Iterative Joint Channel Estimation and Decoding in Flat Correlated Rayleigh Fading

Christos Komninakis and Richard D. Wesel

Electrical Engineering Department
University of California, Los Angeles
Los Angeles, CA 90095-1594
{chkomn, wesel}@ee.ucla.edu

Abstract

Performance evaluation of trellis and turbo-codes in flat correlated Rayleigh fading traditionally assumes ideal channel interleaving and availability of perfect Channel State Information (CSI) at the receiver. This paper presents a joint iterative decoding and channel estimation scheme, which performs very well in the flat Rayleigh fading channel with moderate to high Doppler rates, without channel interleaving or external CSI. The decoding algorithm models the phase trajectory of the channel as a Markov process, the dynamics of which depend on the Doppler rate. At each iteration it considers all possible phase trajectories to determine the maximum a posteriori probability (MAP) data sequence. This decoding procedure is a natural extension of standard turbo decoding, and permits decoding without pilot symbols, at the expense of introducing a significant complexity increase.

Keywords: Phase Estimation, Concatenated Coding, Fading Channels.

1 Introduction

The standard coding technique used to mitigate flat correlated Rayleigh fading is trellis coding in conjunction with interleaving, which provides time diversity. There are two main coding paradigms following this approach. First, the codes in [1, 2] rely on interleaving of the transmitted constellation symbols, which places the code design focus on maximizing the minimum symbol-wise Hamming distance, often referred to as the minimum built-in time diversity of the code. In a more recent scheme known as Bit-Interleaved Coded Modulation (BICM) [3, 4], it was recognized that higher code diversity can be obtained by bit-interleaving the encoder outputs prior to mapping them to a constellation point. Thus code design strives to maximize bit-wise Hamming distance.

In both approaches, however, ideal Channel State Information (CSI) is assumed to be available at the decoder, in the form of perfect knowledge of the amplitude and phase of the fading process. For PSK transmission in particular, CSI refers predominantly to the channel phase, since a very good estimate of the channel amplitude is the received amplitude itself. In practice the receiver obtains the CSI either via a decision directed Phase-Locked Loop (PLL), or, if this is undesirable due to block-based transmission (as with turbo-codes) via pilot symbols, periodically inserted in the data stream [5].

Nevertheless, there are limitations to those CSI acquisition techniques. For instance, in the very low SNR region, where turbo-codes work, the phase acquisition task becomes much more difficult for a usual PLL. Furthermore, in a correlated Rayleigh fading channel with moderate or high Doppler, where significant phase variation is present from symbol to symbol, the rate sacrifice, caused by densely injecting pilot symbols in the coded data stream to obtain good channel estimates, quickly becomes prohibitive. Thus, in this paper we take a different approach: the receiver estimates jointly the data and a quantized version of the phase distortion induced by the correlated Rayleigh fading channel. To do that, we introduce a finite-state Markov phase model, which approximates the values of the real-world fading phase with their quantized counterparts, and the dynamics of the phase fading process with a stationary Markov structure, whose transition probabilities depend on the Doppler.

An analogous idea was proposed in [6], where a Markov chain was used to model the SNR variations of the channel, while perfect phase knowledge was assumed. However, in the case of coded PSK treated here, the phase distortion caused by fading poses a much more severe challenge to the receiver than does amplitude fading with ideally phase-coherent demodulation. Our proposed receiver combines the code trellis with the Markov model for the fading phase to form a supertrellis, along which the joint estimation of the data and the channel phase is performed in an iterative fashion. [7] was based on a similar (supertrellis) concept, but for binary Markov channels. In our case, since the data and the channel are being jointly estimated, no explicit CSI and no pilot symbols are required. With the reliability of the correct data increasing with successive iterations, essentially every coded symbol gradually becomes better known and serves somewhat as a conventional pilot symbol.

The remainder of the paper is organized as follows. Section 2 discusses the channel model for correlated Rayleigh fading and the Markov model that approximates it. Section 3 describes the proposed turbo-coding system based on the supertrellis, while Section 4 shows includes simulation results and discussion of the constituent encoder design. Finally, Section 5 concludes the paper.

