

# Trellis Turbo-Codes in Flat Rayleigh Fading with Diversity

Christos Komninakis and Richard D. Wesel

*Abstract*— This paper compares the performance of two schemes for joint channel estimation and turbo-decoding of high-rate trellis turbo-codes in flat Rayleigh fading, where antenna diversity is available at the receiver. The first method relies on iterative quantized phase estimation, and the second on optimum filtering of pilot and coded symbols. For Doppler rates of practical interest, simulations indicate that optimum pilot filtering is superior to approximating the channel phase with a quantized Markov model. The optimum filtering approach also has lower complexity. For an absolute measure of performance, proximity to channel capacity is also discussed. Iterative joint channel estimation and turbo-decoding with either method is demonstrated to achieve almost all the capacity gain due to receiver antenna diversity.

*Index Terms*— Flat Rayleigh fading, soft diversity combining, turbo-codes,

## I. INTRODUCTION

The iterative exchange of soft information between constituent decoders (“turbo-principle”) in turbo-codes [1] has inspired iterative channel estimation and equalization schemes, such as [2] for flat Rayleigh fading, and [3] for trellis codes in frequency selective channels, and also iterative space-time trellis decoding [4], [5] for systems with multiple transmitter and receiver antennas. In those systems, randomized soft information is exchanged among a generalized equalizer and a Soft-Input Soft Output (SISO) trellis decoder, regarding the channel as a finite-state machine concatenated with that of the code.

In frequency-flat time-correlated fading channels a turbo-code along with some channel estimation algorithm, preferably iterative, can approach capacity. Basic binary Markov modeling of the fading channel is examined in [6], and Rayleigh fading channels with various Doppler rates were discussed in [7], [8] and [9], [2]. Here we consider joint iterative channel estimation and turbo-decoding in flat correlated Rayleigh fading, especially when spatial receiver diversity is available in the form of multiple receiver antennas, as is possible in the uplink of a cellular system.

Specifically, we consider turbo-codes with spectral efficiency of 1 bit/sec/Hz, and show two main ways to integrate the multi-channel estimation and the turbo-decoding processes. Both methods use pilot symbols periodically injected into the coded data stream. The first method is best suited to high rate PSK turbo-codes; it relies on implicit quantization of the fading channel phase and obtains probabilities for every quantized phase using the Forward-Backward algorithm at every iteration [2]. The second relies on Wiener filtering of pilots and (in subsequent iterations) of tentatively decoded symbols, similar

to [7] and [8], and can handle trellis turbo-codes of any constellation. Both methods, henceforth dubbed the “quantized approach” and the “filtering approach” are adapted to handle diversity reception with more than one antennas, whereby the decoder uses soft diversity combining of the available channel estimates.

The filtering approach is similar to [8], but extended to non-binary QAM constellations and employing a diversity combining turbo-decoder. The quantized approach uses the finite-state Markov model in [2] to approximate the values and the statistical properties of the channel phase and compute soft estimates for its quantized version via the Forward-Backward algorithm [10]. With no diversity, this joint estimation and decoding scheme approaches upper bounds to capacity to within 1.3 dB [2]. With spatial diversity, the capacity bounds become complicated computationally, so we assess the performance gain due to diversity in the light of the available capacity gain of a receiver with perfect CSI. This simplistic analysis indicates that both iterative algorithms (quantized phase estimation and Wiener filtering) indeed harvest most of the additional channel capacity available due to diversity.

The organization of the paper is as follows. Section II discusses the transmission and coding schemes and the channel model for pilot-aided turbo-coded transmission in flat correlated Rayleigh fading with receiver diversity. Section III presents a general receiver structure and the two algorithms for joint iterative decoding and channel estimation. Section IV presents the simulated BER performance of the two joint estimation and turbo-decoding schemes, along with the capacity gains available with antenna diversity at the receiver, setting a reference frame for the simulation results. Section V concludes the paper.

## II. TRANSMISSION AND CHANNEL MODEL

To demonstrate joint iterative estimation and decoding of pilot-aided turbo-codes when receiver diversity is available, we use a rate-1 bit/sec/Hz, 4-PSK trellis turbo-code, as in Fig. 1. The constituent encoders are the best 8-state, rate-2/2 code fragments, identified via exhaustive search in [11, Table I], each producing one systematic and one parity bit per 2-bit input, and their outputs are mapped onto a Gray-labeled 4-PSK constellation, following the general turbo-trellis coding paradigm of [11]. Notice that for each block of  $2N$  input bits, the  $N \frac{D+Z}{D}$  pilots and coded symbols form each constituent encoder are transmitted contiguously into the channel; the switch flips every  $N \frac{D+Z}{D} T$  seconds, where  $T$  is the baud period, thus maintaining the correlation necessary for channel estimation at the receiver. Also note the difference between the channel interleaver (CIL), which may have regular block structure, and

The first author is with Broadcom Corp., El Segundo, CA, 90245. The second author is with the Department of Electrical Engineering, University of California at Los Angeles, Box 951594, Los Angeles, CA 90095-1594.

E-mail: {chkomn, wesel}@ee.ucla.edu.

Supported by NSF grants CCR-9732376 and ECS-9820765, NSF CAREER award CCR-9733089, Texas Instruments, and the Xetron Corporation.

the random spread interleaver with parameters (30, 30) of the turbo-code (TIL).

