Research Article

A Gaussian Mixture Approach to Blind Equalization of Block-Oriented Wireless Communications

Frederic Lehmann (EURASIP Member)

Institut TELECOM, TELECOM SudParis, Department CITI, UMR-CNRS 5157, 91011 Evry Cedex, France

Correspondence should be addressed to Frederic Lehmann, frederic.lehmann@it-sudparis.eu

We consider blind equalization for block transmissions over the frequency selective Rayleigh fading channel. In the absence of pilot symbols, the receiver must be able to perform joint equalization and blind channel identification. Relying on a mixed discrete-continuous state-space representation of the communication system, we introduce a blind Bayesian equalization algorithm based on a Gaussian mixture parameterization of the a posteriori probability density function (pdf) of the transmitted data and the channel. The performances of the proposed algorithm are compared with existing blind equalization techniques using numerical simulations for quasi-static and time-varying frequency selective wireless channels.

1. Introduction

Blind equalization has attracted considerable attention in the communication literature over the last three decades. The main advantage of blind transmissions is that they avoid the need for the transmission of training symbols and hence leave more communication resources for data.

The pioneering blind equalizers proposed by Sato [1] and Godard [2] use low-complexity finite impulse response filters. However, these methods suffer from local and slow convergence and may fail on ill-conditioned or time-varying channels.

Other authors have proposed symbol-by-symbol soft-input soft-output (SISO) equalizers based on trellis search algorithms. Two such approaches have been proposed so far to achieve SISO equalization in a blind or semiblind context. The first approach relies on fixed lag smoothing [3] and was further simplified in [4] by allowing pruning and decision feedback techniques. The second approach uses fixed interval smoothing [5, 6]. The aforementioned methods employ a trellis description of the intersymbol interference (ISI) [7], where each discrete ISI state has its associated channel estimate. Another fixed interval method, based on expectation-maximization channel identification, has appeared recently [8], but this technique is restricted to static channels.

In this paper, we will consider fixed interval smoothing, which is adapted to block-oriented communications. After modeling the ISI and the unknown channel at the receiver side, we obtain a combined state-space formulation of our communication system. Specifically, the ISI is modeled as a discrete state space with memory, while the (potentially time-varying) channel is modeled as an autoregressive (AR) process.

Main Contributions and Related Work. The main technical contribution of this work is the introduction of a blind equalization technique based on Gaussian mixtures. A major problem in blind equalization is that multiple modes arise in the *a posteriori* channel pdf, which originate from the phase ambiguities inherent to digital modulations. Assume that an equalizer is able to represent only a single mode, as is usually the case, it is likely that the wrong mode is retained during a fading event or due to the occasional occurrence of high observation noise. In such a situation, a classical equalizer is not able to recover the correct phase determination in a blind mode, since no pilot symbol is available. Therefore, all the subsequent decisions in the frame will be erroneous with high probability. The main feature of the proposed algorithm is that the multimodality of the channel *a posteriori* pdf is explicitly taken into account thanks to a Gaussian mixture parameterization. We derive
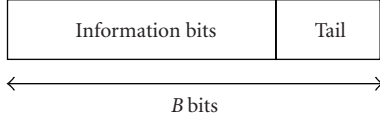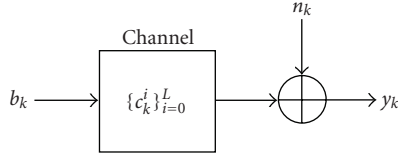
Figure 1: Data format.

Figure 2: Channel model.

the corresponding SISO smoothing algorithm suitable to solve our blind equalization problem.

Note that the idea of Gaussian mixture processing has been presented in [9] in the context of MIMO decoding and that the results in this paper have been partially presented in [10]. Also the idea of exploiting Gaussian mixtures for blind equalization appeared previously in a different form [11].

Organization. Section 2 describes the adopted system model. Section 3 introduces the Gaussian mixture-based blind equalization technique. Finally, in Section 4, the performances of the proposed technique are assessed through numerical simulations and compared with existing blind equalization techniques.

