EM-BASED SOFT NONCOHERENT EQUALIZATION OF DOUBLY SELECTIVE CHANNELS USING TREE SEARCH AND BASIS EXPANSION

Sung-Jun Hwang and Philip Schniter

Dept. ECE, The Ohio State University, 2015 Neil Avenue, Columbus, OH 43210.
E-mail: {hwangsu,schniter}@ece.osu.edu

ABSTRACT

We propose a soft noncoherent equalizer for coded transmissions over unknown time- and frequency-selective, or doubly selective, channels. Such a soft equalizer can be used in conjunction with a soft decoder as part of a turbo reception scheme. Like a number of existing designs, ours is motivated by the use of the expectation maximization (EM) algorithm to estimate the unknown channel parameters, and manifests as an iteration between soft channel estimation and soft coherent equalization. As a departure from the existing designs, which assume a Gauss-Markov (i.e., autoregressive) channel and employ a trellis whose number-of-states grows exponentially in the channel length, we use a basis expansion (BE) model for the channel and a soft tree-search in place of the trellis, leveraging recent ideas from the MIMO literature. The complexity of our schemes grows only linearly in the block length and quadratically in the number of BE coefficients. Numerical experiments show performance that is close to genie-aided bounds and robust to mismatch in channel-fading rate.

1. INTRODUCTION

In this paper, we consider the problem of decoding a data sequence transmitted over an *unknown* time- and frequency-selective channel, i.e., doubly selective (DS) channel. In particular, we are interested in the case of coded transmissions with possibly long codewords (as with LDPC or turbo codes). A practical and near-optimal strategy for equalization in this scenario follows from the turbo principle [1], which suggests to iterate between separate soft equalization and soft decoding steps. In this case, the equalizer's role becomes that of producing posterior bit probabilities from the received samples and any extrinsic information previously supplied by the decoder.

Optimal soft noncoherent equalization can in principle be accomplished using either forward-backward [2] or fixed-lag processing [3], but doing so requires the computational equivalent of calculating a channel estimate for every possible symbol sequence. Thus, researchers have focused on practical yet sub-optimal alternatives. In one approach, approximations to the soft outputs are computed using reduced-state averages enabling a trellis where a low-order Gauss-Markov channel assumption facilitates per-branch linear prediction or Kalman filtering [3–6]. In another approach, soft noncoherent equalization is broken into two steps: soft channel estimation and coherent soft equalization, which are usually iterated [7–11]. The latter approach is connected to, and in many cases a direct application of, the expectation-maximization (EM) algorithm [12].

A potential drawback of the aforementioned practical approaches is the use of trellis with at least Q^{N_H-1} states, where Q is the alphabet size and N_H is the channel spread (measured in symbols). For channels with moderate-to-long spreads, the number of required states can be impractically large. Another potential drawback of the aforementioned practical approaches is the use of low-order Gauss-Markov models, i.e., autoregressive (AR) models, for channel variation. While such models are known to accurately model the "time-domain" channel trajectories encountered in single-carrier systems, they are not accurate models for the "frequency-domain" channel trajectories encountered in multi-carrier systems, where, e.g., one subcarrier may be deeply faded while adjacent subcarriers are not.

In this paper, we propose an EM-based soft noncoherent equalization algorithm that is well suited to long and quickly varying channel responses. Our approach is appropriate for, e.g., underwater acoustic as well as radio-frequency (RF) channels. Compared to previously described literature, our approach differs principally in two regards: 1) in place of trellis techniques, we use soft tree search, and 2) in place of the Gauss-Markov model, we use a generic basis expansion model [13, 14]. Our use of soft tree-search builds on recent ideas from the multiple-input multiple-output (MIMO) literature (e.g., [15, 16]), yielding a flexible and efficient tradeoff between performance and complexity, and our generic BE model allows accurate modeling of many channel classes, including time-domain, frequency-domain, and sparse channels (see, e.g., [17]). Our combination of soft tree-search and BE channel modeling yields a soft noncoherent equalizer whose complexity is linear in the block length and quadratic in the number of BE coefficients.

As in [8–10], we employ the "Bayesian" EM algorithm [12] to iteratively estimate the unknown channel parameters while treating the unknown bits as "missing data". In doing so, posterior bit probabilities are generated as a by-product. Numerical experiments with LDPC-coded transmissions over a DS Rayleigh-fading channel show our receiver performing within 2 dB of genie-aided performance bounds. Additional experiments demonstrate the robustness of our scheme to mismatch in the assumed Doppler spread.

Finally, we mention two related works. The noncoherent equalizer in [11] also iterates between soft coherent equalization and soft BE-coefficient estimation, but chooses LMMSE equalization over MAP equalization. In contrast, [18] does not split noncoherent equalization into separate coherent-equalization and BE-coefficient estimation steps; rather, it implements noncoherent equalization directly using a tree search with a noncoherent metric.