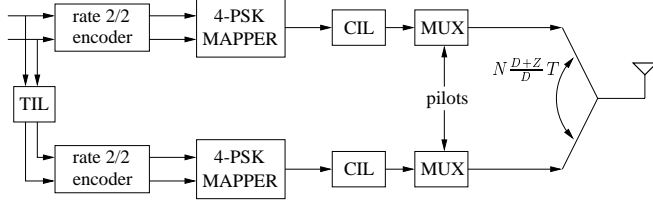


Fig. 1. Transmitter block diagram in a pilot-aided trellis turbo-coded system.

The channel with one transmitter and  $L \geq 1$  receiver antennas produces independent realizations of the fading coefficients and the noise at each of the receiver branches, as shown in Fig. 2. Each of the diversity channels is modeled as an inde-

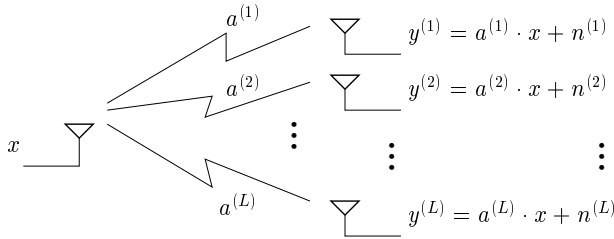


Fig. 2. Flat fading channel with  $L$  branches of receiver diversity.

pendent non-dispersive (flat) Rayleigh fading channel, correlated in time, based on Clarke’s fading model [12]. According to this model, after matched filtering and proper sampling, the discrete representation of the received signal at time  $t$  and diversity branch  $i$ ,  $i = 1, \dots, L$ , is:

$$y_t^{(i)} = a_t^{(i)} \cdot x_t + n_t^{(i)}, \quad i = 1, \dots, L, \quad t = 0, 1, 2, \dots \quad (1)$$

where  $x_t$  is the transmitted constellation point. The channel coefficients and the noise are independent in the different channels, but the fading in each channel remains time-correlated:

$$E a_{t_1}^{(i)} [a_{t_2}^{(j)}]^* = \mathcal{J}_0(2\pi f_D T |t_1 - t_2|) \cdot \delta(i - j) \quad (2)$$

$$E n_{t_1}^{(i)} [n_{t_2}^{(j)}]^* = N_o \cdot \delta(t_1 - t_2) \cdot \delta(i - j) \quad (3)$$

and the noise process has variance  $\sigma^2 = N_o/2$  per dimension. Each time-correlated channel fading process  $\{a_t^{(i)}\}$ ,  $i = 1, \dots, L$  is modeled as a zero-mean circular complex Gaussian random process. The marginal distributions of  $|a_t^{(i)}|$  and  $\phi_t^{a^{(i)}} = \angle a_t^{(i)}$  for each  $i$  are Rayleigh and uniform respectively [13], hence the term “Rayleigh fading”. The correlation properties of each fading processes  $a_t^{(i)}$ ,  $i = 1, \dots, L$ —see (2)—depend on the common Doppler rate  $f_D T$  for all receiver branches, since they are probably mounted on the same physical device and move simultaneously. The working assumption is that the scattering environment in the vicinity of the receiver antennas is rich, such that adequate separation of the antennas provides independent Rayleigh channels in each.

### III. GENERAL ITERATIVE RECEIVER

Given the turbo-coded, pilot-aided transmission model of Fig. 1, the block diagram of a generic receiver is depicted in Fig. 3. By performing joint multi-channel estimation and turbo-decoding, this receiver simultaneously exploits antenna diversity and time-diversity due to the turbo-code. Each constituent decoder iteratively performs “soft diversity combining” of  $L$  soft channel estimates.

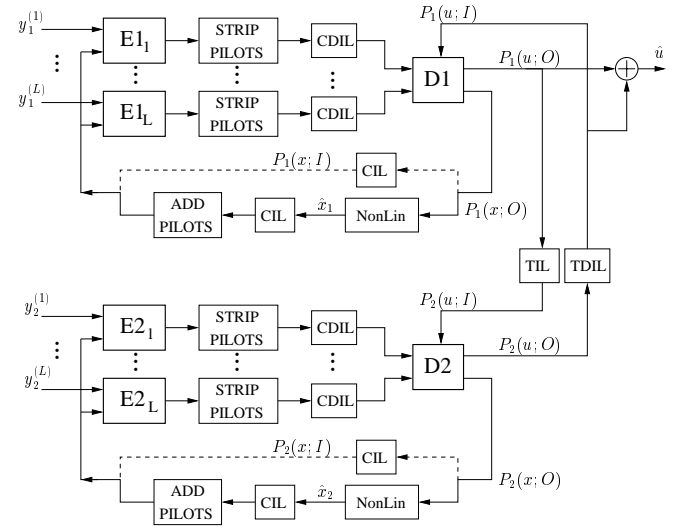


Fig. 3. Receiver block diagram for turbo-decoder with diversity and iterative channel estimation.

The blocks “D1” and “D2” are the constituent decoders (SISOs in the nomenclature of [14]) for the turbo-code, and they implement the Forward-Backward algorithm [10]. They exchange extrinsic information  $(P(u; I), P(u; O))$  about the 2-bit input symbols  $u$  through the uniform random interleaver/deinterleaver pair TIL/TDIL. At the same time they produce extrinsic information  $P(x; O)$  about the coded symbols  $x$ , in order to assist the channel estimation procedure on all diversity branches. Estimation is performed by the two arrays of identical modules  $E1_i, E2_i$ ,  $i = 1, \dots, L$ . For the functionality of the estimation blocks, this paper considers two options.

