$th symbol $d_k$ (with binary $n$-tuple) of an input information sequence $\{d_k\}$, $k \geq 0$, is moved to a new position specified by $z_k^{(m)}$ in the sequence after interleaving. The sequences of interleaved information symbols are encoded by each component encoder with a binary two-state trellis encoder followed by a multi-ary signal mapper. The coded symbols from each component encoder, $\{e_k^{(m)}\}$, are then mapped to a signal constellation of size $2^{n+1}$, producing a sequence of modulated symbols, e.g., $\{d_k^{(m)}\}$ from component encoder $m$. We refer to the sequence of modulated symbols from each component encoder as a CT-TCM component code. The CT-TCM code is finally generated by heavily puncturing the symbols of each component code under a given puncture pattern before they are transmitted for spectral efficiency.

At the receiver, the received codes from the initial transmission and the retransmissions are jointly decoded based on the turbo decoding principle as shown in Fig. 1. The CT-TCM decoder comprises $M$ local a posteriori probability (APP) decoders, one for each CT-TCM component code. The $M$ local APP decoders operate successively. For the $m$th decoder, $T$
is for delay of one iteration, $z^{(m)}$ is the symbol interleaver, and the other variables are the log-likelihood (LL) values and defined as follows.

- $L^{(m)}$ represents the a posteriori LL values for all information symbols after decoding the $n$th component code.
- $L^{(m)}$ represents the LL values $\{\log p(y_k|x_k)\}$ based on individual channel observations of the $n$th component code, where $x_k$ is the $k$th modulated symbol and $y_k$ denotes the $k$th received symbol; its elements $\{L^{(m)}_{c,k}\}$ are calculated as
  \[
  L^{(m)}_{c,k} = \begin{cases} 
  \log p(y_k|x_k), & \text{for unpunctured symbols} \\
  0, & \text{for punctured symbols.}
  \end{cases}
  \]
- $L^{(m)}$ represents the a priori LL values for all information symbols for the $n$th component code. It is initialized to zeroes, implying no a priori information.
- $L^{(m)}$ represents the extrinsic information produced by the $n$th component code, defined by
  \[
  L^{(m)}_e = L^{(m)} - L^{(m)}_a.
  \]

The performance of the CT-TCM hybrid ARQ (CT-TCM-HARQ) scheme in an additive white Gaussian noise (AWGN) channel has been investigated by simulation in [12]. In this paper, we focus on the performance analysis of this scheme over Rayleigh fading channels. We first provide a brief review of the CT-TCM-HARQ in Section II, which is followed by a throughput analysis of the scheme over a quasi-static Rayleigh fading channel in Section III. Numerical results and performance comparisons between the CT-TCM-HARQ and other TCM-based hybrid ARQs are presented in Section IV. Finally, our conclusions are provided in Section V.

## II. Hybrid ARQ Scheme Using CT-TCM Codes

In terms of the aforementioned CT-TCM principle, for a given input data block, a set of CT-TCM codes can be generated by only changing the puncture patterns used at the sender. It is reasonable to use the resulted CT-TCM codes in the initial transmission and further retransmissions of the data block. Therefore, the key issue for designing a hybrid ARQ using the CT-TCM codes is the assignment of puncture patterns in each transmission attempt of a data block. Based on the CT-TCM code principle [4], the puncture pattern for generating a CT-TCM code first satisfies the following constraints: (1) At any time $k \geq 0$, one and only one modulated symbol, from the component encoders, carrying the same information $d_k$, is selected and transmitted. That is, for any $k$, only one symbol of $\{x^{(m)}_k, 0 \leq m < M\}$ is transmitted and the other modulated symbols are punctured under the given puncturer. (2) The punctured modulated symbols are uniformly distributed in each component code. In terms of the above constraints, a qualified puncturer forms a pattern of a binary $M \times M$ indication matrix that contains only one non-zero entry per column. Assume $\Gamma$ a qualified puncture pattern with $\Gamma(l,m) = 1$ $(0 \leq l, m < M)$ the non-zero entry in row $l$, column $m$. Under this puncturer, all modulated symbols of the $n$th component code are punctured except those at the positions $\{k | k \mod M = l\}$. Table I presents four examples of puncture patterns, $\Gamma_i$ $(0 \leq i \leq 3)$, which satisfy the above constraints for a CT-TCM system with $M = 4$ and 8PSK modulation. If $\Gamma_1$ is used by the CT-TCM encoder, the transmitted sequence only contains the following modulated symbols $x^{(1)}_k, x^{(3)}_{k+1}, x^{(0)}_{k+2}, x^{(1)}_{k+3}, x^{(2)}_{k+4}, x^{(3)}_{k+5}, x^{(3)}_k, \ldots$