Notations. Throughout the paper, bold letters indicate vectors and matrices. $\mathcal{N}_{\mathbb{C}}(\mathbf{x} : \mathbf{m}, \mathbf{P})$ will denote a complex Gaussian distribution of the variable $\mathbf{x}$, with mean $\mathbf{m}$ and covariance matrix $\mathbf{P}$. $\mathbf{I}_m$ denotes the $m \times m$ identity matrix, while $\mathbf{0}_m$ denotes the $m \times m$ all-zero matrix. The symbol $\otimes$ denotes the Kronecker product. The operator $\det(\cdot)$ will denote the determinant of a matrix.

2. System Model

The transmitted data are organized in GSM-like bursts containing B bits, as illustrated in Figure 1. For ease of exposition, we consider binary phase shift keying (BPSK) modulation, so that the bit is transmitted at instant k, $b_k \in \{-1, +1\}$. The tail is a short all-one vector of length equal to the channel memory, used to set the final ISI state of the current data burst to a known value. At the same time, this also sets the initial ISI state of the next data burst to the same known value. Since blind equalization is of interest, no additional pilot symbols are inserted in the data stream.

We assume a discrete Rayleigh fading channel of memory L as depicted in Figure 2. The elements of the impulse response $\{c_k^i\}_{i=0}^L$ are modeled as independent zero-mean complex Gaussian random variables. For a static channel, the

channel state is defined as $\mathbf{x}_k = [c_k^0, c_k^1, \ldots, c_k^L]^T$ and evolves according to the trivial dynamical system

$$\mathbf{x}_k = \mathbf{x}_{k-1}. \tag{1}$$

For a time-varying channel, let B_d and T denote the maximum Doppler shift and the symbol duration, respectively. The channel autocorrelation can be modeled as follows [12]:

$$E\left[c_k^i c_{k-n}^{i\,*}\right] = J_0(2\pi n B_d T), \quad i = 0, \ldots, L, \tag{2}$$

where J_0 is the zero-order Bessel function of first kind. A good approximation of the channel statistics is obtained with an order two autoregressive process [13] of the following form:

$$c_k^i = \phi_1 c_{k-1}^i + \phi_2 c_{k-2}^i + \pi_k^i, \quad i = 0, \ldots, L, \tag{3}$$

by letting

$$\phi_1 = 2r \cos(\omega), \quad \phi_2 = -r^2, \tag{4}$$

where $r = 0.809^{2\pi B_d T}$ and $\omega = 0.781 \times 2\pi B_d T$. The driving noise term π_k^i is chosen as a zero-mean white complex Gaussian process of variance

$$q = \frac{1 + \phi_2}{1 - \phi_2}\left[(1 - \phi_2)^2 - \phi_1^2\right], \tag{5}$$

so as to normalize the channel coefficients to unit variance [14]. Consequently, the channel state, in the case of a time-varying wireless channel, is defined as $\mathbf{x}_k = [c_k^0, c_{k-1}^0, c_k^1, c_{k-1}^1, \ldots, c_k^L, c_{k-1}^L]^T$ and evolves according to the dynamical system

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \boldsymbol{\pi}_k, \tag{6}$$

where state transition matrix is given by

$$\mathbf{F} = \begin{bmatrix} \boldsymbol{\Phi} & \mathbf{0}_2 & \ldots & \mathbf{0}_2 \\ \mathbf{0}_2 & \boldsymbol{\Phi} & \ldots & \mathbf{0}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \ldots & \boldsymbol{\Phi} \end{bmatrix}, \quad \boldsymbol{\Phi} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}, \tag{7}$$

and the process noise vector is given by

$$\boldsymbol{\pi}_k = \left[\pi_k^0, 0, \pi_k^1, 0, \ldots, \pi_k^L, 0\right]^T. \tag{8}$$

The received complex noisy observation at instant k has the following form:

$$y_k = \sum_{i=0}^L c_k^i b_{k-i} + n_k, \tag{9}$$

where n_k is a complex white Gaussian noise sample with single-sided power spectral density N_0.