In the first approach, henceforth dubbed the “quantized approach”, (dashed feedback path in Fig. 3) no operation other than channel deinterleaving is performed on the probability vectors  $P(x; O)$  before they are fed back into the estimator modules. They, in turn, produce probability vectors about the quantized channel phase, in exactly the same fashion as the Q-SISO blocks described in [2]. In the second iterative estimation algorithm, the “filtering approach”, (solid feedback line in Fig. 3) the vector of soft information  $P(x; O)$ —each entry of this vector is the extrinsic probability of a possible constellation point  $x$ —passes through the nonlinearity “NonLin”, which provides a coded symbol estimate  $\hat{x}$ . Then, those estimates  $\hat{x}$  are deinterleaved and pilot symbols are injected into the stream, replicating the procedure followed at the transmit-

ter. Then, this stream of estimated coded symbols  $x$  and pilots is used by the estimator modules, which perform optimal filtering in order to produce channel estimates  $\hat{a}^{(i)}$ ,  $i = 1, \dots, L$  to be used by the turbo-decoders in the next iteration, in a fashion similar to the single channel work in [7], [8].

Both algorithms are explained in detail below, and issues such as estimation in the first iteration, or the nature of the non-linear operation “NonLin” are discussed. Description of both algorithms amounts to analyzing the functionality of the blocks denoted  $E1_i, E2_i$ ,  $i = 1, \dots, L$  in Fig. 3 for both the “quantized approach” and the “filtering approach”, as well as the corresponding metric used by the constituent decoders “D1” and “D2” in conjunction with each channel estimation method.

#### A. Quantized Phase Approach

For PSK transmission the acquisition of phase coherence can be a more critical problem than fading amplitude estimation. Thus, the quantized phase algorithm invests most of the computational effort into obtaining an accurate probability distribution  $P_i(q; O)$  on the quantized phases  $q = Q(\phi^{a^{(i)}})$  of each diversity channel  $i = 1, \dots, L$  via the Forward-Backward algorithm. Thus, each estimation module  $E1_i, E2_i$  of Fig. 3 becomes a Q-SISO, described in [2]. For an estimate of the fading amplitude  $|a^{(i)}|$  each Q-SISO uses a simple, symbol-by-symbol optimum affine estimator from the received amplitude, namely:

$$|\widehat{a^{(i)}}| = A \cdot |y^{(i)}| + B, \quad i = 1, \dots, L \quad (4)$$

where  $A, B$  are real coefficients depending on the SNR (hence the same for all  $i$ ), and they are computed to ensure unbiasedness and MMSE of the amplitude estimates.

For the phase estimation, each Q-SISO considers the quantized version of the channel phase  $\phi^{a^{(i)}}$  into  $K$  phase intervals. All the Q-SISOs (there are  $2L$  of them in the receiver) create the same Markov model, which approximates the values and the statistical properties of each  $\phi^{a^{(i)}}$  [2]. The transition probabilities  $P(q' \rightarrow q)$  between phase sectors depend on the common Doppler rate  $f_D T$ , and can be precomputed. Each Q-SISO runs the Forward-Backward algorithm on the trellis of the quantized Markovian channel phases, and produces probabilities  $P_i(q; O)$  for those quantized phases  $q$ , on each diversity channel  $i$ ,  $i = 1, \dots, L$ . To do that, the  $i^{\text{th}}$  Q-SISO needs to compute the branch metric  $\gamma_t^{(i)}(q', q)$  as follows:

$$\gamma_t^{(i)}(q', q) \stackrel{\text{def}}{=} \Pr\left(y_t^{(i)} = \left|y_t^{(i)}\right| e^{j\theta^{(i)}}, q_t^{(i)} = q \mid q_{t-1}^{(i)} = q'\right) \quad (5)$$

$$= P(q' \rightarrow q) \cdot \sum_x P(x; I) \cdot P\left(y_t^{(i)} \mid x_t = x, q_t^{(i)} = q\right) \quad (6)$$

So, using the precomputed Markovian transition probabilities  $P(q' \rightarrow q)$ , as well as the extrinsics  $P(x; I)$  fed back by the constituent turbo-decoder “D1” or “D2”, each Q-SISO provides the desired extrinsic information  $P_i(q; O)$  about the

quantized channel phase in each diversity branch, as shown in Fig. 3.

At the first iteration, when no prior  $P(x; I)$  is available, all inputs  $P(x; I)$  for coded symbols to the Q-SISOs are set to zero (in the log domain), denoting equal probabilities. For points  $t$  in time belonging to a pilot slot, each Q-SISO uses the most skewed pmf, setting the probability of the pilot symbol to a very large number (certainty) and the others to zero. Thus, for pilots, the summation in the Q-SISO is trivial (only one possible  $x$  has non-zero probability). At times  $t$  corresponding to coded data (not pilots) the summation of (5) is necessary to compute the correct branch metric  $\gamma_t^{(i)}(q', q)$  and run the Forward-Backward algorithm in each Q-SISO.