Additionally, considering the combined decoding of successively received packets in retransmissions for optimal coding gain, it generally requires that the received information symbols of a data block are evenly distributed in the combined CT-TCM codes. Based on this condition, the puncture patterns used in two consecutive transmission attempts of a data block can be selected from the qualified puncture patterns whose combination also satisfies constraint (2). That is, if the combined puncture pattern is used as the puncturer at the CT-TCM sender, the modulated symbols that have not been punctured are still uniformly distributed in each CT-TCM component code. For example, in the CT-TCM HARQ with $M = 4$, if we use puncture pattern $\Gamma_0$ in Table I for the initial transmission of a data block. With respect to the above selection rule, the puncture pattern used for the second transmission of the data block is preferred to be $\Gamma_1$ in Table I. At the receiver, the combination of the received sequences with puncture patterns $\Gamma_0$ and $\Gamma_1$ is equivalent to a sequence generated by using puncture pattern

\[
\begin{pmatrix}
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1
\end{pmatrix}
\]

at the sender, which provides evenly puncturing on each CT-TCM component code. The similar method is used to determine the puncture patterns for the following retransmissions.

Denote $\{\gamma_i\}, 1 \leq i \leq M$, a set of puncture patterns satisfying all above selection rules for a CT-TCM encoder with $M$ component encoders. Based on the above considerations, the hybrid ARQ with the CT-TCM codes performs as follows:

(i) Encoding and transmitting: the information sequence is sent to the CT-TCM encoder. Puncture pattern $\gamma_1$ is used to puncture the generated component codes. However, the punctured modulated symbols of each component code are not dropped. Instead they are stored in a buffer at the sender for further use. A packet containing the modulated symbols selected by puncture pattern $\gamma_1$ is transmitted to the receiver.

(ii) Decoding: the received sequence together with the sequences from the initial transmission and previous retransmissions are jointly decoded based on the turbo decoding principle. If there is no error detected in decoding, an ACK is returned to the sender. Otherwise, the received sequence is stored in the buffer at the receiver, and a negative acknowledgement (NAK) is returned to the sender.

### Table I

| Examples of puncture patterns for CT-TCM with $M = 4$. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| $\Gamma_0$ | $\Gamma_1$ | $\Gamma_2$ | $\Gamma_3$ |
| 1 0 0 0 | 0 1 0 0 | 0 1 0 0 | 0 0 0 1 |
| 0 1 0 0 | 0 0 0 1 | 0 1 0 0 | 0 0 0 1 |
| 0 0 0 1 | 0 1 0 0 | 0 0 0 1 | 0 1 0 0 |
| 0 0 0 1 | 0 1 0 0 | 0 0 0 1 | 0 1 0 0 |
(iii) Retransmitting: once the sender receives an NAK, another puncture pattern $\gamma_i$, $2 \leq i \leq M$, is performed over the stored component codes. A new modulated sequence out of the stored component codes is selected by $\gamma_i$ and transmitted. $i$ is increased by one after each retransmission. Then, the scheme repeats procedures (ii)–(iii). If $M$ transmissions are reached, the receiver buffer is cleared, and the scheme repeats (i)–(iii).

One of the advantages of the CT-TCM-HARQ is that the generation of retransmitted packets involves no complex computation or procedures. Another advantage, as we have demonstrated in [12], is that significantly improved coding gain is achieved in comparison to the repetition method.

### III. Throughput Analysis

For a general hybrid ARQ scheme, provided that the bit error rate (BER) performance in each transmission attempt is known, the average number of transmissions per data block by the hybrid ARQ scheme can be derived as a function of the BER performance [13], [14]. Similar method is also applicable in the throughput analysis for the CT-TCM-HARQ scheme. Let the events of success and failure in decoding the received CT-TCM code generated by puncture pattern $\gamma_i$ ($1 \leq i \leq M$) that is used the $j$th ($j \geq 1$) time be denoted as $S_i^{(j)}$ and $D_i^{(j)}$, respectively. Let $P(S_i^{(j)})$ and $P(D_i^{(j)})$ represent the probabilities that each event occurs, respectively. Let $P(V_i^{(j)})$ and $P(F_i^{(j)})$, represent the probabilities of success and failure in decoding the combination of all received CT-TCM codes for a data block that are generated by the puncture patterns $\gamma_1, \gamma_2, \ldots, \gamma_i$ by $j$ times, respectively. We assume that the probability of the undetected errors contained in any decoded CT-TCM codes failed in decoding. Denote such conditional event as $E$. Since it is definitely that no successful transmission occurs before using $\gamma_i$ $j$ times, we have