2. SYSTEM MODEL

At the transmitter, we assume that information bits $\{b_m^{(j)}\}$ are rate-R coded, interleaved, and mapped to 2^Q-ary QAM symbols. Groups of N_s information symbols are then combined with pilot and guard symbols to form symbol blocks of length $N \geq N_s$. (For brevity,

This work supported by National Science Foundation CAREER grant CCR-0237037 and Office of Naval Research grant N00014-07-1-0209.

we omit the details of the pilot pattern; see [17, 18] for pilot details.) We denote the j^{th} symbol block by $\boldsymbol{s}^{(j)} = [s_0^{(j)}, \ldots, s_{N-1}^{(j)}]^T$ and the corresponding coded bit vector by $\boldsymbol{x}^{(j)} = [x_0^{(j)}, \ldots, x_{N_s Q-1}^{(j)}]^T$. The symbol are then linearly block-modulated by either a single-carrier scheme or a multi-carrier scheme, represented by $\boldsymbol{G} \in \mathbb{C}^{N_t \times N}$ with $N_t \geq N$, to form the transmitted signal $\boldsymbol{t}^{(j)} \triangleq [t_0^{(j)}, \ldots, t_{N_t-1}^{(j)}]^T = \boldsymbol{G}\boldsymbol{s}^{(j)}$. The construction of $\boldsymbol{G}$ will be described later.

At the channel output, the samples in the j^{th} received block $\boldsymbol{r}^{(j)} \triangleq [r_0^{(j)}, \ldots, r_{N_r-1}^{(j)}]^T$ are assumed to take the form

$$r_n^{(j)} \;=\; \sum_{l=0}^{N_h-1} h_{n,l}^{(j)} t_{n-l}^{(j)} + v_n^{(j)}, \tag{1}$$

where $h_{n,l}^{(j)}$ is the time-n response of the channel to an impulse applied at time-$(n-l)$, where N_h is the discrete channel delay spread, and where $\{v_n^{(j)}\}$ is zero-mean circular white Gaussian noise (CWGN) with covariance σ^2.

The received vector $\boldsymbol{r}^{(j)}$ is then linearly (single- or multi-carrier) demodulated via matrix $\boldsymbol{\Gamma} \in \mathbb{C}^{N \times N_r}$ to yield

$$\boldsymbol{y}^{(j)} \;=\; \underbrace{\boldsymbol{\Gamma}\boldsymbol{\mathcal{H}}^{(j)}\boldsymbol{G}}_{\triangleq\, \boldsymbol{H}^{(j)}} \boldsymbol{s}^{(j)} + \boldsymbol{z}^{(j)}. \tag{2}$$

In (2), $\boldsymbol{z}^{(j)} = \boldsymbol{\Gamma}\boldsymbol{v}^{(j)}$ and $\boldsymbol{\mathcal{H}}^{(j)} \in \mathbb{C}^{N_r \times N_t}$ is a convolution matrix constructed from the channel's time-varying impulse response according to $[\boldsymbol{\mathcal{H}}^{(j)}]_{n,n-l} = h_{n,l}^{(j)}$. Thus $N_r = N_t + N_h - 1$ and $\boldsymbol{\mathcal{H}}^{(j)}$ is banded with bandwidth N_h. Note that $\boldsymbol{H}^{(j)}$ represents the composite effect of modulation, channel propagation, and demodulation. When the single- or multi-carrier scheme is appropriately designed, $\boldsymbol{H}^{(j)}$ can be closely approximated by a "circularly banded" matrix [19] with bandwidth N_H:

- In the single-carrier case, $\boldsymbol{G} = \boldsymbol{I}_N$ (so $N_t = N$) and

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{I}_{N_h-1} & \boldsymbol{0} & \boldsymbol{I}_{N_h-1} \\ \boldsymbol{0} & \boldsymbol{I}_{N-N_h+1} & \boldsymbol{0} \end{bmatrix}.$$

  Thus $\boldsymbol{H}^{(j)}$, with bandwidth $N_H = N_h$, contains the impulse response coefficients $\{h_{n,l}^{(j)}\}$.

- In the multi-carrier case, $\boldsymbol{G} = \mathcal{D}(\boldsymbol{g})\boldsymbol{F}_t^H$, where $\boldsymbol{F}_t^H \in \mathbb{C}^{N_t \times N}$ is a period-N unitary IDFT matrix cyclically extended in the row dimension, and where $\mathcal{D}(\boldsymbol{g})$ is a diagonal matrix created from a time-domain transmission pulse $\boldsymbol{g} \in \mathbb{C}^{N_t}$. Then $\boldsymbol{\Gamma} = \boldsymbol{F}_r\mathcal{D}(\boldsymbol{\gamma})$, where $\boldsymbol{F}_r \in \mathbb{C}^{N \times N_r}$ is a period-N unitary DFT matrix cyclically extended in the column dimension, and $\boldsymbol{\gamma} \in \mathbb{C}^{N_r}$ is a time-domain reception pulse. With appropriate design of $\boldsymbol{g}$ and $\boldsymbol{\gamma}$ [19], the frequency-domain channel matrix $\boldsymbol{H}^{(j)}$ has bandwidth $N_H = N_{\mathsf{D}} \triangleq \lceil 2 f_{\mathsf{D}} T_s N \rceil$ where f_{D} denotes the single-sided Doppler spread (in Hz), and T_s denotes the channel-use interval (in sec). The off-diagonal elements of $\boldsymbol{H}^{(j)}$ induce inter-carrier interference (ICI).

We assume that the last $N_H - 1$ symbols in $\boldsymbol{s}$ are guards, so that $\boldsymbol{H}^{(j)}$ acts causally on the first $N - N_H + 1$ symbols.