2 Rayleigh Fading and Markov-phase Model

We consider PSK transmission in a correlated Rayleigh fading channel with Doppler frequency f_D. After demodulation, the received value at time t is:

$$y_t = \alpha_t \cdot x_t + n_t, \tag{1}$$

where x_t is the transmitted M-PSK constellation point, n_t is an i.i.d. (white) complex Gaussian random variable with variance σ^2 per dimension, and α_t is a correlated complex Gaussian, with power normalized to 1. For a continuum of scatterers in the vicinity of the receiver and an omni-directional receiver antenna, the real and imaginary parts X_t and Y_t of the fading process $\alpha_t = X_t + jY_t = |\alpha_t|\, e^{j\phi_t^a}$ are mutually uncorrelated Gaussian processes, each of which has the well-known U-shaped normalized power spectral density [8]:

$$S_{xx}(f) = S_{yy}(f) = \frac{1}{\sqrt{1 - \left(\frac{f}{f_D T}\right)^2}}. \tag{2}$$

Our main objective is to track the phase variations of the channel, so the receiver jointly estimates the data and a quantized version Q_t of the channel phase ϕ_t^a at every time t from the total received angle $\phi_t^r = \phi_t^x + \phi_t^a + \phi_t^*$, as seen in Fig. 1. In this Figure, ϕ_t^a is the fading-induced angle, ϕ_t^x is the transmitted constellation point angle, as the M-PSK trellis code transitions from state c' to c, i.e. $x_t(c' \to c) = 1 \cdot e^{j\phi_t^x}$, and ϕ_t^* denotes the noise-induced additional angle, whose distribution $P(\phi^*)$ is given by the formula:

$$P(\phi^*; \lambda) = \frac{e^{-\lambda^2}}{2\pi} \cdot \left[1 + \sqrt{\pi}\lambda \cos\phi^* e^{(\lambda\cos\phi^*)^2} \mathrm{erfc}(-\lambda\cos\phi^*)\right] \tag{3}$$

where erfc(.) is the complementary error function and the parameter λ of the distribution depends on the fading amplitude: $\lambda = \frac{|a|}{\sigma\sqrt{2}}$.

Now the discrete time Markov chain $\{Q_t\}$ takes values in a finite set $\mathcal{Q} = \{q_0, q_1, \ldots, q_{K-1}\}$ of "quantized phase distortion channel states" with:

$$q_i = \frac{2\pi i}{K},\ i = 0, 1, 2, \ldots, K-1, \tag{4}$$

where, introducing a quantization operator $\Pi(.)$:

$$Q_t = q_i \Leftrightarrow \Pi(\phi_t^a) = q_i \Leftrightarrow \phi_t^a \in \left[q_i - \frac{\pi}{K}, q_i + \frac{\pi}{K}\right),\ i = 0, 1, 2, \ldots, K-1, \tag{5}$$

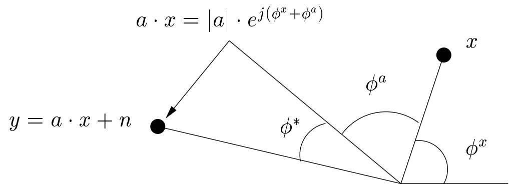

Figure 1: Addition of angles in fading

The transition probabilities $P_{i,j}$, $i,j = 0,1,\dots,K-1$ of the Markov chain are independent of t by stationarity, and can be computed from the joint pdf of two successive sampled fading phases:

$$\begin{aligned} P_{i,j} &= \Pr(Q_{t+T} = q_j \mid Q_t = q_i) &(6)\\ &= \Pr\left(\phi^a_{t+T} \in \left[q_j - \frac{\pi}{K}, q_j + \frac{\pi}{K}\right) \mid \phi^a_t \in \left[q_i - \frac{\pi}{K}, q_i + \frac{\pi}{K}\right)\right) &(7)\\ &= \frac{\int_{q_i-\pi/K}^{q_i+\pi/K} \int_{q_j-\pi/K}^{q_j+\pi/K} p(\phi^a_t, \phi^a_{t+T}) d\phi^a_t d\phi^a_{t+T}}{\int_{q_i-\pi/K}^{q_i+\pi/K} p(\phi^a_t) d\phi^a_t} &(8) \end{aligned}$$

with the marginal pdf being uniform, and the joint pdf [8]:

$$p(\phi^a_t, \phi^a_{t+T}) = \frac{1 - \mathcal{J}_o{}^2(2\pi f_D T)}{4\pi^2} \left[\frac{\sqrt{1-B^2} + B(\pi - B \arccos(B))}{(1-B^2)^{3/2}} \right], \qquad (9)$$

where $B = \mathcal{J}_o(2\pi f_D T) \cdot \cos(\phi^a_{t+T} - \phi^a_t)$.