$$P(D_i^{(j)}) = P(D_i^{(j)}|E)P(E).$$

Due to the fact that is $P(E) \leq 1$, we obtain

$$P(D_i^{(j)}|E) \geq P(D_i^{(j)}).$$

Similarly, we have $P(F_i^{(j)}|E) \geq P(F_i^{(j)})$ for the term of $P(F_i^{(j)}|D_i^{(1)}, D_i^{(2)}, F_i^{(1)}, \ldots, D_i^{(j)} F_i^{(j)}, \ldots, F_i^{(j)} F_i^{(j)})$. Hence, Eq. (2) can be re-written as follows,

$$P(D_i^{(1)}, D_i^{(2)}, F_i^{(1)}, \ldots, D_i^{(j)}, F_i^{(1)}, \ldots, F_i^{(j)} F_i^{(j)}) \geq P(D_i^{(1)})P(D_i^{(2)}) \cdot P(F_i^{(2)}) \cdot \cdots \cdot P(D_i^{(j)})P(D_i^{(2)})P(F_i^{(j)}).$$

Substituting Eq. (5) into Eq. (1), we obtain

$$T_r \geq 1 + \sum_{i=1}^{M} \Pi_{i=1}^{j-1} \Pi_{i=2}^{M} P(D_i^{(j)})P_i^{M} \cdot \sum_{j=1}^{\infty} \Pi_{j=1}^{j-1} \Pi_{i=1}^{M} P(D_i^{(j)})P_i^{M}.$$  

On the other hand, the following relationships also hold:

$$P(D_i^{(1)}, D_i^{(2)}, F_i^{(1)}, \ldots, D_i^{(j)}, F_i^{(1)}, \ldots, F_i^{(j)} F_i^{(j)}) \leq P(F_i^{(j)})$$

$$P(D_i^{(1)}, D_i^{(2)}, F_i^{(1)}, \ldots, D_i^{(j)}, F_i^{(1)}, \ldots, F_i^{(j)} F_i^{(j)}) \leq P(D_i^{(j)}).$$
Substituting Eq. (7) and Eq. (8) into Eq. (1), we can obtain

$$T_r \leq 1 + \sum_{j=1}^{\infty} P(D_j) + \sum_{i=2}^{M} P(F_i).$$

Combining Eqs. (6) and (9), we derive the bounds of the average number of transmissions $T_r$ per data block by the CT-TCM-HARQ.

Using $T_r$, the throughput efficiency of the CT-TCM-HARQ can be calculated by $((1/T_r) \times \text{the number of information bits per transmitted symbol})$. For the CT-TCM scheme based on 8PSK modulation and rate 2/3 CT-TCM codes, the maximum achievable throughput efficiency (with ideal channel) is 2 bits/symbol. The throughput efficiency of the CT-TCM-HARQ based on such CT-TCM codes is thus given by $2/T_r$.

From [12], the BER in each transmission attempt for the CT-TCM-HARQ over an AWGN channel has been obtained. Based on the BER values and the bounds of $T_r$ presented in Eqs. (6) and (9), we can obtain the throughput bounds of the CT-TCM-HARQ scheme over the AWGN channel. Then, the throughput of the CT-TCM-HARQ scheme in the quasi-static Rayleigh fading channel can be derived by averaging the throughput in the AWGN channel over the probability density function (pdf) of the signal-to-noise-ratio (SNR) in Rayleigh fading. Let $\alpha$ denote the Rayleigh-distributed fading envelope. Let $E_b/N_0$ denote the ratio of bit energy to noise power spectral density. The SNR in the quasi-static Rayleigh fading channel is written as $\beta = \alpha^2 E_b/N_0$. The pdf of the Rayleigh-distributed fading envelope $\alpha$ is

$$f(\alpha) = \frac{\alpha}{\sigma^2} e^{-\alpha^2/2\sigma^2}.$$ (10)