The receiver infers the information bits $\{b_m^{(j)}\}$ using the "turbo" principle: "soft" information on the coded bits $\boldsymbol{x}^{(j)}$, in the form of log-likelihood ratios (LLRs), is iteratively refined through alternating soft-equalization and soft-decoding steps, as shown in Fig. 1. Our equalizer is "noncoherent" in that it operates without externally provided channel state information.

The equalizer employs an N_b-term basis expansion (BE) model for the variation of the composite channel over the block. In particular, the d^{th} "cyclic" diagonal of $\boldsymbol{H}^{(j)}$, i.e., $\boldsymbol{h}_d^{(j)} \triangleq \left[[\boldsymbol{H}^{(j)}]_{0,-d}, [\boldsymbol{H}^{(j)}]_{1,1-d}, \ldots, [\boldsymbol{H}^{(j)}]_{N-1,N-1-d}\right]^T$ for modulo-N indices, is modeled as

$$\boldsymbol{h}_d^{(j)} \;\approx\; \boldsymbol{B}\boldsymbol{\eta}_d^{(j)}, \quad d = 0, \ldots, N_H - 1, \tag{3}$$

where $\boldsymbol{B} \in \mathbb{C}^{N \times N_b}$ is a matrix of basis vectors and $\boldsymbol{\eta}_d^{(j)} \in \mathbb{C}^{N_b}$ is a vector of BE coefficients. While an error-free approximation is possible when $N_b = N$, significant reduction in receiver complexity is possible when $N_b \ll N$. Note that, with single-carrier modulation, the BE models channel variation in the time domain, so that $N_b = N_{\mathsf{D}}$ suffices (with appropriate choice of $\boldsymbol{B}$). With multi-carrier modulation, the BE models channel variation in the frequency domain, so that $N_b = N_h$ suffices. In either case, we have $N_b N_H = N_h N_{\mathsf{D}}$.

Under the BE approximation (3), the received vector $\boldsymbol{y}^{(j)}$ from (2) becomes

$$\boldsymbol{y}^{(j)} \;=\; \boldsymbol{A}^{(j)}\boldsymbol{\theta}^{(j)} + \boldsymbol{z}^{(j)}, \tag{4}$$

where $\boldsymbol{\theta}^{(j)} \triangleq [\boldsymbol{\eta}_0^{(j)T}, \ldots, \boldsymbol{\eta}_{N_H-1}^{(j)T}]^T \in \mathbb{C}^{N_b N_H}$ and

$$\boldsymbol{A}^{(j)} \;\triangleq\; \left[\mathcal{D}_0(\boldsymbol{s}^{(j)})\boldsymbol{B}, \ldots, \mathcal{D}_{N_H-1}(\boldsymbol{s}^{(j)})\boldsymbol{B}\right]. \tag{5}$$

In (5) we use $\mathcal{D}_d(\cdot)$ to denote the diagonal matrix constructed from the d^{th} cyclic down-shift of its vector argument.

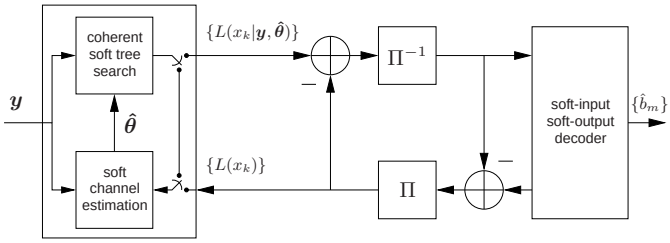

Fig. 1. Turbo receiver with EM-based soft noncoherent equalizer.

3. SOFT NONCOHERENT EQUALIZATION

Before going into the details, we give a high-level description of our soft noncoherent equalizer. At each turbo iteration, the equalizer generates (an approximation of) the posterior LLRs $\{L(x_k^{(j)}|\boldsymbol{y}^{(j)})\}_{k=0}^{N_s Q-1}$ from the prior LLRs $\{L(x_k^{(j)})\}_{k=0}^{N_s Q-1}$ obtained from the decoder; on the first turbo iteration, the prior LLRs are set to zero. Since calculation of the posterior-LLRs is complicated by the fact that the channel state $\boldsymbol{\theta}^{(j)}$ is unknown, we take the practical approach of *jointly* estimating the channel $\boldsymbol{\theta}^{(j)}$ and coded bits $\boldsymbol{x}^{(j)}$ using the Bayesian EM algorithm [8], which iteratively computes the channel estimate $\hat{\boldsymbol{\theta}}_{\mathsf{MAP}}^{(j)} \triangleq \arg\max_{\hat{\boldsymbol{\theta}}^{(j)}} p(\hat{\boldsymbol{\theta}}^{(j)}|\boldsymbol{y}^{(j)})$ and generates the posteriors $\{L(x_k^{(j)}|\boldsymbol{y}^{(j)}, \hat{\boldsymbol{\theta}}_{\mathsf{MAP}}^{(j)})\}$ as a by-product. Because this procedure is invariant to the block index j, we suppress the "$^{(j)}$" notation in the sequel.

