Iterative Demodulation and Decoding of Trellis Coded CPM

Krishna R. Narayanan
Department of Electrical Engineering
Texas A&M University,
College Station, TX 77843, USA

ABSTRACT

Interleaved trellis coded systems with continuous phase modulation (TCCPM) are considered. It is shown that the combination of a trellis code, interleaver and the CPM modulator results in a significantly enhanced distance spectrum for the transmitted signals. Upper bounds on the bit error rate are derived by considering the combination of the trellis code, interleaver and the CPM modulator as serially concatenated code. Finally, iterative coherent and non-coherent receivers are proposed and their performance is evaluated through computer simulations.

I Introduction

Many communication systems employ non-linear power amplifiers and, hence, require a low peak-to-average power ratio for the modulated signal. Examples of such systems include satellite and mobile communications systems. Continuous phase modulation (CPM) is a constant envelope modulation technique and, hence, a good choice for such systems. Trellis coded modulation (TCM) is a technique that combines modulation and coding in order to achieve coding gains without sacrificing data rate or bandwidth. When TCM is combined with CPM, high power and bandwidth efficiency can be achieved.

The distance spectrum and error performance of convolutionally encoded CPM on additive white Gaussian noise (AWGN) channels has been analyzed by many researchers. However, most of the work has not considered the use of an interleaver between the trellis code and the CPM modulator for AWGN channels. Similarly, they have also ignored the effect of the recursive nature of CPM modulator in designing TCCPM schemes. The performance of TCCPM for fading channels has been evaluated in [1],[2],[3]. A fading channel typically exhibits correlation in time and, when the correlation is high, the channel exhibits prolonged deep fades. Therefore, some form of interleaving has to be used to combat the fading correlation. When TCM is used with PSK, PAM, or QAM, the modulated symbols are interleaved using a block or convolutional interleaver. Since interleaving attempts to destroy the correlation in the fading process, performance bounds can be derived by assuming that the channel is uncorrelated. However, when CPM is used, the modulated signal cannot be interleaved because interleaving destroys the continuous phase property of the modulated signal. Therefore, a system was proposed in [1] where an interleaver is used between the trellis code and the CPM modulator. At the receiver, a soft-output CPM demodulator generates metrics for the coded symbols. These metrics are then deinterleaved and used to derive the branch metrics for the Viterbi decoder that is used to decode the trellis code. This deinterleaving makes the channel 'appear' uncorrelated to the trellis code. Performance bounds can then be developed by suitably modifying the transfer function of the trellis code and assuming that the channel is uncorrelated [2].

All of the aforementioned approaches completely ignore the contribution of the interleaver to the overall distance spectrum of the transmitted signals. From the recent advances in concatenated schemes, we know that the interleaver drastically influences the distance spectrum and, hence, will significantly change the system performance [4].

Our approach is to treat TCCPM as a serial concatenation of a trellis code and a CPM modulator, which represents a rate-1 recursive inner code. In contrast to the other papers, we explicitly consider the effect of the interleaver on the distance spectrum of the transmitted signals. The same approach was used in [5] to analyze the performance of interleaved convolutionally encoded systems with differential phase-shift keying (DPSK). In [5], performance bounds were derived for Rayleigh fading channels by assuming perfect interleaving of the modulator output. In this paper, we focus on TCCPM schemes for fading channels. As mentioned before, unlike the case of PSK, for CPM the modulated symbols cannot be interleaved prior to transmission. Consequently, we cannot assume an interleaved fading channel in the case of CPM signals. This makes a significant difference in the performance bounds developed in this paper in comparison to the ones in [5].

Finally, we consider iterative receivers for coherent and non-coherent reception of TCCPM. The coherent receiver is identical to the one in [6]. A non-coherent receiver for TCCPM was proposed in [3] where a non-coherent front end is used to generate metrics for the coded bits and a Viterbi decoder is used for the trellis code. In this paper, we derive a non-coherent detector for iterative demodulating and decoding TCCPM. Our receiver structure is similar to the one in [7] and based on the maximum-likelihood block detection principle in [8].

II System Model

The TCCPM system considered in this paper is similar to the system in [1], [2], but with a few differences. At each time instant, $m$ bits of the binary data sequence $a$ are encoded...
by a rate-$m/n$ convolutional code and binary outputs of the convolutional encoder are multiplexed into the sequence $b$. The binary sequence $b$ is then interleaved into the sequence $\hat{b}$. Then, $p$-tuples of the sequence $\hat{b}$ are mapped on to symbols $c_k \in \{\pm 1, \pm 3, \ldots, \pm (2^p - 1)\}$ and the sequence $\hat{c}$ is input to the CPM modulator. If $p = n$, the symbol rate is the same as the uncoded data rate and, hence, there is no bandwidth expansion. We refer to such systems as trellis coded CPM. If $p < n$, the required bandwidth is higher for the coded system than the uncoded system and, we refer to such systems as convolutionally encoded CPM. Note that we consider bit interleaving of the outer code words in contrast to the symbol interleaving used in [1], [2].