The model described above is essentially an approximation in a dual sense: First, it maps all real fading angles $\phi^a \in [-\pi, \pi)$ to a finite number of "quantized fading phase states" q_i, $i = 0,1,\dots,K-1$. Moreover, the model approximates the dynamics of the continuous process $\{\phi^a_t\}_{t=0,1,\dots,\infty}$ with a discrete Markov chain, taking values in the finite-state space $\mathcal{Q}$ and having stationary probabilities $p_i = 1/K$ and transition probabilities $P_{i,j}$.

3 Forward-Backward Algorithm on Supertrellis

We combine the structure of the Markov model created in the previous section for the quantized channel phase with the structure of the trellis code to form a supertrellis. At time t, the state S_t of the supertrellis is an ordered pair consisting of the "fading phase state" Q_t and the code state C_t, giving $S_t = (Q_t, C_t) = (q, c) = m$, with $m = 0,1,\dots,2^\nu K - 1$, for a code with ν memory elements. It is important to note here that if reasonable joint estimation of data and channel phase is to be performed, it must hold that:

$$K \geq 2 \cdot M, \qquad (10)$$

where M is the size of the PSK constellation, and preferably (but not necessarily) that K is an integer multiple of M. Clearly, $K < 2M$ will degrade performance, while finer quantization of the channel phase (larger K) will result in better estimation, but will also increase complexity rather dramatically, since the number of states of the supertrellis is $2^\nu K$, for diminishing performance gain. Simulations have shown that the value $K = 2M$, e.g. 8 quantized channel phases for QPSK, represents the most reasonable tradeoff between performance and complexity.

For iterative decoding [9], the crucial quantity to be computed is:

$$\begin{aligned} \gamma_t(m', m) &= \Pr\left(S_t = (q,c), \phi^r_t = \theta \middle| S_{t-1} = (q', c')\right)\\ &= \Pr\left(S_t = (q,c) \middle| S_{t-1} = (q', c')\right) \cdot f\left(\phi^r_t = \theta \middle| S_t = (q,c), S_{t-1} = (q', c')\right). &(11) \end{aligned}$$

For the first RHS term of Eq. (11) we have:

$$\begin{aligned}\Pr\left(S_t=(q,c)\mid S_{t-1}=(q',c')\right) &= \\ &= \Pr\left(\Pi(\phi_t^a)=q\mid \Pi(\phi_{t-1}^a)=q'\right)\cdot \Pr\left(u_t \text{ such that } C_t=c\mid C_{t-1}=c'\right) \quad (12)\\ &= P_{i,j}\cdot P(u_t;I) \quad (13)\end{aligned}$$

where $P_{i,j}$ was derived in (6)-(8), and $P(u_t;I)$ is a constant in the absence of a priori information about the input symbols. However, in the iterative algorithm $P(u_t;I)$ denotes the extrinsic information about u_t provided by the other soft decoder. For the second RHS term of (11):

$$\begin{aligned} f(\theta|\,(q,c),(q',c')) &\stackrel{\text{def}}{=} \Pr\left(\phi_t^r=\theta|\; S_t=(q,c),S_{t-1}=(q',c')\right) \quad (14)\\ &= \Pr\left(\phi_t^x+\phi_t^a+\phi_t^*=\theta|\;\Pi(\phi_t^a)=q,\phi_t^x=\angle x_t(c'\to c)\right) \quad (15)\\ &= \Pr\left(\phi_t^a+\phi_t^*=\theta-\angle x_t(c'\to c)|\;\phi_t^a\sim\mathcal{U}\left[q-\frac{\pi}{K},q+\frac{\pi}{K}\right]\right) \quad (16)\\ &= \frac{K}{2\pi}\cdot\int_{\theta-\angle x(c'\to c)-q-\frac{\pi}{K}}^{\theta-\angle x(c'\to c)-q+\frac{\pi}{K}} p(\phi^*)d\phi^* \quad (17)\end{aligned}$$

where $P(\phi^*)$ was given in Eq. (3). Note that, since the receiver has no knowledge of the true $|a|$, it uses the high SNR approximation $|a_t|\approx|y_t|$ (true for PSK signals) to compute the parameter λ with insignificant performance degradation.