3.1. Equalization via the Bayesian EM Algorithm

We now detail the application of the EM algorithm to soft noncoherent equalization using the system model of Section 2. Given the iteration-i parameter estimate $\hat{\boldsymbol{\theta}}[i]$, the "incomplete data" $\boldsymbol{y}$, and the

"missing data" $\boldsymbol{s}$, the iteration-$(i+1)$ parameter estimate generated by the Bayesian EM algorithm is (for $i \geq 1$) [12]

$$\hat{\boldsymbol{\theta}}[i+1] \triangleq \arg\max_{\hat{\boldsymbol{\theta}}} \mathrm{E}\big\{ \ln p(\boldsymbol{y}, \boldsymbol{s}|\hat{\boldsymbol{\theta}}) \,\big|\, \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i] \big\} + \ln p(\hat{\boldsymbol{\theta}}). \quad (6)$$

The posterior probabilities of $\hat{\boldsymbol{\theta}}[1], \hat{\boldsymbol{\theta}}[2], \hat{\boldsymbol{\theta}}[3], \ldots$ are known to be non-decreasing, implying that the EM estimates will converge to $\hat{\boldsymbol{\theta}}_{\mathsf{MAP}}$ when $p(\boldsymbol{\theta}|\boldsymbol{y})$ is unimodal in $\boldsymbol{\theta}$. The "Bayesian" version of the EM algorithm requires a prior on the unknown parameter $\boldsymbol{\theta}$, and we assume that $\boldsymbol{\theta} \sim \mathcal{CN}(\bar{\boldsymbol{\theta}}, \boldsymbol{R}_\theta)$ for our application.

Assuming independence between data and channel, i.e., $p(\boldsymbol{s}|\boldsymbol{\theta}) = p(\boldsymbol{s})$, we can write $p(\boldsymbol{y}, \boldsymbol{s}|\boldsymbol{\theta}) = p(\boldsymbol{y}|\boldsymbol{s}, \boldsymbol{\theta}) p(\boldsymbol{s})$, where $\boldsymbol{y}|\boldsymbol{s}, \boldsymbol{\theta} \sim \mathcal{CN}(\boldsymbol{A}\boldsymbol{\theta}, \sigma^2 \boldsymbol{I})$ and $p(\boldsymbol{s}) = \prod_{n=0}^{N-1} p(s_n)$. Plugging the assumed pdfs into (6) and simplifying yields

$$\hat{\boldsymbol{\theta}}[i+1] = \arg\min_{\hat{\boldsymbol{\theta}}} \tfrac{1}{\sigma^2} \mathrm{E}\big\{ \|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\theta}}\|^2 \big| \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i] \big\} + \|\hat{\boldsymbol{\theta}} - \bar{\boldsymbol{\theta}}\|^2_{\boldsymbol{R}_\theta^{-1}} \quad (7)$$

where we take $\hat{\boldsymbol{\theta}}[0] = \bar{\boldsymbol{\theta}}$. Zeroing the partial derivative w.r.t $\hat{\boldsymbol{\theta}}$ gives

$$\hat{\boldsymbol{\theta}}[i+1] = \big(\boldsymbol{C} + \sigma^2 \boldsymbol{R}_\theta^{-1}\big)^{-1} \big(\bar{\boldsymbol{A}}^H \boldsymbol{y} + \sigma^2 \boldsymbol{R}_\theta^{-1} \bar{\boldsymbol{\theta}}\big) \quad (8)$$

$$= \bar{\boldsymbol{\theta}} + \big(\boldsymbol{C} + \sigma^2 \boldsymbol{R}_\theta^{-1}\big)^{-1} \big(\bar{\boldsymbol{A}}^H \boldsymbol{y} - \boldsymbol{C}\bar{\boldsymbol{\theta}}\big) \quad (9)$$

for $\bar{\boldsymbol{A}} \triangleq \mathrm{E}\{\boldsymbol{A}\,|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}$ and $\boldsymbol{C} \triangleq \mathrm{E}\{\boldsymbol{A}^H \boldsymbol{A}\,|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}$. Collecting the posterior quantities $\bar{s}_n \triangleq \mathrm{E}\{s_n|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}$ and $c_n \triangleq \mathrm{E}\{|s_n|^2\,|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}$ into the vectors $\bar{\boldsymbol{s}} \triangleq [\bar{s}_0, \ldots, \bar{s}_{N-1}]^T$ and $\boldsymbol{c} \triangleq [c_0, \ldots, c_{N-1}]^T$, we can recall (5) and write

$$\bar{\boldsymbol{A}} = \big[\mathcal{D}_0(\bar{\boldsymbol{s}})\boldsymbol{B}, \ldots, \mathcal{D}_{N_H-1}(\bar{\boldsymbol{s}})\boldsymbol{B}\big] \quad (10)$$

$$\boldsymbol{C} = \bar{\boldsymbol{A}}^H \bar{\boldsymbol{A}} - \begin{bmatrix} \boldsymbol{\Delta}_0 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{\Delta}_{N_H-1} \end{bmatrix} \quad (11)$$

$$\boldsymbol{\Delta}_d \triangleq \boldsymbol{B}^H \Big(\mathcal{D}_d(\bar{\boldsymbol{s}})^H \mathcal{D}_d(\bar{\boldsymbol{s}}) - \mathcal{D}_d(\boldsymbol{c}) \Big) \boldsymbol{B}. \quad (12)$$