We define the ISI state s_k as the decimal form of the binary subsequence $[b_k, b_{k-1}, \ldots, b_{k-L+1}]^T$, which can take 2^L

discrete values. Let f_k denote the ISI state transition function defined by the following relation:

$$s_k = f_k(s_{k-1}, b_k). \tag{10}$$

It is well known that f_k can be represented graphically by a trellis diagram containing 2^L states [7].

For a time-varying channel, it is now clear from (6) and (9) that our communication system can be described as a mixed discrete-continuous state space of the form $(s_k, \mathbf{x}_k)$, whose dynamics are given by

$$s_k = f_k(s_{k-1}, b_k),$$
$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \boldsymbol{\pi}_k, \tag{11}$$
$$y_k = \mathbf{h}_k(s_{k-1}, s_k)^T \mathbf{x}_k + n_k.$$

The second equation in (11) can be interpreted as a linear dynamical system with state transition matrix given by (7) and zero-mean Gaussian process noise, with covariance matrix

$$\mathbf{Q} = q\mathbf{I}_{L+1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{12}$$

The third equation in (11) can be interpreted as a linear observation process, where the observation matrix $\mathbf{h}_k(s_{k-1}, s_k)$ has the form

$$\mathbf{h}_k(s_{k-1}, s_k) = [b_k, 0, b_{k-1}, 0, \ldots, b_{k-L}, 0]^T. \tag{13}$$

From (1) and (9), a slightly different state-space representation is obtained for the static channel as a special case of (11) by choosing $\mathbf{F} = \mathbf{I}_{L+1}$, $\mathbf{Q} = \mathbf{O}_{L+1}$, and

$$\mathbf{h}_k(s_{k-1}, s_k) = [b_k, b_{k-1}, \ldots, b_{k-L}]^T. \tag{14}$$

3. Blind SISO Equalization Using a Gaussian Mixture Approach

In this section, we derive a fixed-interval smoother by propagating a mixture of N Gaussians per ISI state forward and backward in the ISI trellis. Consequently, the ISI state s_k and the channel state $\mathbf{x}_k$ will be jointly estimated. Finally, the desired *a posteriori* probabilities for the bits b_k are obtained by a simple marginalization step.

3.1. Forward Filtering. A recursive expression of $p(s_k, \mathbf{x}_k \mid y_{1:k})$, where $y_{1:k} = (y_1, y_2, \ldots, y_k)$ is obtained by noting that

$$p(s_k, \mathbf{x}_k, y_{1:k}) = \sum_{s_{k-1}} p(s_k \mid s_{k-1}) p(y_k \mid \mathbf{h}_k(s_{k-1}, s_k), \mathbf{x}_k)$$
$$\times \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(s_{k-1}, \mathbf{x}_{k-1}, y_{1:k-1}) d\mathbf{x}_{k-1}, \tag{15}$$

where the discrete summation extends over the states s_{k-1}, for which a valid transition (s_{k-1}, s_k) exists. In general, the multiplications and integration in (15) cannot be expressed in closed form, therefore, we introduce the following Gaussian mixture parameterization at instant $k-1$:

$$p(s_{k-1}, \mathbf{x}_{k-1}, y_{1:k-1})$$
$$= \sum_{i=1}^N \alpha^i(s_{k-1}) \mathcal{N}_\mathbb{C}\left(\mathbf{x}_{k-1} : \mathbf{x}_{k-1|k-1}^i(s_{k-1}), \mathbf{P}_{k-1|k-1}^i(s_{k-1})\right). \tag{16}$$

In (16), each discrete state s_{k-1} is associated with a mixture of N Gaussians, where N is a design parameter of choice.