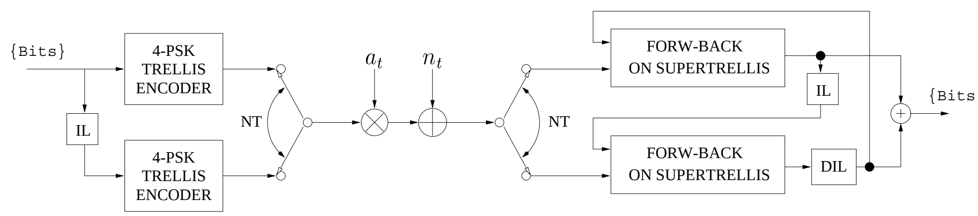

Figure 2: Block diagram of system employing iterative decoder

The system of Fig. 2 is a parallel concatenated (turbo) code, as in [10], with the significant difference that each constituent decoder operates on the supertrellis, running the Forward-Backward algorithm with $\gamma_t(m',m)$ given by Eq. (11), in order to jointly estimate the data and the quantized CSI. For efficiency, an additive version of the algorithm is used, whereby the respective logarithms replace all quantities involved, multiplication becomes summation and summation becomes the max* operation, defined in [11, 9]. In that fashion, the complexity per iteration is roughly twice that of a Viterbi algorithm with $2^\nu K$ states. Observe that this system relies on the joint estimation of the channel and the data, and does not require explicit CSI or pilot symbols. However, if pilot symbols are available, the adaptation of the algorithm outlined above to take pilot symbols into account is straightforward.

The need to preserve the channel correlation in order to estimate the fading phase precludes the use of channel interleaving with our iterative scheme. This leads to lack of diversity. However, we introduce implicit diversity by employing parallel code concatenation [12], having two constituent trellis encoders connected through a uniform symbol interleaver and iterative decoding.

4 Simulation Results

Fig. 3 shows the simulated BER performance of the system of Fig. 2 in a correlated Rayleigh fading channel with a relatively high Doppler rate $f_DT=0.05$. The constituent codes are 8-state, rate-1/2, Gray-labeled QPSK trellis codes, with maximum Hamming distance and thus Euclidean distance also. As a benchmark, we also plot the performance of the same turbo-code enjoying the benefits of ideal channel interleaving (infinite diversity) and perfect CSI. Since the existence of the uniform interleaver inside the turbo-encoder partially compensates for the absence of channel interleaving, we conjecture that the large performance gap between our iterative scheme and the one having ideal CSI is mainly due to lack of accurate CSI and to a lesser extent due to the absence of deep channel interleaving.

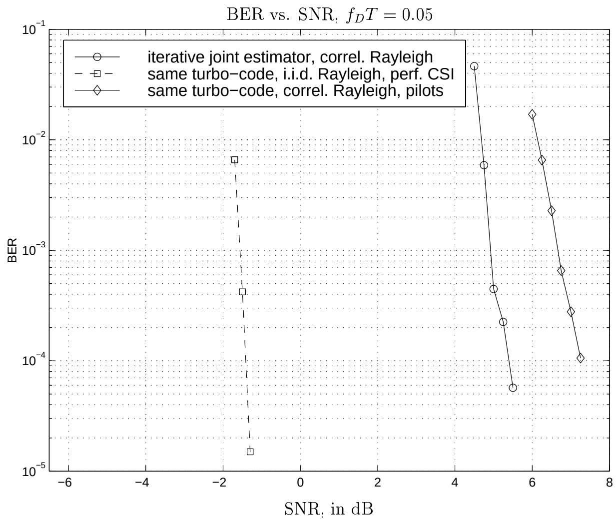

Figure 3: Performance of our iterative estimator compared with that of the same turbo-code with perfect CSI and interleaving (left) and with 3 pilot symbols for every 5 coded symbols (right).

Indeed, to demonstrate the difficulty of obtaining CSI in a practical system operating in moderate to high Doppler rates such as $f_D T = 0.05$, we also simulated a pilot-symbol assisted system (PSAS), using 3 pilot symbols every 5 data symbols. Specifically, we employed a slightly more sophisticated variant of this PSAS, as proposed in [13], in conjunction with the same 8-state turbo-code, without joint estimation. Observe that in this fast fading channel, even with the significant sacrifice in rate by $3/8 = 37.5\%$, the pilot-aided system cannot perform as well as our joint data and channel estimation with no pilots at all. The intuitive explanation for this is that with our joint iterative estimation every data symbol essentially becomes a pilot symbol, as its reliability increases with successive iterations of the turbo decoding algorithm.