At every time  $t$ , each constituent decoder (“D1” or “D2” in Fig. 3) receives from the  $L$  Q-SISOs after deinterleaving the  $L$  vectors  $P_i(q; I)$ ,  $i = 1, \dots, L$ , each of length  $K$ —the number of quantized phase states. These  $K$ -entry vectors  $P_i(q; I)$  contain soft information about the quantized channel phase  $q^{(i)}$  on each of the diversity branches, so denote their entries:

$$P\left(q^{(i)}; I\right), \quad q^{(i)} = 0, \dots, K-1, \quad i = 1, \dots, L. \quad (7)$$

Let boldface letters denote the  $L$ -length vectors of the received (deinterleaved) samples and the quantized phases:  $\mathbf{y}_t = [y_t^{(1)} \dots y_t^{(L)}]$  and  $\mathbf{q}_t = [q_t^{(1)} \dots q_t^{(L)}]$ . Then, the “vector phase-state”  $\mathbf{q}_t$  of the  $L$ -branch diversity channel can take one of any  $K^L$  possible values. From the independence between diversity channels, the probability of the channel vector-state is  $P(\mathbf{q}_t = \mathbf{q}) = \prod_{i=1}^L P(q^{(i)}; I)$ .

To run the Forward-Backward algorithm on the code trellis, each constituent decoder computes the branch metric  $\gamma_t(c', c)$ , based on the vector observations  $\mathbf{y}_t$ , and using the *soft* channel estimates (i.e., probabilities that the channel phase is in one of  $K$  sectors). This is effectively soft-diversity combining:

$$\gamma_t(c', c) = \Pr(\mathbf{y}_t, C_t = c \mid C_{t-1} = c') \quad (8)$$

$$= P(u; I) \cdot \sum_{\mathbf{q}} P(\mathbf{q}_t = \mathbf{q}) \cdot P(\mathbf{y}_t \mid x(c' \rightarrow c), \mathbf{q}) \quad (9)$$

$$= P(u; I) \cdot \sum_{q^{(1)}, \dots, q^{(L)}} \prod_{i=1}^L P(q^{(i)}; I) \cdot P\left(y_t^{(i)} \mid x(c' \rightarrow c), q^{(i)}\right)$$

Of the terms in the above formula,  $P(u; I)$  is the extrinsic information about the input symbol  $u$  and is provided by the *other* constituent decoder, while  $P(q^{(i)}; I)$  is provided by the  $i^{\text{th}}$  Q-SISO. The summation involves  $K^L$  distinct terms, each of which is the product of two terms. In both the above equation and in (6) the likelihood of the received value  $y_t^{(i)}$  on each diversity channel given the transmitted  $x$  and the fading sector  $q$  on that diversity branch is:

$$P\left(y_t^{(i)} \mid x_t = x, q\right) = \frac{K}{2\pi} \cdot \int_{\theta^{(i)} - \mathcal{L}_{x-q} - \frac{\pi}{K}}^{\theta^{(i)} - \mathcal{L}_{x-q} + \frac{\pi}{K}} P(\phi^*; \lambda_t^{(i)}) d\phi^*,$$

where  $P(\phi^*; \lambda)$  is the noise angle distribution given in [2] and we use the amplitude approximation of (4) to estimate the parameter  $\lambda$  as  $\widehat{\lambda}_t^{(i)} = \frac{A|y_t^{(i)}|+B}{\sigma\sqrt{2}}$  from the received amplitude.

Due to the soft combining of channel estimates, the complexity of the quantized approach is quite high. For  $L = 2$  diversity branches at the receiver, every full iteration of the joint estimator and turbo-decoder involves  $2L = 4$  Forward-Backward runs on the  $K$ -state Q-SISOs (8 states here) with branch metrics computed as in (5)-(6), as well as two Forward-Backward runs on the 8-state constituent decoders ‘‘D1’’ and ‘‘D2’’ with branch metrics computed as in (8). Note that multiplications in the above formulas become summations in the log-domain and summations become the  $\max^*$  operation, defined in [14]. However, complexity is still an issue, so we limit the number of quantized phases to  $K = 8$  (twice the cardinality of the 4-PSK constellation). The performance of this receiver is presented in section IV.

### B. Optimum Filtering Approach

This section describes iterative joint channel estimation and turbo decoding, based on optimum filtering of received symbols (only pilots at the first iteration, coded symbols and pilots alike in subsequent iterations). Iterative filtering solutions have been proposed in [7] and in [8] for low-rate turbo-codes with BPSK constellations. The treatment here is general, and in fact can also cover non-PSK constellations, unlike the quantized phase approach in section III-A.

With the filtering approach, the solid feedback lines in Fig. 3 are active. The channel estimation modules  $E1_i, E2_i, i = 1, \dots, L$  obtain estimates for the complex channel gains  $a_1^{(i)}, a_2^{(i)}, i = 1, \dots, L$  via Wiener filtering, and provide those estimates (after deinterleaving) to the appropriate constituent decoder ‘‘D1’’ or ‘‘D2’’. Intuitively, as decoder iterations proceed and knowledge of the coded symbols  $x$  becomes more reliable, the channel estimators provide improved estimates of the channel gains  $a^{(i)}$ , in turn contributing to a better performance by the constituent decoders.

