Differentially-Encoded Turbo Coded Modulation with APP Channel Estimation

Sheryl Howard  
Dept. of Electrical Engineering  
University of Utah  
Salt Lake City, UT 84112  
email: s-howard@ee.utah.edu

Christian Schlegel  
Dept. of Electrical Engineering  
University of Alberta  
Edmonton, AB Canada T6G 2G7  
email: schlegel@ee.ualberta.ca

ABSTRACT
A simple serially concatenated turbo code using differential 8PSK encoding as the inner code and a [3,2,2] parity code as the outer code is studied. This system is decoded coherently according to turbo principles with iterative exchange of extrinsic probabilities. Decoding over channels without prior synchronization is demonstrated to be feasible even with significant phase offset, using a simple channel estimator that utilizes the extrinsic output symbol probabilities from the differential APP decoder, which is termed APP channel estimation.

KEY WORDS
differential turbo coded modulation, channel estimation

1 Introduction

Coherent detection assumes perfect knowledge of the carrier phase at the receiver. Under coherent detection, if the received phase is rotated from the transmitted phase due to channel or local oscillator noise, the receiver is unable to decode correctly.

Differential turbo coded modulation concatenates an error-control code with a differential PSK encoder which serves as the inner modulation code. Iterative (turbo) decoding [2], [3] is applied to decode the received signal.

Iterative decoding has been applied to a serial concatenation of an 8-state, rate 2/3 convolutional code and regular 8-PSK modulation, using hard decision decoding with the Viterbi algorithm in [4]. Differential QPSK modulation, concatenated with a convolutional code, has been decoded iteratively with differential demodulation over multiple symbols in [5]. Differential BPSK modulation, and double differential encoding (DDE), serially concatenated with a rate 1/2 convolutional code, is considered in [7] using coherent detection.

Differential modulation has been applied to space-time coding [8] and serial concatenation of error-control coding with excellent performance results in [9].

In this paper, we examine an efficient yet simple serially concatenated system composed of a [3,2,2] parity check code as the outer code and differential 8PSK encoding as the inner code. This system is decoded iteratively according to turbo principles, without using differential demodulation. Both coherent detection and decoding without channel information are considered. A simple channel estimation method utilizing the extrinsic transmitted symbol probabilities from the inner APP decoder is introduced, which we term APP channel estimation.

The paper is organized as follows: Section 2 describes the serially concatenated parity code/differential 8PSK system, from both the encoding and decoding perspective. Section 3 discusses EXIT analysis of this system. Section 4 describes a method of obtaining channel estimates when the system is decoded without channel information, termed APP channel estimation. Section 5 presents simulation results. Conclusions are discussed in Section 6.

2 System Description

Figure 1 displays the encoder. A sequence of information bits \( u \) of length \( L \) is encoded through the [3,2,2] parity code into a sequence of coded bits \( v \) of length \( \frac{2}{3}L \). The coded bits \( v \) are then bitwise interleaved. These interleaved bits \( v' \) are mapped to \( \frac{2}{3} \) 8PSK symbols \( w \), where \( w_k = \sqrt{E_s} e^{j\theta_k} \). The 8PSK symbols \( w \) serve as input to the differential encoder. Differential encoding consists of multiplying the current input symbol \( w_r \) with the previously transmitted symbol \( x_{k-1} \) to obtain the current transmitted symbol \( x_k = x_{k-1} w_k \). Initially, \( x_0 = w_0 \).

The differential 8PSK symbols \( x \) are then transmitted across an AWGN channel with noise variance \( N_0/2 \) in each dimension. Under coherent detection, the received symbols

![Figure 1. Serial Turbo Encoder for Differential Turbo Coded Modulation](image-url)
\[ y \] consist of the transmitted symbols plus complex noise with variance \( N_0 \), \( i.e. \), \( y = x + n \).

We view the differential encoder as an inner code of the serially concatenated system. It can be seen as a recursive non-systematic convolutional code, with a regular, fully-connected 8-state trellis.