The low-pass equivalent CPM signal $s(t, c)$ for $nT < t \leq (n + 1)T$ is given by

$$s(t, c) = \sqrt{\frac{2nE_k}{nT}} \exp \left(2j\left(\frac{E_k}{nT}\right)T \right)$$

The information is contained in the phase $\theta(t, c) = \theta_0 + 2\pi\beta\sum_{m=-\infty}^{\infty} c(t - mT)$, where $q(t)$ is a frequency shaping pulse such that $q(t) = 1/2$ for $t > JT$, $\theta_0 = \pi\beta\sum_{m=\infty}^{\infty} c_1$ is the accumulated phase at the beginning of the $n$th epoch, and $\beta$ is the modulation index. We will assume that $\beta$ is rational. If $J = 1$, the CPM scheme is called full-response CPM, and if $J > 1$ it is called partial-response CPM. In this paper, we will only consider full-response CPM. A full response CPM signal can also be represented by $s(t, \theta_c)$ to emphasize the finite-state nature of CPM and, hence, suggest a trellis representation similar to that of convolutional codes.

When the modulated signal is transmitted over a frequency non-selective channel, under the assumption that the channel gain remains constant over one modulated symbol, the received signal can be expressed as

$$r(t) = \alpha_k s(t) + n(t), \quad kT < t \leq (k + 1)T$$

where $\alpha_k = \alpha(kT)$. For the AWGN channel, $\alpha_k = 1, \forall k$ and for the flat Rayleigh fading channel $\alpha_k$ is a random variable with a Rayleigh probability density function (PDF).

### III CPM modulator as recursive inner code

Consider the low pass equivalent representation of $M$-ary CPM schemes in Section II. It was shown in [9] that an equivalent way to represent the CPM signal is in terms of the physical tilted phase, defined as $\psi(t) = \theta(t) + \frac{\pi(M-1)}{M} \bmod 2\pi$. The CPM modulator can then be represented as a finite state machine with input $u_n = x_n + (M - 1)/2$ and state $S_n = [\psi_n, u_n]$ as shown below

$$s(t, c) = f(\sigma_n, t)$$

$$\sigma_{n+1} = g(\sigma_n, u_n)$$

Note that since $c_k \in \{\pm 1, \pm 2, \ldots, \pm (2^n - 1)\}$, $u_k \in \{0, 1, \ldots, M - 1\}$. It was shown in [9] that $g$ is time-invariant and can be synthesized as a differential encoder with arithmetic operations in the ring of integers modulo $B$. That is, $\psi_n = 2\pi v_n/B$ where $v_n$ is given by

$$v_k = v_{k-1} \oplus u_{k-1} \quad (4)$$

where $\oplus$ denotes modulo-$B$ addition. It was also shown in [9] that for $nT < t \leq (n + 1)T$, $f$ is a memoryless mapper that maps $\sigma_n$ on to any one of the $BM^J$ possible CPM signals from the CPM signal set $S$. From (3) and (4), we can see that the CPM modulator is equivalent to a recursive rate-$J/J + 1$ convolutional encoder, also called the continuous phase encoder (CPE) followed by a memoryless mapper. We can consider the combination of the trellis code, interleaver and the recursive CPE as a SCC. In the following, we take this approach. Since the inner code in the SCC is recursive, large interleaving gains similar to conventional SCCs should be possible [4].

### IV Performance Analysis

The combination of the convolutional code, interleaver and the modulator can be considered as an equivalent block code, whose code words are transmitted over the correlated fading channel. Therefore, we need to derive the pairwise error probability between the transmitted CPM signal $x_0 = A \exp j\phi_0(t)$ and another signal $x_1 = A \exp j\phi_1(t)$ for a correlated fading channel. By extending the derivation in [10], the pairwise error probability for flat Rayleigh fading channels can be shown to be [11]

$$P(x_0 \rightarrow x_1) \leq \left(1 - \frac{1}{\det(I + \frac{1}{n_e}GD)}\right)$$

where $D$ is a diagonal matrix with entries $D_{ii} = \int_{t_i}^{t_i+T} |e^{j\phi(t)} - e^{j\phi(t)}|^2 dt$, and $I_L$ is the $L \times L$ identity matrix. Further, $\tau = [t_1, t_2, \ldots, t_L]$ is a vector of $L$ symbol positions at which $x_0$ differs from $x_1$. $\Gamma$ denotes an $L \times L$ correlation matrix with $\Gamma_{ij} = E[(\alpha_i - m_{\alpha_i})(\alpha_j - m_{\alpha_j})]$, where $m_{\alpha_i}$ denotes the mean of $\alpha_i$.