Last but not least, a problem that needs to be discussed is the selection of constituent codes for the joint estimation scheme described above. From the coding perspective, the rate-1/2 encoders chosen above are a suboptimal choice, because the systematic bit is repeated twice, and repetition coding is not the most efficient coding possible. However, this choice of low-rate encoders is unavoidable, if we are to maintain reasonable complexity. The rate of the constituent encoders has to be low enough for a given constellation size, so that the joint data and phase estimation algorithm can distinguish whether a change in the received phase is to be attributed to the code or to a change in the channel. For 4-PSK using our current encoders, only two possible outputs are possible at any given time, and one of them is always more likely given the previous channel phase state and the transition probabilities. If, on the other hand, one wishes to have two input bits (i.e. four possible outputs) in each constituent code, then a larger constellation (8-PSK) must be used and it must be $K = 16$, which yields a supertrellis with 128 states. Therefore, although our scheme does not lose rate directly because of pilots that bear no information, the rate reduction that makes channel estimation possible is inherent in the requirements of the constituent encoder design. On a higher level, this can be viewed as incorporating the training in the code design, instead of having higher rate codes and then explicitly injecting pilot symbols in the coded data stream.

5 Conclusions

This paper proposed a PSK turbo-coded system that operates in the flat correlated Rayleigh fading channel and requires no CSI or pilot symbols. The data estimation is performed in an iterative (turbo) fashion, jointly with channel phase estimation. A Markov model for a quantized version of the channel phase was introduced, approximating the non-Markovian dynamics of the channel phase process. Subsequently, this channel phase Markov model was combined with the code trellis to form a supertrellis, along which data decoding and channel estimation are jointly performed. The resulting performance for moderate to high Doppler rates is very good, even superior to that of a system that uses a big number of pilot symbols for channel estimation.

References

[1] D. Divsalar and M. K. Simon. The design of trellis coded mpsk for fading channels: Performance criteria. *IEEE Trans. on Comm.*, 36(9):1004–1012, September 1988.

[2] C.-E. W. Sundberg and N. Seshadri. Coded modulation for fading channels: An overview. *European Transactions on Telecommunications*, 4(3):309–323, May–June 1993.

[3] E. Zehavi. 8-PSK Trellis Codes for a Rayleigh Fading Channel. *IEEE Trans. on Comm.*, pages 873–884, May 1992.

[4] G. Caire, G. Taricco, and E. Biglieri. Bit-Interleaved Coded Modulation. *IEEE Trans. on Inform. Theory*, 44(3):927–945, May 1998.

[5] J. K. Cavers. An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels. *IEEE Trans. on Veh. Tech.*, 40(4):686–693, Nov. 1991.

[6] H. S. Wang and N. Moayeri. Finite-State Markov Channel: A Useful Model for Radio Communication Channels. *IEEE Trans. on Veh. Tech.*, 44(1):163–171, February 1995.

[7] J. Garcia-Frias and J.D. Villasenor. Turbo decoders for Markov channels. *IEEE Communications Letters*, 2(9):257–9, Sept. 1998.

[8] W. C. Y. Lee. *Mobile Communications Engineering*. MacGraw Hill, 1982.

[9] S. Benedetto , D. Divsalar, G. Montorsi, and F. Pollara. Soft-Input Soft-Output Modules for the Construction and Distributed Iterative Decoding of Code Networks. *Europ. Trans. on Telecom.*, 9(2):155–172, March-April 1998.

[10] S. Benedetto , D. Divsalar, G. Montorsi, and F. Pollara. A Soft-Input Soft-Output APP Module for the Iterative Decoding of Concatenated Codes. *IEEE Communications Letters*, 1(1), Jan. 1997.

[11] S. Benedetto and G. Montorsi. Design of Parallel Concatenated Convolutional Codes. *IEEE Trans. on Comm.*, 44(5):591–600, May 1996.

[12] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes (1). In *ICC '93*, pages 1064–1070, 1993.

[13] A. G. Dabak, S. Hosur, and T. Schmidl. Iterative Channel Estimation for Wideband CDMA Systems. *Thirty-Sixth Annual Allerton Conference on Communication, Control and Computing*, pages 441–9, Sep. 1998.