Theorem 1. *A closed form expression of $p(s_k, \mathbf{x}_k, y_{1:k})$ is obtained as follows:*

$$p(s_k, \mathbf{x}_k, y_{1:k})$$
$$= \sum_{s_{k-1}} \sum_{i=1}^N \alpha^i(s_{k-1}, s_k) \mathcal{N}_\mathbb{C}\left(\mathbf{x}_k : \mathbf{x}_{k|k}^i(s_{k-1}, s_k), \mathbf{P}_{k|k}^i(s_{k-1}, s_k)\right), \tag{17}$$

where the means $\mathbf{x}_{k|k}^i(s_{k-1}, s_k)$ and covariance matrices $\mathbf{P}_{k|k}^i(s_{k-1}, s_k)$ associated with the state transition (s_{k-1}, s_k) are obtained from the following recursions:

$$\mathbf{x}_{k|k-1}^i(s_{k-1}) = \mathbf{F}\mathbf{x}_{k-1|k-1}^i(s_{k-1}),$$

$$\mathbf{P}_{k|k-1}^i(s_{k-1}) = \mathbf{F}\mathbf{P}_{k-1|k-1}^i(s_{k-1})\mathbf{F}^H + \mathbf{Q},$$

$$\mathbf{K}_k^i(s_{k-1}, s_k)$$
$$= \mathbf{P}_{k|k-1}^i(s_{k-1})\mathbf{h}_k(s_{k-1}, s_k)^*$$
$$\times \left(\mathbf{h}_k(s_{k-1}, s_k)^T \mathbf{P}_{k|k-1}^i(s_{k-1})\mathbf{h}_k(s_{k-1}, s_k)^* + N_0\right)^{-1},$$

$$\mathbf{x}_{k|k}^i(s_{k-1}, s_k)$$
$$= \mathbf{x}_{k|k-1}^i(s_{k-1}) + \mathbf{K}_k^i(s_{k-1}, s_k)$$
$$\times \left(y_k - \mathbf{h}_k(s_{k-1}, s_k)^T \mathbf{x}_{k|k-1}^i(s_{k-1})\right),$$

$$\mathbf{P}_{k|k}^i(s_{k-1}, s_k)$$
$$= \mathbf{P}_{k|k-1}^i(s_{k-1}) - \mathbf{K}_k^i(s_{k-1}, s_k)\mathbf{h}_k(s_{k-1}, s_k)^T$$
$$\times \mathbf{P}_{k|k-1}^i(s_{k-1}), \tag{18}$$

and the weights $\alpha^i(s_{k-1}, s_k)$ are given by

$$\alpha^i(s_{k-1}, s_k) = \alpha^i(s_{k-1})p(s_k \mid s_{k-1})$$
$$\times \mathcal{N}_\mathbb{C}\left(y_k : \mathbf{h}_k(s_{k-1}, s_k)^T \mathbf{x}_{k|k-1}^i(s_{k-1}),\right.$$
$$\mathbf{h}_k(s_{k-1}, s_k)^T \mathbf{P}_{k|k-1}^i(s_{k-1})$$
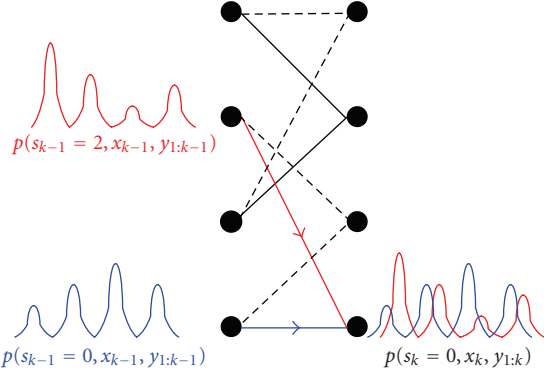$$\left. \times \mathbf{h}_k(s_{k-1}, s_k)^* + N_0\right). \tag{19}$$

FIGURE 3: Example of forward propagation of Gaussian mixtures (with $N = 4$) on a 4-state trellis.