In the first iteration, when no knowledge about the transmitted symbols  $x$ , other than the pilots, is available, the estimators  $E1_i, E2_i, i = 1, \dots, L$  are pilot-only filters (POFs, following [8]). These FIR filters are identical for all  $2L$  diversity branches, and are derived as in Appendix A of [15], given the filter length, the Doppler rate  $f_D T$ , the number  $D$  of coded symbols between pilot slots, the number  $Z$  of pilot symbols per slot, and the SNR. So, the  $D$  distinct POFs  $W_k^o$  in (A.2) of [15] (where  $k = 0, 1, \dots, D - 1$  refers to the time index within the coded data slot between two adjacent pilot slots that the filter  $W_k^o$  is intended for) are computed once, at the beginning of the simulation, and subsequently used for the first iteration of every turbo-coded block, by the  $2L$  estimators. They filter received samples corresponding to pilots only (denote them  $\tilde{y}^{(i)}$  after derotation of  $y^{(i)}$  with the known pilot symbol  $x$ ) and produce channel estimates  $a_1^{(i)}, i = 1, \dots, L$

(and  $a_2^{(i)}, i = 1, \dots, L$ ) to be used by the constituent decoders ‘‘D1’’ (and ‘‘D2’’ respectively).

Now, the constituent decoder ‘‘D1’’ (the same holds for ‘‘D2’’, so we suppress the subscripts) computes the metric  $\gamma_t(c', c)$ , using the channel estimates  $a^{(i)}, i = 1, \dots, L$ , after deinterleaving. The optimal way to utilize those  $L$  channel estimates is to perform maximal ratio combining of the  $L$  received samples  $y_t^{(i)}, i = 1, \dots, L$ . Collecting those in the vector  $\mathbf{y}_t$ :

$$\begin{aligned} \gamma_t(c', c) &= \Pr(\mathbf{y}_t, C_t = c | C_{t-1} = c') & (10) \\ &= P(u_t \text{ such that } c' \rightarrow c) \cdot P(\mathbf{y}_t | \hat{\mathbf{a}}_t, x(c' \rightarrow c)) \\ &= P(u; I) \cdot (-G_1) \cdot \|\mathbf{y}_t - \hat{\mathbf{a}}_t x(c' \rightarrow c)\|^2 \\ &= P(u; I) \cdot G_2 \cdot (2\mathcal{R}\{\mathbf{y}_t^* \hat{\mathbf{a}}_t x(c' \rightarrow c)\} - \|\hat{\mathbf{a}}_t\|^2 |x(c' \rightarrow c)|^2), & (11) \end{aligned}$$

where  $\hat{\mathbf{a}}_t$  is the  $L$ -length vector comprising the estimates  $a^{(i)}, i = 1, \dots, L$ , and  $\mathcal{R}\{\cdot\}$  denotes the real part, and  $G_1, G_2$  are irrelevant nonnegative constants. Note that in the case of PSK transmission, where all symbols  $x$  are of equal energy, the metric (11) can take the simpler form:

$$\gamma_t(c', c) = P(u; I) \cdot \mathcal{R}\{\mathbf{y}_t^* \hat{\mathbf{a}}_t x(c' \rightarrow c)\}. \quad (12)$$

The inner product  $\mathbf{y}_t^* \hat{\mathbf{a}}_t$  in (12) involves  $L$  complex multiplications, so the decoder metric for the iterative filtering approach is much simpler than the one in (8) for the ‘‘quantized approach’’.

However, some additional complexity ensues in the filtering approach after the first iteration. The estimation modules  $E1_i, E2_i, i = 1, \dots, L$  need estimates of the coded symbols (call those estimates  $\hat{x}$ ) and cannot exploit the vectors  $P(x; O)$  of extrinsic information about  $x$  provided by the decoders ‘‘D1’’ and ‘‘D2’’. This is why the solid feedback paths of Fig. 3 include the nonlinearity ‘‘NonLin’’, which uses the probability vector  $P(x; O)$  for every  $t$  and provides a hard estimate  $\hat{x}$ . The filters  $E1_i, E2_i$  use  $\hat{x}$  to improve their channel estimates in the next iteration.

An immediate way to obtain  $\hat{x}$  is to simply choose as  $\hat{x}$  the constellation point  $x_m$  with the maximum  $P(x; O)$ , or:

$$\begin{aligned} \hat{x} = x_m &\iff & (13) \\ P(x_m; O) &\geq P(x_l; O), \quad l = 0, \dots, M - 1, \quad l \neq m, \end{aligned}$$

where  $M$  is the constellation size. We refer to nonlinearity (13) as hard tentative decoding, similar to [7]. A more sophisticated way to obtain  $\hat{x}$  uses the soft information in  $P(x; O)$  to obtain the posterior average of  $x$ , similar to [8] and [7] for BPSK. Recall that the entries of the vector  $P(x; O)$  at each time are the log-probabilities of every constellation point  $x$ . Hence:

$$\hat{x} = E[x] = \sum_{l=0}^{M-1} \Pr(x_l) \cdot x_l = \sum_{l=0}^{M-1} e^{P(x_l; O)} \cdot x_l. \quad (14)$$

In the simulations results presented in the next section we use the nonlinearity of (14), but the performance degradation if the simpler (13) is used is very small.

After the estimates for the coded symbols  $\hat{x}$  have been produced for the whole block, the receiver replicates the operations of the transmitter, by interleaving them and adding pilot symbols (refer to Fig. 3, solid feedback path). So, at the input of the estimation blocks  $E1_i, E2_i, i = 1, \dots, L$ , we now have an approximated replica of the transmitted sequence. In all iterations but the first, the estimators use this sequence to derotate the received samples  $y_1^{(i)}, y_2^{(i)}, i = 1, \dots, L$ , so that Wiener filtering of the derotated  $\tilde{y}_1^{(i)}, \tilde{y}_2^{(i)}$  can follow, to produce new channel estimates  $\hat{a}_1^{(i)}, \hat{a}_2^{(i)}$ . We call each of those Wiener filters—they are the same on all  $2L$  diversity branches—an “all-symbol filter” (ASF, as in [8]), to distinguish it from the pilot-only filters (POFs)  $W_k^o$ , used only at the first iteration.