Thus, the pdf of $\beta$ is derived from the pdf of $\alpha$ as follows:

$$f(\beta) = \frac{1}{2\sigma^2 E_b/N_0} e^{-\beta/2\sigma^2 E_b/N_0},$$ (11)

where $2\sigma^2$ is the second moment of $\alpha$, i.e., $E[\alpha^2] = 2\sigma^2$. For simplicity, we choose $E[\alpha^2] = 1$. Let $f(R_{AWGN}|\beta)$ denote the throughput of the CT-TCM-HARQ scheme over the AWGN channel conditional on a certain value of SNR $\beta$. Then, the throughput in the quasi-static Rayleigh fading channel, denoted by $R$, can be obtained by numerical integration of the throughput in the AWGN channel as follows:

$$R = \int_0^\infty f(R_{AWGN}|\beta) f(\beta) d\beta.$$ (12)

IV. PERFORMANCE EVALUATIONS

We evaluate the performance of the CT-TCM-HARQ over a quasi-static Rayleigh fading channel in terms of the following parameters for the CT-TCM-HARQ. Without loss of generality, we choose $M=4$, 8PSK modulation for each information packet containing 1024 symbols (2048 bits), and a maximum of four transmissions per information packet. Additionally, for revealing the essential performance of the CT-TCM-HARQ, it is reasonable to assume that a stop-and-wait ARQ protocol is used as the ARQ strategy and that the feedback messages (ACKs and NAKs) returned by the receiver are always correctly received by the sender.

We first verify the throughput analysis of the CT-TCM-HARQ scheme in the quasi-static Rayleigh fading channel by simulation. Based on the BER values obtained in [12], we have calculated the upper and lower throughput bounds of the CT-TCM-HARQ over the AWGN channel. The results show that the bounds are very tight and hard to be distinguished when they are plotted. Thus, from either upper or lower bounds for the AWGN channel, we then calculate the approximate throughput of the CT-TCM-HARQ over the quasi-static Rayleigh fading channel by the aforementioned numerical integration method. We present in Fig. 2 the throughput obtained by analysis and simulation in the quasi-static Rayleigh fading channel. The results demonstrate the simulation results match the analytical results very well. For further performance evaluation over the quasi-static Rayleigh fading channel, in Figs. 2 and 3, we compare the throughput and the residual
frame error rate (FER) obtained by the CT-TCM-HARQ with the results obtained by a Type-I ARQ which uses the CT-TCM code as the underlying FEC scheme, respectively. In this Type-I ARQ, the same CT-TCM code generated by puncture pattern \( \Gamma_0 \) in Table I is used for the initial transmission and all further retransmissions. It is observed that compared with the Type-I ARQ scheme with the repetition coding of the CT-TCM code, the CT-TCM-HARQ save about one dB SNR for a given FER value or a given throughput. This demonstrates that by using different punctured parts in different retransmissions, significantly improved coding gain can be achieved in the CT-TCM-HARQ over the quasi-static Rayleigh fading channel.

In Fig. 4 and Fig. 5, we compare the BER and throughput performances over a fully interleaved Rayleigh fading channel, between the CT-TCM-HARQ and the hybrid ARQ schemes of [9] which use the Turbo-based TCM (TTCM) proposed in [1] with 8PSK modulation as the underlying FEC scheme, and also set the length of information packet to be 1024 symbols. Here, we compare with the following two TTCM hybrid ARQ schemes from [9]:

- Scheme T-IB : a Type-I hybrid ARQ, which uses the same TTCM code based on a rate 2/3 Turbo code in each retransmission of an information packet. For decoding, the receiver uses equal diversity combining to combine all copies of the received packet into a single packet of the same block size by adding the demodulated log-likelihood values of each packet.
- Scheme T-II : a Type-II hybrid ARQ with incremental redundancy, which uses a rate 1/3 as the mother code mapped to 8PSK constellation, and sends the information bits in the first transmission and then incrementally sends the parity-check bits in retransmissions.

Fig. 4 demonstrates that improved coding gain is achieved by CT-TCM-HARQ than the TTCM based Type-I and Type-II hybrid ARQs. For throughput performance, we only consider a comparison between our CT-TCM-HARQ scheme and Scheme T-IB of [9] in Fig. 5, because both of them achieve a maximum throughput efficiency of 2 bits/symbol. It is observed from Fig. 5 that the CT-TCM-HARQ scheme provides higher throughput than Scheme T-IB of [9], particularly in the low SNR range.

V. CONCLUSION

In this paper, we investigated the performance of a novel hybrid ARQ using the CT-TCM code in both slow and fast Rayleigh fading channels. The numerical results have demonstrated that such a method of using CT-TCM codes in hybrid ARQ outperforms an existing hybrid ARQ using Turbo-type TCM codes over the fully interleaved Rayleigh channels.

REFERENCES