Proof. Injecting (16) into (15), one obtains

$$
\begin{aligned}
&p(s_k, \mathbf{x}_k, y_{1:k}) \\
&= \sum_{s_{k-1}} \sum_{i=1}^{N} \alpha^i(s_{k-1}) p(s_k \mid s_{k-1}) p(y_k \mid \mathbf{h}_k(s_{k-1}, s_k), \mathbf{x}_k) \\
&\quad \times \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\left(\mathbf{x}_{k-1} : \mathbf{x}_{k-1|k-1}^i(s_{k-1}), \mathbf{P}_{k-1|k-1}^i(s_{k-1})\right) d\mathbf{x}_{k-1}.
\end{aligned}
\tag{20}
$$

In the above expression, we easily recognize the integral as the well-known prediction step of Kalman filtering [14]. Moreover, the multiplication by $p(y_k \mid \mathbf{h}_k(s_{k-1}, s_k), \mathbf{x}_k)$ is the correction step of Kalman filtering. Therefore, $p(s_k, \mathbf{x}_k, y_{1:k})$ can be written as (17). □

Figure 3 illustrates how the Gaussian mixture $p(s_k = 0, \mathbf{x}_k, y_{1:k})$ is computed on a 4-state trellis. The components of the Gaussian mixtures $p(s_{k-1} = 0, \mathbf{x}_{k-1}, y_{1:k-1})$ and $p(s_{k-1} = 2, \mathbf{x}_{k-1}, y_{1:k-1})$ undergo a Kalman prediction and correction given the hypothesized data symbol on the valid trellis branch $(s_{k-1} = 0, s_k = 0)$ and $(s_{k-1} = 2, s_k = 0)$, respectively. The resulting Gaussian mixture $p(s_k = 0, \mathbf{x}_k, y_{1:k})$ is obtained as a weighted sum of the resulting mixtures.

3.2. Complexity Reduction Algorithm (CRA).

A problem with (17) is that each discrete state s_k is now associated with a mixture of more than N Gaussians. This means that the number of terms in the Gaussian mixture will grow with time. In order to keep the computational complexity constant for each time instant, we need to approximate the exact expression given by (17) as follows:

$$
p(s_k, \mathbf{x}_k, y_{1:k}) \approx \sum_{i=1}^{N} \alpha^i(s_k) \mathcal{N}_{\mathbb{C}}\left(\mathbf{x}_k : \mathbf{x}_{k|k}^i(s_k), \mathbf{P}_{k|k}^i(s_k)\right), \tag{21}
$$

so that again N Gaussians with weight $\alpha^i(s_k)$, mean $\mathbf{x}_{k|k}^i(s_k)$, and covariance $\mathbf{P}_{k|k}^i(s_k)$, $i = 1, \ldots, N$ are associated with each

state s_k, as in (16). We do this by applying the CRA proposed in [15]. Assume that $\mathcal{N}_1$ (resp. $\mathcal{N}_2$) is a multivariate Gaussian, whose weight, mean, and covariance are given by w_1, $\mathbf{x}_1$, and $\mathbf{P}_1$ (resp. w_2, $\mathbf{x}_2$, $\mathbf{P}_2$). In [15], a practical measure of similarity between the two densities is given by

$$
D = w_1 w_2 [\mathcal{I}(\mathcal{N}_1 \| \mathcal{N}_2) + \mathcal{I}(\mathcal{N}_2 \| \mathcal{N}_1)], \tag{22}
$$

where $\mathcal{I}(\cdot \| \cdot)$ denotes the Kullback-Leibler distance. Then, pairs of similar Gaussians with minimal D are repeatedly merged until N Gaussians remain using the following approximation:

$$
w_1 \mathcal{N}_{\mathbb{C}}(\mathbf{x}_k : \mathbf{x}_1, \mathbf{P}_1) + w_2 \mathcal{N}_{\mathbb{C}}(\mathbf{x}_k : \mathbf{x}_2, \mathbf{P}_2) \approx w \mathcal{N}_{\mathbb{C}}(\mathbf{x}_k : \mathbf{x}, \mathbf{P}), \tag{23}
$$