To calculate $\bar{\boldsymbol{s}}$ and $\boldsymbol{c}$, we use the definitions

$$\bar{s}_n = \sum_{s \in \mathcal{S}} s \, \Pr\{s_n = s|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\} \quad (13)$$

$$c_n = \sum_{s \in \mathcal{S}} |s|^2 \Pr\{s_n = s|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}, \quad (14)$$

where $\mathcal{S}$ is the symbol constellation, and calculate the symbol posteriors $\{p(s_n|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i])\}_{n=0}^{N-1}$ from the bit LLRs, defined as

$$L(x_k|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]) \triangleq \ln \frac{\Pr\{x_k = 1|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}}{\Pr\{x_k = 0|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}}. \quad (15)$$

For example, it can be shown that $\bar{s}_n = \tanh\{\frac{1}{2} L(x_{2n}|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i])\} + j \tanh\{\frac{1}{2} L(x_{2n+1}|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i])\}$ and $c_n = 2$ with QPSK, i.e., $\mathcal{S} = \{\pm 1 \pm j\}$. Then, to calculate the LLRs, one can (in principle) use [15]

$$L(x_k|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]) = \ln \frac{\sum_{\boldsymbol{x}: x_k = 1} \exp \mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i])}{\sum_{\boldsymbol{x}: x_k = 0} \exp \mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i])} \quad (16)$$

$$\mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i]) \triangleq \ln p(\boldsymbol{y}|\boldsymbol{x}, \hat{\boldsymbol{\theta}}[i]) p(\boldsymbol{x}) \quad (17)$$

$$= -\tfrac{1}{\sigma^2} \|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\theta}}[i]\|^2 + \boldsymbol{l}^T \boldsymbol{x}, \quad (18)$$

for $\boldsymbol{x} = [x_0, \ldots, x_{N_s Q-1}]^T$ and $\boldsymbol{l} \triangleq [L(x_0), \ldots, L(x_{N_s Q-1})]^T$.

We can identify (13)–(18) as the "expectation-step" and (9)–(12) as the "maximization-step" of the EM recursion. The E-step calculates the posterior symbol quantities $\bar{\boldsymbol{s}}$ and $\boldsymbol{c}$ assuming the channel estimate $\hat{\boldsymbol{\theta}}[i]$, and the M-step updates the channel estimate based on the output of the E-step. We can also interpret the EM-based noncoherent equalizer as iterating between "soft" channel estimation, i.e., calculation of $\hat{\boldsymbol{\theta}}[i]$ from LLRs and $\boldsymbol{y}$, and soft coherent tree search, i.e., calculation of LLRs from $\hat{\boldsymbol{\theta}}[i]$ and $\boldsymbol{y}$, as illustrated in Fig. 1. The LLRs used in the first EM iteration, $\{L(x_k)\}_{k=0}^{N_s Q-1}$, come from the decoder, and the the LLRs produced in the last EM iteration, $\{L(x_k|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[K])\}_{k=0}^{N_s Q-1}$, are passed to the decoder.

3.2. Modifications for Practical Implementation

One can observe two principal challenges for practical implementation of the EM-based equalizer described in Section 3.1. First, calculation of the posterior LLRs $\{L(x_k|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i])\}_{k=0}^{N_s Q-1}$ via (16) is computationally impractical. So, we apply the "max-log" approximation $\ln \sum_{\boldsymbol{x}: x_k = 1} \exp \mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i]) \approx \max_{\boldsymbol{x}: x_k = 1} \mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i])$ to (16), as in [15], giving

$$L(x_k|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]) \approx \max_{\mathcal{X}[i] \cap \{\boldsymbol{x}: x_k = 1\}} \mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i]) - \max_{\mathcal{X}[i] \cap \{\boldsymbol{x}: x_k = 0\}} \mu(\boldsymbol{x}|\hat{\boldsymbol{\theta}}[i]), \quad (19)$$

where $\mathcal{X}[i]$ denotes the set of bit sequences $\{\boldsymbol{x}\}$ having dominant posterior probability $p(\boldsymbol{x}|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i])$. Then, the (usually very small) set $\mathcal{X}[i]$ can be efficiently found via *coherent* tree search based on the M-algorithm [16]. This procedure has complexity $\mathcal{O}(M 2^Q N N_b N_H)$, where M is the search breadth.

The second challenge results from the matrix inversion in (9). To reduce complexity, we make the approximation

$$c_n \approx |\bar{s}_n|^2, \quad (20)$$

which leads to $\boldsymbol{\Delta}_d \approx \boldsymbol{0}\ \forall d$ and thus $\boldsymbol{C} \approx \bar{\boldsymbol{A}}^H \bar{\boldsymbol{A}}$. We justify this by claiming that, as the turbo iterations proceed, the symbol estimates become more reliable, thereby reducing the symbol variance $c_n - |\bar{s}_n|^2$ and satisfying (20). (In Section 4, we demonstrate the effect of (20) numerically.) With this approximation, (9) simplifies to

$$\hat{\boldsymbol{\theta}}[i+1] = \bar{\boldsymbol{\theta}} + \big(\bar{\boldsymbol{A}}^H \bar{\boldsymbol{A}} + \sigma^2 \boldsymbol{R}_\theta^{-1}\big)^{-1} \bar{\boldsymbol{A}}^H (\boldsymbol{y} - \bar{\boldsymbol{A}}\bar{\boldsymbol{\theta}}), \quad (21)$$

allowing the application of the efficient sequential-Bayes recursion summarized in Table 1. (A similar recursion was derived in our previous work [18] on *noncoherent* tree search.)