The ASF  $Z^o$  operates for every time  $t = 0, \dots, N - 1$  and at every diversity branch to yield updated channel estimates:

$$\hat{a}_t^{(i)} = Z^o \cdot [\tilde{y}_{t-L_1}^{(i)} \tilde{y}_{t-L_1+1}^{(i)} \dots \tilde{y}_{t+L_1}^{(i)}]^T, \quad i = 1, \dots, L,$$

which are subsequently deinterleaved and provided to decoder “D1” or “D2” for the next algorithm iteration.

The complexity of the “filtering approach” is significantly less than that of the quantized approach for reasonable lengths of the POFs ( $2L_0Z$ ) and the ASF ( $2L_1 + 1$ ). But the receiver structure of Fig. 3 remains valid and describes both algorithms on a high level. Here, the estimator blocks  $E1_i, E2_i, i = 1, \dots, L$  are FIR filters and not Forward-Backward algorithms, so there are only two Forward-Backward algorithms to run per iteration, those of the constituent decoders.

#### IV. SIMULATION RESULTS

This section presents the simulated BER performance of the “quantized approach” presented in section III-A and the “filtering approach” of section III-B in a flat correlated Rayleigh fading channel with diversity order of  $L = 2$ . The turbo-coding scheme used throughout is the same, as shown in Fig. 1. The blocklength is  $N = 4100$ , and pilots are injected into the coded data stream at a rate of  $Z = 1$  pilot every  $D$  coded 4-PSK symbols. The generic diagram of the iterative diversity receiver is that of Fig. 3, with the estimation blocks  $E1_i, E2_i, i = 1, 2$ , and the decoders “D1” and “D2” as described previously. Figs. 4 and 5 show two groups of simulations for Doppler rates  $f_D T = 0.01$  and  $f_D T = 0.05$  respectively.

We observe that for both Doppler rates, the filtering approach provides better performance than the quantized approach by more than 1 dB, and is also less intensive computationally, as pointed out before. For the quantized approach we used a Markov model for the channel phase with  $K = 8$  phase states, while for the filtering approach the simulation parameters were  $L_0 = 5$  pilots on either side for the POF (i.e. total of 10 pilot symbols) for both Doppler rates, and  $L_1 = 16$  for

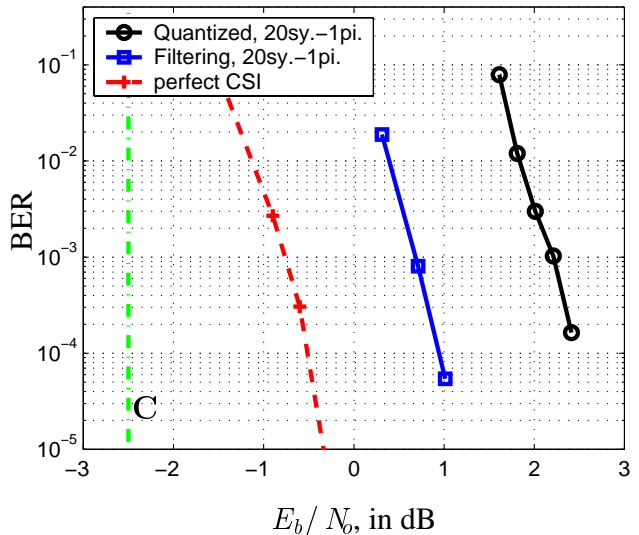


Fig. 4. BER with  $f_D T = 0.01$  for the two channel estimation methods, and given perfect CSI. The vertical line marks the capacity given perfect CSI.

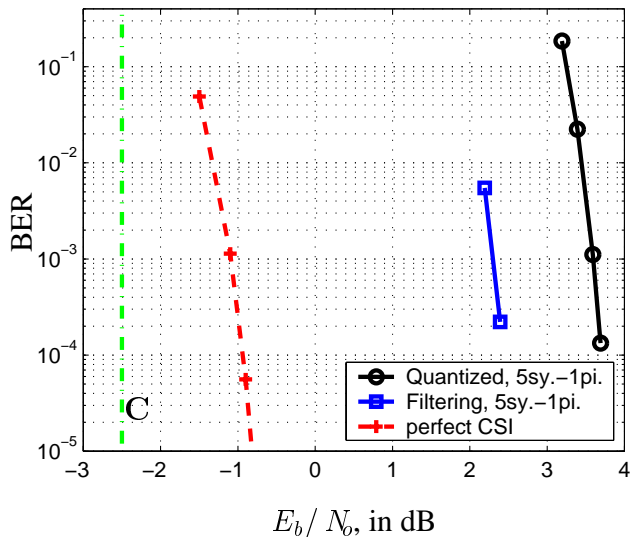


Fig. 5. BER with  $f_D T = 0.05$  for the two channel estimation methods, and given perfect CSI. The vertical line marks the capacity given perfect CSI.

$f_D T = 0.01$  and  $L_1 = 10$  for  $f_D T = 0.05$  for the ASF (i.e. totals of 33 and 21 filtered symbols respectively). The number of iterations was 12 for  $f_D T = 0.01$ , and 10 for  $f_D T = 0.05$ , for both estimation methods, so the superior performance of the filtering approach in both Doppler rates comes with reasonable complexity.