Decoding of the serially concatenated system proceeds iteratively according to turbo decoding principles [2], [3]. Figure 2 displays the decoding process, with the APP channel estimation block shown in the dashed rectangle. For now, we consider coherent detection, assuming that we have channel phase synchronization and operate without channel estimation. The received channel symbols \( y \) are converted into channel metrics \( p(y_k|x_k) = \frac{1}{\sqrt{N_0}} e^{-|y_k-x_k|^2/N_0} \) which are fed into APP decoder 1 for the differential code, along with a priori information on the 8PSK symbols \( w \) from APP decoder 2. In the first iteration, APP decoder 1 has no a priori information available from decoder 2, and the a priori values are set to uniform.

\[ \text{Figure 2. Serial Turbo Decoder for Differential Turbo Coded Modulation with APP Channel Estimation} \]

Using the BCJR [12] algorithm, APP decoder 1 calculates symbol probabilities on both the 8PSK symbols \( w \) and the transmitted D8PSK symbols \( x \), and the extrinsic 8PSK symbol probabilities on \( w \) are passed on to APP decoder 2. Extrinsic probabilities are first obtained by dividing out the corresponding a priori symbol probabilities \( P_a(w_k) = e^{(j2\pi y_k/M)} \) and then normalizing such that the extrinsic symbol probabilities sum to one.

\[ P_e(w_k) = \frac{P(w_k)}{P_a(w_k)}, \quad \alpha = \left( \sum_{l=0}^{M-1} P_a(w_k = l) \right)^{-1} \]

These extrinsic symbol probabilities now must be converted to bit probabilities before being deinterleaved, as the interleaver works bitwise. This is done through marginalization. The bit probabilities \( P_e(v') \) are now deinterleaved and fed into APP decoder 2 as a priori bit probabilities \( P_a(v) \). The [3,2,2] parity code is simple enough that its APP decoder can be implemented as 6 equations, giving extrinsic probabilities that express the parity constraints as

\[ P_e(v_1 = 1) \propto P_a(v_2 = 0)P_a(v_3 = 1) + P_a(v_2 = 1)P_a(v_3 = 0) \]

and analogously for \( v_1 = 0, v_2 \) and \( v_3 \).

The bit probabilities are then interleaved to provide a priori bit probabilities \( P_a(v') \), which are converted back to symbol probabilities \( P_a(w) \) for the next iteration of APP decoder 1. The probability of a symbol is simply the normalized product of its component bit probabilities.

Iterative decoding continues in this fashion, with APP decoders 1 and 2 exchanging extrinsic probabilities until convergence is reached.

### 3 EXIT Analysis

Turbo coded systems can be analyzed very elegantly by a method known as EXIT analysis [13], [14]. The reliability of the extrinsic soft information generated by each component decoder is measured by the mutual information \( I(Z;E) \) between the extrinsic information \( E \) and actual symbols \( z \) associated with that soft information.

Likewise, the reliability of the a priori information \( A \) into the same decoder is measured by \( I(Z;A) \). Plots of \( I(Z;A) \) versus \( I(Z;E) \), known as extrinsic information transfer (EXIT) charts, can be used to study the convergence behavior of iterative decoding systems.

Mutual information is unchanged by the interleaving process; interleaving scrambles the symbols but leaves the first order distribution unchanged. Furthermore, the interleaver destroys any correlation between successive symbols. Using this separation assumption, the component decoder EXIT charts may be combined into a single EXIT graph which accurately describes the behavior of the iterative turbo decoding process.

The outer parity decoder produces soft information \( p_e(v) \) and \( p_a(v) \) on the bit level, which are processed as LLRs \( E_2 \) and \( A_2 \). Since the inner differential code operates on 8PSK symbols, \( p_e(w) \) and \( p_a(w) \) must be converted from the interleaved bit probabilities \( p(v') \).