set $\{\boldsymbol{\Sigma}_{-1}^{-1},\ \hat{\boldsymbol{\theta}}_{-1}[i]\} \triangleq \{\sigma^{-2}\boldsymbol{R}_\theta,\ \bar{\boldsymbol{\theta}}\}$;
for $n = 0, 1, 2, \ldots, N-1$,
  $\boldsymbol{a}_n = [\bar{s}_n \boldsymbol{b}_n^H, \cdots, \bar{s}_{n-N_H+1} \boldsymbol{b}_n^H]^H$;
  $\boldsymbol{d}_n = \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{a}_n$;
  $\alpha_n = (1 + \boldsymbol{a}_n^H \boldsymbol{d}_n)^{-1}$;
  $\boldsymbol{\Sigma}_n^{-1} = \boldsymbol{\Sigma}_{n-1}^{-1} - \alpha_n \boldsymbol{d}_n \boldsymbol{d}_n^H$;
  $\hat{\boldsymbol{\theta}}_n[i] = \hat{\boldsymbol{\theta}}_{n-1}[i] + \alpha_n (y_n - \boldsymbol{a}_n^H \hat{\boldsymbol{\theta}}_{n-1}[i]) \boldsymbol{d}_n$;
end

Table 1. N-step recursion to compute $\hat{\boldsymbol{\theta}}[i]$ via the temporary vectors $\hat{\boldsymbol{\theta}}_0[i], \ldots, \hat{\boldsymbol{\theta}}_{N-1}[i]$. Here, $\boldsymbol{b}_n^H$ denotes the n^{th} row of $\boldsymbol{B}$.

The simplified EM implementation described above consumes $\mathcal{O}(KN(N_b N_H)^2)$ complex multiplications per N-block, which is linear in the block size and quadratic in the number of BE parameters $N_b N_H$. The number of EM iterations, K, is typically small (e.g., $K = 3$ for our numerical results).

3.3. Modifications for Improved Performance

We now propose one final modification to the equalization algorithm. Though (18) specifies the use of the priors $\{L(x_k)\}_{k=0}^{N_sQ-1}$ at every EM iteration, we have observed that performance improves significantly if the most recently calculated posteriors $\{L(x_k|\boldsymbol{y}, \hat{\boldsymbol{\theta}}[i-1])\}_{k=0}^{N_sQ-1}$ are used in place of the priors in (18) for iterations $i \geq 2$.

4. NUMERICAL RESULTS

For the numerical experiments, Jakes method [20] was used to generate realizations of a wide-sense stationary uncorrelated scattering (WSSUS) Rayleigh fading channel with uniform delay-power profile $\sigma_l^2 = N_h^{-1}$ and temporal autocorrelation $\rho_m = J_0(2\pi f_{\mathsf{D}} T_s m)$. Here, f_{D} denotes the single-sided Doppler spread (in Hz), T_s denotes the channel use interval (in sec), and $J_0(\cdot)$ denotes the 0^{th}-order Bessel function of the first kind. The values $f_{\mathsf{D}}T_s = 0.002$ and $N_h = 3$ were assumed unless otherwise noted.

The transmitter employed rate-$R = \frac{1}{2}$ irregular low density parity check (LDPC) codes with average column-weight 3, generated via publicly available software [21]. The coded bits were mapped to QPSK symbols (i.e., $Q = 2$) and partitioned into data blocks of length N_s, each of which was merged with N_p leading pilots and $N_h - 1$ trailing zeros to form a transmission block of length $N = N_s + N_p + N_h - 1$. So that each codeword spanned $J = 32$ data blocks, $(JQN_s, RJQN_s)$-LDPC codes were employed. Throughout, we used block length $N = 64$ with $N_p = 6$ pilots and $N_s = 56$ information symbols per block. Though our BE-based soft noncoherent equalizer can be applied to either single- or multi-carrier communication, we consider only single-carrier experiments here. This restriction enables comparison to noncoherent equalizers that use a low-order Gauss-Markov channel model to facilitate Kalman channel estimation (e.g., [22]).

Our soft noncoherent equalizer used a Karhunen-Lóeve (KL) BE channel model [14] with $N_b = 3$ to model the channel's time variation. In other words, $\boldsymbol{B}$ was constructed column-wise from the N_b principal eigenvectors of $\boldsymbol{R}_h \triangleq \mathrm{E}\{\boldsymbol{h}_d \boldsymbol{h}_d^H\}$ and diagonal $\boldsymbol{R}_\theta$ was constructed from the N_b principal eigenvalues of $\boldsymbol{R}_h$. Since the channel was Rayleigh, we used $\bar{\boldsymbol{\theta}} = \boldsymbol{0}$. The LDPC decoder of MacKay and Neal [23] was used with a maximum of 60 "inner" iterations, and equalization/decoding were iterated using a maximum of 8 "outer" iterations. We specify the *maximum* number of iterations because the receiver breaks out of both the inner and outer loops as soon as the LDPC syndrome check indicates error-free decoding.