Consider the trellis coded CPM system with bit interleaving in Section II. Let us assume that the CPM signal $x_0$ corresponding to the input sequence $a_0$ is transmitted. Let $x_{i,L}$ denote the $i$th CPM signal which differs from $x_0$ in exactly $L$ epochs. If $x_{i,L}$ corresponds to the input sequence $a_{i,L}$, then the difference sequence $e = a_0 \oplus a_{i,L}$ is called the error sequence. Let $W(e)$ denote the input weight of the error sequence. The union bound on the probability of error for correlated flat fading channels is

$$P(e) \leq \sum_{a_0} \sum_{L=1}^{\infty} \sum_{x_{i,L}} \sum_{N} \frac{W(e)}{N} P(x_0 \rightarrow x_{i,L})$$

$$\leq \sum_{a_0} \sum_{L=1}^{\infty} \sum_{x_{i,L}} \sum_{\tau} \sum_{D} \frac{W(e)}{N} P(L, \tau, D)$$

where $P(L, \tau, D)$ denotes the pairwise error probability between two sequences that differ in $L$ time instants, with $\tau$ and $D$ defined earlier, and $N$ is the block length. In general, for CPM schemes $P(x_0 \rightarrow x_{i,L})$ depends on $x_0$ and, hence, the transmitted code word cannot be assumed to be the all-zeros code word. In [11] approximations to evaluate (7) in such situations are discussed. In this paper, we consider only
minimum shift keying (MSK) for which \( P(x_0 \rightarrow x_{1,L}) \) does not depend on \( x_0 \) and, hence, (7) can be shown to be
\[
P_b(c) \leq \sum_{L=L_{\text{min}}}^{\infty} \sum_{\tau} \sum_{D} \frac{A(L, \tau, D)}{N} P(L, \tau, D)
\]
where \( A(L, \tau, D) \) is the average number of CPM signals (or, codewords) which differs from the transmitted sequence in \( L \) epochs, with \( \tau \) and \( D \) defined before. Since the CPE is recursive, with the assumption of a uniform interleaver [4], it can be shown that for small values of \( L \), \( A(L, \tau, D) \propto N^{1-d/2} \) [4],[11]. This suggests that increasing the block length \( N \) results in a drastic reduction of the number of codewords with small diversity \( L \) and, hence, in a large interleaving gain.

V Receiver Structures

V-A Coherent Receiver

The coherent receiver considered is identical to the one in [4],[6]. The receiver performs iterative demodulation and decoding of the received signal using the Bahl, Jelinek, Cocke, and Raviv (BCJR) algorithm for both demodulation and decoding. The receiver assumes that perfect channel state information is available. After each iteration, the decoded data is checked for errors using a cyclic redundancy check (CRC) and if the CRC passes, the iterations are stopped; otherwise, the iterative process continues up to a maximum of 8 iterations.

V-B Non-coherent Receiver

From Sect. IV and Sect. VI we can see that a very significant interleaving gain is possible for the TCCPM scheme and, hence, TCCPM schemes can operate at very low \( E_b/N_o \). At such low \( E_b/N_o \), channel parameter estimation becomes very difficult. Non-coherent receivers do not require channel state information and, hence, are attractive for use at such low \( E_b/N_o \). Iterative non-coherent detection is a rather new topic and there is very little work in this area. An iterative non-coherent receiver structure for iterative demodulation and decoding of DPSK signals in AWGN channels was proposed in [7]. The overall receiver structure is identical to the coherent receiver except that a non-coherent soft-output demodulator is used instead of the BCJR algorithm.

Unlike in the case of coherent reception, the BCJR algorithm cannot be used for optimum soft-output demodulation of the CPM signals. This is mainly due to the fact that the optimum metric for non-coherent demodulation is not additive and, hence, cannot be used in the BCJR algorithm. In general, non-coherent demodulation of CPM signals is accomplished either by using a sub-optimum metric in conjunction with an optimum trellis based algorithm or by using the optimum metric with a sub-optimum trellis based algorithm. One such approach is block detection, where a small window of data is used to generate soft-output instead of the entire data sequence [8]. A sub-optimal iterative non-coherent demodulator using this principle was used in [3] for Rayleigh fading channels. Here, we derive the optimum soft-output non-coherent demodulator for CPM signals for Rayleigh fading channels using block detection. Our approach here is based on non-coherent block detection of CPM signals proposed by Simon and Divsalar in [8].