Figure 3 shows the EXIT chart for our system with the differential 8PSK curve as the inner decoder \( I(A_1=I(V';A_1)) \) on the horizontal axis, \( I(E_1=I(V';E_1)) \) on the vertical axis) and the [3,2,2] parity check curve as the outer decoder, with swapped axes. Only the inner decoder EXIT curves depend on SNR.

The significant advantage of EXIT analysis is that the turbo decoder performance near the ‘turbo cliff’ region may be predicted without running simulations of the complete turbo decoder; EXIT transfer curves are obtained for each individual decoder. From Figure 3, we see that the mutual information values for the serially concatenated system, indicated by the trajectory in blue, match well with the predicted individual decoder EXIT curves. At SNR 5 dB, an open iteration channel exists and convergence occurs in 15 iterations. Each vertical-horizontal step indicates one complete iteration of decoding.
4 Decoding Without Channel Information

We now consider the case when the received channel phase is unknown and we decode without channel information. The differential outer code allows the receiver to extract soft information on the symbols \( w \) and \( x \) even in the absence of channel knowledge. This is achieved through the APP decoder, without differential decoding.

APP decoder 1 generates extrinsic input symbol probabilities \( P(w_k) \), as well as extrinsic output symbol probabilities \( P(x_k) \) which will be used to feed a channel estimator for use in the following iteration. This channel estimator must be of low complexity so as not to overwhelm the system and is discussed below. An optimal linear estimator such as the minimum mean square error (MMSE) estimator is therefore not feasible and a simpler filtering estimator is considered [9].

Assuming the channel model to be \( y_k = h_k x_k + n_k \), where \( h_k \) is a complex time-varying gain, we find from the first moment equation that

\[
E[y_k] = h_k E[x_k]
\]

where \( E[x_k] \) is taken to be the expectation over the a posteriori symbol probabilities at time \( k \). A channel estimate may be found as

\[
\hat{h}_k = \frac{y_k x_k^*}{E[x_k]^2}
\]

where \( \hat{x}_k \) is the empirical mean, modified to unit modulus, \( i.e., |\hat{x}| = 1 \), according to \( P(x) \) generated by APP decoder 1. As the a posteriori probabilities \( P(x) \) form the channel estimates \( \hat{h} \), we term this procedure APP channel estimation. Figure 2 shows the iterative decoding process with the APP channel estimation block enclosed in the dashed rectangle. APP decoder 1 sends its extrinsic \( P_e(u) \) to APP decoder 2 and APP 2 generates a priori \( P_a(u) \) for use in the next iteration. The channel estimate \( \hat{h}_k \) is used to calculate coherent channel metrics for APP 1 in the next iteration as

\[
P(y_k|x_k) \propto \exp \left( -|y_k - \hat{h}_k x_k|^2 \right)
\]

Each iteration improves the extrinsic values \( P(x) \) from APP 1, and an improved channel estimate \( \hat{h} \) can be determined at each iteration.

We consider a channel with time-varying phase offset and unity gain, \( i.e., \ h_k = e^{j\phi_k} \). Two different channel phase models are examined: 1) a constant phase offset, and 2) a random walk phase process.

1) For a constant phase offset, \( \phi = e^{j\phi_k} \). The individual APP channel estimates \( \hat{h}_k \) provide phase estimates \( \hat{\phi}_k \), which are averaged to obtain a constant phase estimate \( \phi \).

Figure 5 shows simulation results of the channel phase estimation for a constant phase offset vs. iterations. A phase offset of \( \pi/16 \) radians can be compensated for with our APP channel estimation method in 10 iterations at SNR 4.8 dB; a phase offset of \( \pi/8 \) requires 25 iterations.