where

$$
\begin{aligned}
w &= w_1 + w_2, \\
\mathbf{x} &= \frac{w_1 \mathbf{x}_1 + w_2 \mathbf{x}_2}{w_1 + w_2}, \\
\mathbf{P} &= \frac{w_1}{w_1 + w_2}\left[\mathbf{P}_1 + (\mathbf{x}_1 - \mathbf{x})(\mathbf{x}_1 - \mathbf{x})^H\right] \\
&\quad + \frac{w_2}{w_1 + w_2}\left[\mathbf{P}_2 + (\mathbf{x}_2 - \mathbf{x})(\mathbf{x}_2 - \mathbf{x})^H\right].
\end{aligned}
\tag{24}
$$

3.3. Backward Filtering.

Let T denote the total number of available observations and $y_{k+1:T} = (y_{k+1}, y_{k+2}, \ldots, y_T)$. A time-reversed version of the forward filter in Section 3.1 can also be derived. We seek a recursive expression of the likelihood $p(y_{k+1:T} \mid s_k, \mathbf{x}_k)$, propagated backward in time. We obtain the following recursion:

$$
\begin{aligned}
&p(y_{k+1:T} \mid s_k, \mathbf{x}_k) \\
&= \sum_{s_{k+1}} p(s_{k+1} \mid s_k) \times \int p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) \\
&\quad \times p(y_{k+1} \mid \mathbf{h}_{k+1}(s_k, s_{k+1}), \mathbf{x}_{k+1}) \\
&\quad \times p(y_{k+2:T} \mid s_{k+1}, \mathbf{x}_{k+1}) d\mathbf{x}_{k+1},
\end{aligned}
\tag{25}
$$

where the discrete summation extends over the states s_{k+1}, for which a valid transition (s_k, s_{k+1}) exists.

Theorem 2. *Assume that the following Gaussian mixture parameterization:*

$$
\begin{aligned}
&p(y_{k+2:T} \mid s_{k+1}, \mathbf{x}_{k+1}) \\
&= \sum_{i=1}^{N} \beta^i(s_{k+1}) \mathcal{N}_{\mathbb{C}}\left(\mathbf{x}_{k+1} : \mathbf{x}_{k+1|k+2:T}^i(s_{k+1}), \mathbf{P}_{k+1|k+2:T}^i(s_{k+1})\right)
\end{aligned}
\tag{26}
$$

is available at instant $k + 1$, a closed form expression of $p(y_{k+1:T}s_k, \mathbf{x}_k)$ is obtained as

$$p(y_{k+1:T} \mid s_k, \mathbf{x}_k)$$

$$= \sum_{s_{k+1}} \sum_{i=1}^{N} \beta^i(s_k, s_{k+1}) \mathcal{N}_{\mathbb{C}} \left(\mathbf{x}_k : \mathbf{x}_{k|k+1:T}^i(s_k, s_{k+1}), \right. \quad (27)$$

$$\left. \mathbf{P}_{k|k+1:T}^i(s_k, s_{k+1}) \right),$$

where the means $\mathbf{x}_{k|k+1:T}^i(s_k, s_{k+1})$ and covariance matrices $\mathbf{P}_{k|k+1:T}^i(s_k, s_{k+1})$ associated with the state transition (s_k, s_{k+1}) are obtained from the following recursions:

$$\mathbf{K}_{k+1}^i(s_k, s_{k+1})$$

$$= \mathbf{P}_{k+1|k+2:T}^i(s_{k+1}) \mathbf{h}_{k+1}(s_k, s_{k+1})^*$$

$$\times \left(\mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{P}_{k+1|k+2:T}^i(s_{k+1}) \mathbf{h}_{k+1}(s_k, s_{k+1})^* + N_0 \right)^{-1},$$

$$\mathbf{x}_{k+1|k+1:T}^i(s_k, s_{k+1})$$

$$= \mathbf{x}_{k+1|k+2:T}^i(s_{k+1}) + \mathbf{K}_{k+1}^i(s_k, s_{k+1})$$

$$\times \left(y_{k+1} - \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{x}_{k+1|k+2:T}^i(s_{k+1}) \right),$$