Let \( x(t, \tilde{c}) \) denote the transmitted CPM signal corresponding to the sequence \( \tilde{c} \). Then, the received signal is given by
\[
r(t) = x(t, \tilde{c}) | \alpha(t)| e^{j \theta(t)} + n(t)
\]
We assume that \( |\alpha(t)| \) and \( \theta(t) \) remain constant over the interval \( kT - N_1T < t < kT + N_2T \). Further, let \( |\alpha_k| \) and \( \theta_k \) denote the magnitude of the channel gain and phase of the channel gain, respectively and \( r^{k+N_2}_{k-N_1} \) the received signal over the interval \( kT - N_1T < t < kT + N_2T \).

To generate the LLRs for \( \tilde{c}_k \), the non-coherent receiver observes the received signal over the window \( (kT - N_1T < t < (kT + N_2T) \) and evaluates the conditional probability \( P(r^{k+N_2}_{k-N_1} | x(t, \tilde{B})) \) for all possible sequences \( B \) of length \( N_1 + N_2 + 1 \), i.e. \( B = [\tilde{c}_k, \tilde{c}_{k+1}, \ldots, \tilde{c}_{k+N_2}] \). In general, for \( M \)-ary CPM there are \( M^{N_1+N_2+1} \) possible vectors \( B \). For MSK, there are \( 2^{N_1+N_2+1} \) values for \( B \) denoted by \( B_i = [\tilde{c}_{i,k}, \tilde{c}_{i,k+1}, \ldots, \tilde{c}_{i,k+N_2}] \) for \( i = 0, 1, \ldots, 2^{N_1+N_2+1} - 1 \). The LLR for bit \( \tilde{c}_k \) is then given by
\[
\text{LLR}(\tilde{c}_k) = \sum_{i=0}^{2^{N_1+N_2+1}-1} \sum_{j,k=1}^{N_2} P(r^{k+N_2}_{k-N_1} | B_i) P_{ap}(B_i) P(\alpha_k) \frac{d(\alpha_k)}{d\theta_k}
\]
The term \( \exp \left\{ \int |r(t)|^2 \right\} \) in the above equation depends only on the received signal and, hence, is a constant. Therefore,

\[
P \left( r_{k+N_1}^{k+N_2} B_1(t), |\alpha_k|, \theta_k \right) = A' \exp \left\{ -\left(N_1 + N_2 + 1\right) \frac{E_s |\alpha_k|^2}{N_o} \right\} \exp \left\{ 2 \frac{N_o}{N_o} |\alpha_k| \cos(\theta_k) \text{Re}(\Delta_k(x_i)) \right\}
\]

where \( \Delta_k(x_i) = \int_{kT-N_1T}^{kT+N_2T} r^*(t)x_i(t)dt \). The phase \( \theta_k \) is uniformly distributed over \( [0, 2\pi] \) and, hence \[8\]

\[
\begin{align*}
P \left( r_{k+N_1}^{k+N_2} |x_i(t)|, |\alpha_k| \right) &= A' \exp \left\{ -\left(N_1 + N_2 + 1\right) \frac{E_s |\alpha_k|^2}{N_o} \right\} I_o \left( 2 \frac{N_o}{N_o} |\alpha_k| |\Delta_k(x_i)| \right)
\end{align*}
\]

To evaluate \( \Delta_k(x_i) \), the phase of the CPM signal \( x_i \) at the \((k-N_1-1)\)th epoch, \( \phi_{k-N_1-1} \), should be known. However, it can be seen that \( \Delta_k(x_i) \) is independent of \( \phi_{k-N_1-1} \) and, hence, \( \phi_{k-N_1-1} \) can be set to zero. Techniques to recursively evaluate \( \Delta_k(x_i) \) can be found in \[8\].