In Fig. 2, the soft noncoherent equalizer proposed in Section 3, which we henceforth refer to as

- coherent tree-search coupled with soft BE-channel estimation after K iterations: "$(\mathsf{cT{+}sBE})^K$,"

was compared to several other soft noncoherent equalizers:

- coherent tree-search coupled with *exact* soft BE-channel estimation after K iterations: "$(\mathsf{cT{+}esBE})^K$,"
- coherent tree-search using a soft BE-channel estimate *non-iteratively*: "cT+sBE,"
- coherent tree-search using a soft Gauss-Markov estimate *non-iteratively*: "cT+sGM,"

and two genie-aided performance bounds:

- coherent tree-search based on a *perfect* estimate of the channel $\boldsymbol{H}$: "cT+pH,"
- coherent tree-search based on a BE-channel estimate constructed using *perfect LLR feedback*: "cT+pllrBE."

We now elaborate on the procedures cT, sBE, esBE, sGM, pH, and pllrBE mentioned above. As discussed in Section 3.2, *coherent tree search* (cT) uses the M-algorithm to sequentially maximize the metric $\ln p(\boldsymbol{x}|\boldsymbol{y}, \hat{\boldsymbol{H}})$ for externally supplied $\hat{\boldsymbol{H}}$—a direct application of the MIMO technique [16]. In all cases, we employed a search breadth of $M = 64$ for the M-algorithm, and $K = 3$ EM iterations. *Soft BE-channel estimation* (sBE) uses the simplified EM-recursion (21), while *exact soft BE-channel estimation* (esBE) uses the "exact" recursion (9). Finally, *soft Gauss-Markov channel estimation* (sGM) refers to the Kalman technique proposed in [22], for which we employed a first-order Gauss-Markov model.

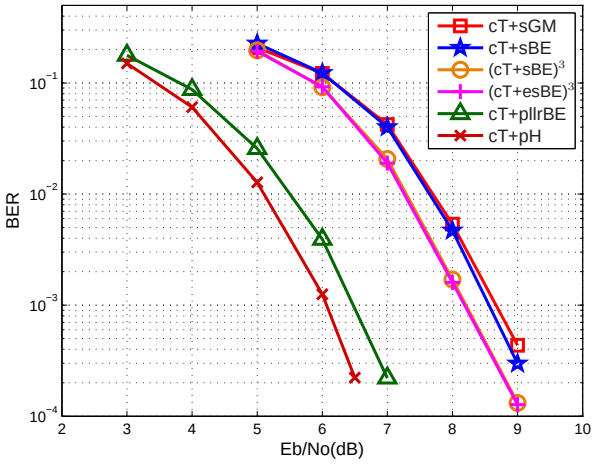

Fig. 2. BER vs. E_b/N_o for various equalization schemes under $f_{\mathsf{D}}T_s = 0.002$.

Figure 2 shows that the proposed $(\mathsf{cT{+}sBE})^3$ approach, a simplification of the expensive $(\mathsf{cT{+}esBE})^3$ approach, gives almost identical performance. Furthermore, the proposed $(\mathsf{cT{+}sBE})^3$ performs only 2 dB from the perfect-CSI bound cT+pH and only 1.7 dB from the perfect-LLR-feedback bound cT+pllrBE. The traces cT+sBE and cT+sGM can be used to compare between the use of BE versus Gauss-Markov channel models; it can be seen that they are almost the same except high SNR, where the BE approach shows slightly better performance. For the frequency-domain channels experienced in multi-carrier applications, we would expect to see a much bigger difference between BE and GM. To see the gain from multiple EM iterations, one can compare $(\mathsf{cT{+}sBE})^3$ (where $K = 3$) to cT+sBE (where $K = 1$); about 0.5 dB improvement can be observed. No additional gains were observed for $K > 3$.

Though the proposed noncoherent equalization scheme operated without knowledge of the channel state, it did assume knowledge of channel distribution in the form of the BE coefficient covariance matrix $\boldsymbol{R}_\theta$. We now examine the robustness of the scheme to knowledge of Doppler spread $f_{\mathsf{D}}T_s$, the most important parameter in the construction of $\boldsymbol{R}_\theta$, by comparing the BER-versus-$f_{\mathsf{D}}T_s$ performance of the equalizer with perfect knowledge of $f_{\mathsf{D}}T_s$ to one which assumes the fixed value $f_{\mathsf{D}}T_s = 0.002$. For this experiment, we fixed $E_b/N_o = 8$ dB. Figure 3 demonstrates that the proposed equalization scheme is robust to mismatch in Doppler-spread: the "mismatched" scheme stays close to the "matched" scheme over the entire range of tested Doppler spreads. Note that, as $f_{\mathsf{D}}T_s$ decreases, the BER for the matched scheme increases due to a lack of diversity; while the BER increase at $f_{\mathsf{D}}T_s = 0.0001$ may at first seem large, Fig. 2 shows that it is equivalent to a loss of 1 dB SNR. Likewise, the