In terms of SNR per information bit ( $E_b/N_0$ ) the performance of both estimation methods (quantized and filtering) degrades for increasing Doppler rate, as seen in Figs. 4 and 5. This happens despite the fact that perfect CSI performance itself improves with increasing Doppler due to increased time-

diversity, and indicates that the channel estimation problem becomes harder for increasing Doppler [2], which makes sense intuitively. In the limit of a channel fading independently, there is unlimited time-diversity to harvest given “genie-provided” CSI, but the estimation task becomes impossible in the absence of any time-correlation in the channel. This explains the observation that the SNR gap between iterative estimation performance and performance achieved with perfect CSI at the receiver increases with increasing Doppler rate (e.g., with filtering, this SNR gap is 1.5 dB for  $f_D T = 0.01$ , and 3.2 dB for  $f_D T = 0.05$ ).

### A. Capacity with Receiver Diversity

This section discusses channel capacity with spatial diversity at the receiver. The results indicate that both algorithms discussed above indeed exploit the SNR advantage offered by diversity, and gain almost as much SNR because of it (with respect to operation without diversity) as the capacity improvement is. To keep the capacity computation tractable, we restrict ourselves to time-independent fading (the assumption of independence across the  $L$  diversity channels remains) and we assume perfect knowledge of the fading gains at the receiver. Thus, collecting all received values into a vector of length  $L$ :

$$\mathbf{y} = \mathbf{a} \cdot x + \mathbf{n}, \quad (15)$$

where all entries of vectors  $\mathbf{a}$  and  $\mathbf{n}$  are assumed independent of each other and in time. With no time constraint and given perfect channel knowledge at the receiver, the capacity of this vector channel is (see [16]):

$$C_L = \int \log_2(1 + \rho \mathbf{a} \mathbf{a}^*) P_{\mathbf{a}}(\mathbf{a}) d\mathbf{a}, \quad (16)$$

where  $P_{\mathbf{a}}$  is the distribution of the fading vector  $\mathbf{a}$ , and  $\rho = P_x / N_o$ , with  $P_x$  the average power of the transmitted constellation point  $x$ . For instance, in the case of no fading ( $L$ -branch AWGN diversity) where  $\mathbf{a} = \mathbf{1}_L$  always, the SNR gain from diversity of order  $L = 2$  is 3 dB.

In Rayleigh fading, each element (fading scale factor)  $a_i$  in the vector  $\mathbf{a}$  is a zero-mean, unit variance complex Gaussian random variable. Hence, the SNR advantage due to diversity will be greater than 3 dB for  $L = 2$ , because of the “sphere-hardening” phenomenon. In other words, the  $L$ -fold receiver diversity helps in a dual fashion: it captures more of the transmitter power and stabilizes against channel spatial fluctuations [16]. Writing the capacity of (16) with Rayleigh fading as

$$C_L = \int \log_2 \left( 1 + \rho \sum_{i=1}^L |a_i|^2 \right) P(|a_1|, \dots, |a_L|) d|a_1| \cdots d|a_L|,$$

the argument of the logarithm points to the optimum “maximal-ratio combiner”. The SNR difference between  $L = 1$  and  $L = 2$  is greater than 3 dB for every  $\rho$ , but higher  $L$  only provides diminishing returns in SNR gain. In particular, at the

nominal rate of 1 bit/sec/Hz, where the turbo-code of this paper operates (excluding pilots) the diversity gain in capacity is 3.5 dB for  $L = 2$ . For ease of comparison, Figs. 6 and 7 include results for Doppler rates  $f_D T = 0.01$  and  $f_D T = 0.05$  respectively, with diversity orders of  $L = 1$  (dashed curves) and  $L = 2$  (solid curves) of both receivers, along with the capacity vertical lines. It is clear that both iterative estimation algorithms benefit from the existence of 2-fold diversity approximately as much ( $\approx 3.5$  dB) as the above capacity result indicates. In other words, both Figs. 6 and 7 show that the second branch of spatial receiver diversity improves the performance of both the “quantized” and the “filtering” iterative receivers by approximately the same amount that it enhances the channel capacity under perfect CSI and ideal interleaving. The filtering approach maintains its superiority over the quantized phase method both with and without diversity, as is obvious from Figs. 6 and 7.

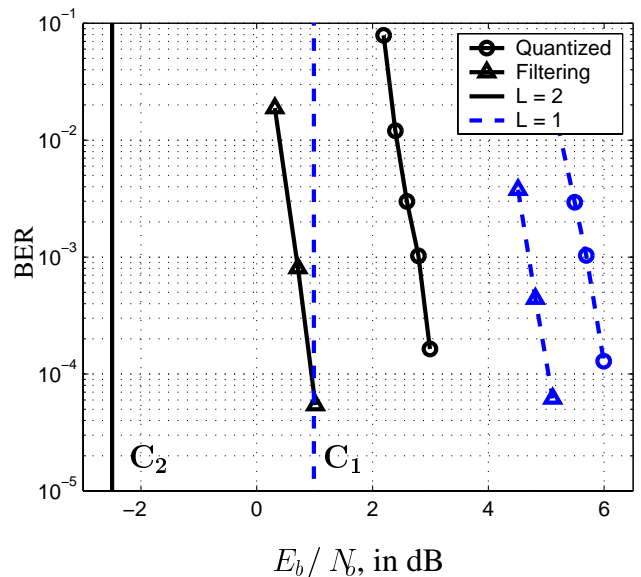


Fig. 6. BER performance of turbo-code in flat Rayleigh fading with  $f_D T = 0.01$ , with diversity orders of  $L = 2$  (solid curves) and  $L = 1$  (dashed curves, no diversity). The “quantized” approach for estimation is marked with “o”, while the “filtering” approach with “ $\Delta$ ”. Vertical lines (solid and dashed) mark the respective capacities with Gaussian inputs and perfect CSI, for  $L = 2$  and  $L = 1$ . Notice that pairs of similarly marked solid and dashed curves maintain an SNR distance of  $\approx 3.5$  dB, or equivalently, that both quantized and filtering simulations improve with diversity about as much as capacity does.