A feature of the differential 8-PSK trellis works to our advantage in this estimation process, that is, the rotational invariance of the differential 8-PSK trellis. In the absence of noise, a channel phase offset of \( \pi/4 \) rotates the symbols in the transmitted sequence and thus cyclically permutes each state in the traversed state sequence. This results in incorrectly decoded transmitted symbols \( x \) but correctly decoded 8-PSK symbols \( w \). The system can thus coherently decode any channel phase rotation of an integer multiple of \( \pi/4 \) without channel estimation. Any phase rotation only needs to be corrected through channel phase estimation to the closest integer multiple of \( \pi/4 \) radians with our system. A phase rotation of \( \pi/8 \) modulo \( \pi/4 \) radians will be the most difficult to estimate, as it lies halfway between two valid states.

Fixing the differential trellis to begin and end in state 0 will cause errors at those points for any phase offset. The rest of the trellis shifts to a rotated sequence, but those points are pegged at state 0. Therefore, we use a “floating” trellis, where both beginning and end states are assumed unknown and set to uniform probabilities.

2) The random walk phase process is a Markov process described by \( \phi_k = \phi_{k-1} + \Theta_k + \Phi_{\text{const}}, \) where \( \phi_k \) is the channel phase offset at symbol interval \( k \), \( \Theta_k \) is a zero-mean Gaussian distributed random phase with variance \( \sigma^2_{\Theta} \), and \( \Phi_{\text{const}} \) is a constant channel phase rotation. Initial channel estimates are found as per equation 4. These initial estimates are then filtered through a moving average filter with exponential decay \( \alpha \) to obtain \( \hat{h}_k = (1-\alpha)\sum_{i=0}^{k} e^{-i\alpha} \hat{h}_i \). The filtered channel estimates \( \hat{h} \) are used to calculate improved channel metrics in the next iteration.

---

Figure 3. EXIT chart with trajectory for serial concatenation of [3,2,2] parity check code with differential 8-PSK code at SNR=5 dB.

Decreasing SNR values lower the differential 8PSK EXIT curve. At the turbo cliff, a narrow channel exists between the component code EXIT curves, allowing only minimal error rate improvement per iteration, resulting in a large number of iterations to reach convergence.
The rotational invariance of the differential 8-PSK trellis to multiples of $\pi/4$ phase rotation is displayed in Figure 5, which shows the random walk channel phase at top and the APP channel phase estimate beneath. The phase estimate 'slips' twice to a phase rotated by $-\pi/4$ rads from the actual channel phase. However, there are no decoding errors, even at the phase discontinuities, due to the rotationally invariant trellis.

5 Simulation Results

Simulation results are provided for the [3,2,2] parity code/differential 8PSK modulation under coherent operation and using channel estimation when the channel phase is unknown. Two different channel phase models are simulated: a constant phase offset and a random walk phase process that varies with each symbol.

Figure 6 shows results for the [3,2,2] parity outer code with differential 8PSK modulation as inner code, with coherent decoding. A block length of 10000 information bits is used. An 8-PSK mapping described in [6], designed to increase the minimum distance of this system, which is 1.172 for natural mapping, is used. Natural mapping provides a 0.2 dB advantage in turbo cliff onset, at the cost of a higher error floor [6].

Simulation results for APP channel estimation with a constant channel phase offset of $\pi/16$ rads are shown in Figure 7. Results for a constant channel phase offset of $\pi/10$ rads are shown in Figure 8. Performance degrades somewhat as we approach an offset of $\pi/8$ rads.

Random walk phase results with channel estimation are provided in Figure 9 with $\alpha=.99$ and $\sigma^2 = .05$ rad$^2$. Near-coherent performance is achieved without channel phase knowledge, using APP channel estimation.

6 Conclusions

We have shown that a very simple serially concatenated system consisting of a [3,2,2] parity code as outer code and differential 8PSK modulation serving as the inner code provides very good results when decoded iteratively according
to turbo principles. This system can be very easily encoded and decoded, and could be used in conjunction with packet transmission, where short messages increase the need for phase offset immunity.

Near-coherent performance is achieved when operating without channel information using a simple channel estimation technique that utilizes the extrinsic information available from the inner differential APP decoder.

Acknowledgment:
Many thanks to Lovisa Björklund for generating the data for the EXIT chart, and to Alex Grant for helpful discussions.

References