For the flat Rayleigh fading channel, \( (13) \) should be averaged over the pdf of \( |\alpha_k| \) given by \( P(|\alpha_k|) = 2|\alpha_k|e^{-|\alpha_k|^2} \) and, therefore,

\[
P \left( r_{k+N_1}^{k+N_2} |x_i(t)| \right) = A' \int 2|\alpha_k|e^{-\left(\frac{N_1+N_2+1}{N_o}\right) |\alpha_k|^2} \exp \left\{ 2 \frac{N_o}{N_o} |\alpha_k| |\Delta_k(x_i)| \right\} d|\alpha_k|
\]

The above integral can be rewritten as

\[
P \left( r_{k+N_1}^{k+N_2} |x_i(t)| \right) = A'' e^u \int 2|\alpha_k|e^{-\left(\frac{1}{2\sigma^2}\right) |\alpha_k|^2} \exp \left\{ 2 \frac{N_o}{N_o} |\alpha_k| |\Delta_k(x_i)| \right\} d|\alpha_k|
\]

where \( A'' \) is a constant and \( u = \frac{N_1+N_2+1}{2\sigma^2} |\Delta_k(x_i)|^2 \). The integrand in \( (15) \) is a Rician pdf that integrates to unity. Therefore,

\[
P \left( r_{k+N_1}^{k+N_2} |x_i(t)| \right) = A'' \exp \left\{ \left(N_1+N_2+1\right) \frac{E_s}{N_o} \right\} \left(N_2 \left[ 1 + \left(\frac{N_1+N_2+1}{N_o}\right) E_s \right] \right)^{-1}
\]

The constant \( A'' \) cancels out in the evaluation of \( (10) \) and, hence, need not be evaluated. Once \( P(r_{k+N_1}^{k+N_2} |x_i(t)|) \) is evaluated, \( (10) \) can be used to evaluate the LLRs. An important feature of this technique is that \( P(r_{k+N_1}^{k+N_2} |x_i(t)|) \) only needs to be computed during the first iteration. In the following stages, only \( P'(B_i) \) needs to be evaluated and the LLRs updated accordingly. Computing \( P'(B_i) \) at most involves only \( N_1 + N_2 \) multiplications and, therefore, the complexity of this non-coherent receiver is rather low.

\sections{Results and Discussion}

A TCCPM scheme with a rate-1/2 non-recursive convolutional outer code with generator polynomials \( 1 + D^2, 1 + D + D^2 \) and MSK was evaluated through simulations.

The truncated union bound on the bit error rate (BER), the lower bound on the BER, and the performance of a ideal coherent receiver is shown in Fig. 1 for a flat Rayleigh fading channel with \( f_oT = 0.01 \). Results are shown for both block and pseudo-random interleavers. The uniform interleaver assumption is not used with the truncated union bound; rather, the first few terms in the union bound are evaluated with the exact interleaver used. It can be seen that the performance of the coherent iterative receiver is close to that of the bounds developed.

The performance of the coherent iterative receiver and the performance of a non-coherent iterative receiver for the AWGN channel are shown in Fig. 2. Similar curves are shown for the flat Rayleigh fading channel with \( f_oT = 0.01 \) is shown in Fig. 3. It can be seen that the non-coherent receiver with \( N_1 = N_2 = 3, 4 \) performs to within 1 dB of the coherent receiver for the AWGN channel and, within 2-3 dB of the coherent receiver for fading channels. For the same \( N_1 \) and \( N_2 \), the performance loss for the fading channel is higher than that for the AWGN channel. A possible reason is that the channel changes over the observation window length \( N_1 + N_2 + 1 \), which will degrade the performance. The performance of the non-coherent receiver for the fading channel did not appear to improve after the first 2 or 3 iterations. This is mainly due to the fact that only a finite window of \( N_1 + N_2 + 1 \) samples is used to generate the soft-output for each coded bit, as compared to the coherent receiver which uses the entire block to update the soft-output at each iteration. Techniques which use a sub-optimum metric but use the entire block to generate soft-outputs may perform better than the proposed non-coherent receiver, when used in an iterative receiver. This topic is left for further research.

The complexity of the non-coherent receiver may be slightly larger than that of the coherent receiver depending on the actual number of iterations used. However, the coherent receiver requires channel estimation and tracking that must be accounted for when making a fair comparison.

\sections{Conclusions}

We have shown that iterative demodulation and decoding can significantly improve the performance of TCCPM on AWGN and flat fading channels. The optimum iterative non-coherent receiver based on the block detection principle was derived and shown to perform within 1 dB of the performance of the ideal coherent receiver on AWGN channels. The performance on flat Rayleigh fading channels was shown to be within a few dB of that of the ideal coherent receiver.

\sections{References}


Figure 1: Upper bound on BER and performance of coherent receiver on flat Rayleigh fading channel; $f_mT = 0.01; N = 512$.

Figure 2: Comparison of coherent and non-coherent receivers for MSK on AWGN channel; $N = 1024$.

Figure 3: Comparison of coherent and non-coherent receivers; Rayleigh fading, $f_mT = 0.01; N = 512$.