## V. CONCLUSIONS

This paper discussed the problem of joint channel estimation and turbo-decoding in frequency-flat, time-correlated Rayleigh fading, with receiver antenna diversity, but without channel state information at the receiver. Two methods for joint channel estimation and turbo-decoding were analyzed. The first constructs a finite-state Markov model for the fading channel phase and performs quantized phase estimation along the trellis implied by the model. The second method relies on opti-

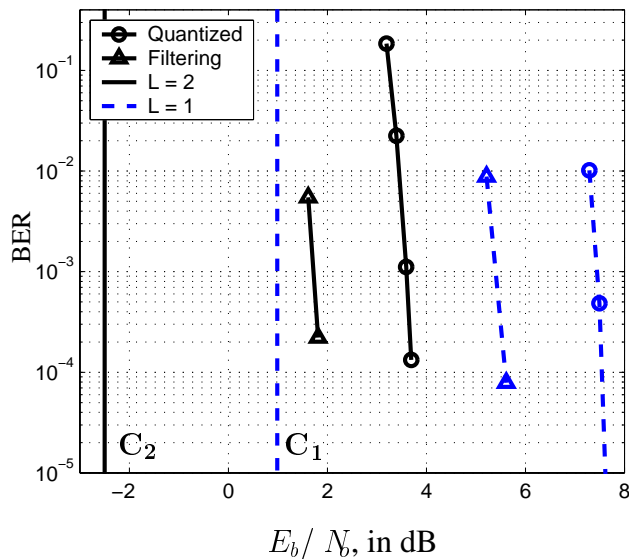


Fig. 7. BER performance in flat Rayleigh fading with  $f_D T = 0.05$ . Naming conventions are the same as in Fig. 6. Again, the second branch of diversity improves the simulations (dashed to solid curve with the same marker) by about as much as it improves the capacity (dashed to solid vertical line).

imum filtering of received samples to estimate the channel coefficients. Quantized or filtered channel estimates are then appropriately combined in the constituent turbo-decoders. The filtering solution performs better, but both methods appear to harvest most of the capacity gain promised by receiver antenna diversity.

#### REFERENCES

- [1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes (1)," in *ICC '93*, 1993, pp. 1064–1070.
- [2] C. Kominakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE Journal on Sel. Areas in Comm.*, to appear, second quarter 2001.
- [3] C. Douillard, M. Jézéquel, C. Berrou, A. Piccart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *Europ. Trans. on Telecom.*, vol. 6, no. 5, pp. 507–11, Sept.-Oct. 1995.
- [4] S. L. Ariyavisitakul, "Turbo space-time processing to improve wireless channel capacity," in *proceedings of ICC 2000, N. Orleans, LA*, pp. 1238–42, 18–22 June 2000.
- [5] G. Bauch, A. F. Naguib, and N. Seshadri, "MAP equalization of spcae-time coded signals over frequency-selective channels," *IEEE Wireless Communications and Networking Conference, WCNC 1999*, vol. 1, pp. 261–5, 1999.
- [6] J. Garcia-Frias and J.D. Villasenor, "Turbo codes for binary Markov channels," in *ICC '98, Atlanta, GA*, 7–11 June 1998, pp. 110–15.
- [7] M. C. Valenti, "Iterative channel estimation for turbo-codes over fading channels," *IEEE Wireless Communications and Networking Conference, WCNC 2000*, vol. 3, pp. 1019–1024, 2000.
- [8] H. J. Su and E. Geraniotis, "Improved performance of a PSAM system with iterative filtering and decoding," *Thirty-Sixth Annual Allerton Conference on Communication, Control and Computing*, pp. 156–166, Sept. 1998.
- [9] C. Kominakis and R. D. Wesel, "Pilot-aided joint data and channel estimation in flat correlated fading," in *Globecom '99*, 5–9 Dec. 1999.
- [10] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. on Inform. Theory*, vol. 20, no. 2, pp. 284–287, March 1974.
- [11] C. Fragouli and R. Wesel, "Turbo encoder design for symbol-interleaved parallel concatenated trellis coded modulation," *IEEE Trans. on Comm.*, vol. 49, no. 3, pp. 425–35, March 2001.
- [12] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.*, pp. 957–1000, July 1968.
- [13] W. C. Jakes, Jr., *Microwave Mobile Communications*, John Wiley & Sons, NY, 1974.
- [14] 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.*, vol. 9, no. 2, pp. 155–172, March–April 1998.
- [15] C. Kominakis, *Joint Channel Estimation and Decoding for Wireless Channels*, Ph.D. thesis, Univ. of California, Los Angeles, Dec. 2000.
- [16] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–35, 1998.