$$\mathbf{P}_{k+1|k+1:T}^i(s_k, s_{k+1})$$

$$= \mathbf{P}_{k+1|k+2:T}^i(s_{k+1}) - \mathbf{K}_{k+1}^i(s_k, s_{k+1})$$

$$\times \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{P}_{k+1|k+2:T}^i(s_{k+1}),$$

$$\mathbf{x}_{k|k+1:T}^i(s_k, s_{k+1}) = \mathbf{F} \mathbf{x}_{k+1|k+1:T}^i(s_k, s_{k+1}),$$

$$\mathbf{P}_{k|k+1:T}^i(s_k, s_{k+1}) = \mathbf{F} \mathbf{P}_{k+1|k+1:T}^i(s_k, s_{k+1}) \mathbf{F}^H + \mathbf{Q}, \quad (28)$$

and the weights $\beta^i(s_k, s_{k+1})$ are given by

$$\beta^i(s_k, s_{k+1})$$

$$= \beta^i(s_{k+1}) p(s_{k+1} \mid s_k)$$

$$\times \mathcal{N}_{\mathbb{C}} \left(y_{k+1} : \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{x}_{k+1|k+2:T}^i(s_{k+1}), \right.$$

$$\left. \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{P}_{k+1|k+2:T}^i(s_{k+1}) \mathbf{h}_{k+1}(s_k, s_{k+1})^* + N_0 \right). \quad (29)$$

The proof is obtained by injecting (26) into (25) and using the same arguments as in the demonstration of Theorem 1.

Figure 4 illustrates how the Gaussian mixture $p(y_{k+1:T} \mid s_k = 1, \mathbf{x}_k)$ is computed on a 4-state trellis. The components of the Gaussian mixtures $p(y_{k+2:T} \mid s_{k+1} = 2, \mathbf{x}_{k+1})$ and $p(y_{k+2:T} \mid s_{k+1} = 3, \mathbf{x}_{k+1})$ undergo a Kalman correction and backward prediction given the hypothesized data symbol on the valid trellis branch $(s_k = 1, s_{k+1} = 2)$ and $(s_k = 1, s_{k+1} = 3)$, respectively. The resulting Gaussian mixture $p(y_{k+1:T} \mid s_k = 1, \mathbf{x}_k)$ is obtained as a weighted sum of the resulting mixtures.
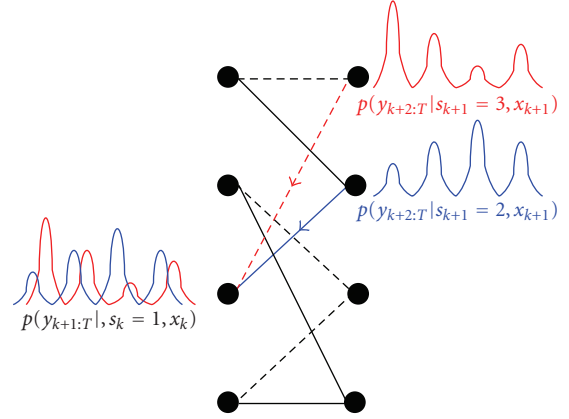

FIGURE 4: Example of backward propagation of Gaussian mixtures (with $N = 4$) on a 4-state trellis.

Again, we need to apply the CRA of Section 3.2. Complexity reduction algorithm to (27), so that $p(y_{k+1:T} \mid s_k, \mathbf{x}_k)$ admits the desired form

$$p(y_{k+1:T} \mid s_k, \mathbf{x}_k)$$

$$\approx \sum_{i=1}^{N} \beta^i(s_k) \mathcal{N}_{\mathbb{C}} \left(\mathbf{x}_k : \mathbf{x}_{k|k+1:T}^i(s_k), \mathbf{P}_{k|k+1:T}^i(s_k) \right). \quad (30)$$