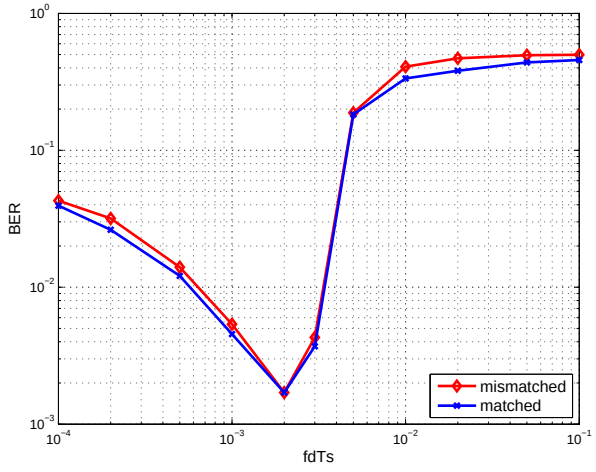

Fig. 3. BER vs. true $f_{\mathsf{D}}T_s$ for $f_{\mathsf{D}}T_s$-matched and fixed-$f_{\mathsf{D}}T_s$ (i.e., mismatched) reception at $E_b/N_o = 8$ dB.

BER for the matched scheme increases sharply with $f_{\mathsf{D}}T_s$ due to the limitations of the $N_b = 3$ BE model. Similar behavior was observed for other soft noncoherent equalizers in [17].

5. CONCLUSION

In this paper, we proposed a soft noncoherent equalizer for unknown DS channels for use in a turbo receiver. Our design was based on the use of the Bayesian EM algorithm to estimate the channel parameters, and it manifested as iterations between a soft coherent equalizer and a soft channel estimator. The receiver modeled the channel via basis expansion (BE), and performed soft coherent equalization via soft tree search. These two operations were accomplished using fast algorithms whose overall complexity grows linearly in the block size and quadratically in the number of BE parameters. Numerical studies show that our equalizer performs within 2dB from genie-aided bounds and remains robust to mismatch in assumed Doppler spread.

6. REFERENCES

[1] R. Koetter, A. C. Singer, and M. Tüchler, "Turbo equalization," *IEEE Signal Processing Mag.*, pp. 67–80, Jan. 2004.

[2] M. J. Gertsman and J. H. Lodge, "Symbol-by-symbol MAP demodulation of CPM and PSK signals on Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 788–799, Jul. 1997.

[3] Y. Zhang, M. P. Fitz, and S. B. Gelfand, "Soft output demodulation on frequency-selective Rayleigh fading channels using AR channel models," in *Proc. Allerton Conf. Commun., Control, and Computing*, Oct. 1995.

[4] R. A. Iltis, "Bayesian algorithms for blind equalization using parallel adaptive filtering," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1017–1032, Feb./Mar./Apr. 1994.

[5] B. D. Hart and S. Pasupathy, "Innovations-based MAP detection for time-varying frequency-selective channels," *IEEE Trans. Veh. Tech.*, vol. 48, no. 9, pp. 1507–1519, Sept. 2000.

[6] L. M. Davis, I. B. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, Dec. 2001.

[7] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, pp. 2406–2413, July 1994.

[8] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation on frequency-flat Rayleigh fading channels via a Bayesian EM algorithm," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1869–1872, Nov. 2001.

[9] M. Yan and B. D. Rao, "Soft decision-directed MAP estimate of fast Rayleigh flat fading channels," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 1965–1969, Dec. 2003.

[10] M. Nissilä and S. Pasupathy, "Adaptive Bayesian and EM-based detectors for frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1325–1336, Aug. 2003.

[11] K. Fang, L. Rugini, and G. Leus, "Iterative channel estimation and turbo equalization for time-varying OFDM systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2008.

[12] A. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc.*, vol. 39, pp. 1–17, 1977.

[13] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adaptive Control & Signal Processing*, vol. 10, no. 2/3, pp. 159–176, Mar. 1996.

[14] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 862–873, June 1999.

[15] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389–399, Mar. 2003.

[16] Y. L. C. de Jong and T. J. Willink, "Iterative tree search detection for MIMO wireless systems," *IEEE Trans. Commun.*, vol. 53, pp. 930–935, June 2005.

[17] S.-J. Hwang and P. Schniter, "Efficient multicarrier communication for highly spread underwater acoustic channels," *IEEE J. Select. Areas In Commun.*, vol. 26, no. 9, pp. 1674–1683, Dec. 2008.

[18] ——, "Fast noncoherent decoding of block transmissions over doubly dispersive channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 2007.

[19] S. Das and P. Schniter, "Max-SINR ISI/ICI-shaping multicarrier communication over the doubly dispersive channel," *IEEE Trans. Signal Processing*, vol. 55, no. 12, pp. 5782–5795, Dec. 2007.

[20] W. C. Jakes, *Microwave Mobile Communications*. Wiley, 1974, [reprinted by IEEE Press].

[21] I. Kozintsev, "Matlab programs for encoding and decoding of LDPC codes over GF(2^m)," http://www.kozintsev.net/soft.html.

[22] S. Song, A. C. Singer, and K.-M. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Signal Processing*, vol. 52, no. 10, pp. 2885–2894, Oct. 2004.

[23] J. D. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronics Letters*, vol. 33, pp. 457–458, Mar. 1997.