3.4. Smoothing. A two-filter smoothing formula is obtained as follows:

$$p(s_k, s_{k+1}, \mathbf{x}_k, y_{1:T})$$

$$= p(s_{k+1} \mid s_k) p(s_k, \mathbf{x}_k, y_{1:k})$$

$$\times \int p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) p(y_{k+1} \mid \mathbf{h}_{k+1}(s_k, s_{k+1}), \mathbf{x}_{k+1}) \quad (31)$$

$$\times p(y_{k+2:T} \mid s_{k+1}, \mathbf{x}_{k+1}) d\mathbf{x}_{k+1}.$$

Theorem 3. *Using the Gaussian mixture approximations for the forward and the backward filter introduced in Sections 3.1 and 3.3, respectively, a closed form expression of $p(s_k, s_{k+1}, \mathbf{x}_k, y_{1:T})$ is obtained as follows:*

$$p(s_k, s_{k+1}, \mathbf{x}_k, y_{1:T})$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma^{i,j}(s_k, s_{k+1}) \mathcal{N}_{\mathbb{C}} \left(\mathbf{x}_k : \mathbf{x}_{k|1:T}^{i,j}(s_k, s_{k+1}), \mathbf{P}_{k|1:T}^{i,j}(s_k, s_{k+1}) \right), \quad (32)$$

where the covariances associated to transition (s_k, s_{k+1}) are

$$\mathbf{P}_{k|1:T}^{i,j}(s_k, s_{k+1})$$

$$= \mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1}) \times \left[\mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1}) + \mathbf{P}_{k|k}^i(s_k) \right]^{-1}$$

$$\times \mathbf{P}_{k|k}^i(s_k), \quad (33)$$

and the means associated to transition (s_k, s_{k+1}) are

$$
\begin{aligned}
&\mathbf{x}_{k|1:T}^{i,j}(s_k, s_{k+1}) \\
&= \mathbf{P}_{k|1:T}^{i,j}(s_k, s_{k+1}) \times \left[\mathbf{P}_{k|k}^i(s_k)^{-1} \mathbf{x}_{k|k}^i(s_k) + \mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1})^{-1} \right. \\
&\qquad\qquad \left. \times \mathbf{x}_{k|k+1:T}^j(s_k, s_{k+1}) \right],
\end{aligned}
\tag{34}
$$

for $1 \le i, j \le N$. The expression of the weights is given by

$$
\begin{aligned}
&\sigma^{i,j}(s_k, s_{k+1}) \\
&= \alpha^i(s_k)\beta^j(s_{k+1})p(s_{k+1} \mid s_k)b^{i,j}(s_k, s_{k+1}) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\Big(y_{k+1} : \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{x}_{k+1|k+2:T}^j(s_{k+1}), \\
&\qquad \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{P}_{k+1|k+2:T}^j(s_{k+1})\mathbf{h}_{k+1}(s_k, s_{k+1})^* + N_0 \Big).
\end{aligned}
\tag{35}
$$

The coefficient $b^{i,j}(s_k, s_{k+1})$ has the following form:

$$
\begin{aligned}
&b^{i,j}(s_k, s_{k+1}) \\
&= \frac{1}{\pi^d \det\left[\mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1}) + \mathbf{P}_{k|k}^i(s_k) \right]} \\
&\quad \times \exp\Big\{ -\left[\mathbf{x}_{k|k}^i(s_k) - \mathbf{x}_{k|k+1:T}^j(s_k, s_{k+1}) \right]^H \\
&\qquad \left[\mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1}) + \mathbf{P}_{k|k}^i(s_k) \right]^{-1} \\
&\qquad \left[\mathbf{x}_{k|k}^i(s_k) - \mathbf{x}_{k|k+1:T}^j(s_k, s_{k+1}) \right] \Big\},
\end{aligned}
\tag{36}
$$

where d denotes the dimension of the continuous valued state variable.

Proof. In (31), the term $p(s_k, \mathbf{x}_k, y_{1:k})$ has been calculated as (21) and the integral, also appearing in (25), has already been computed as

$$
\begin{aligned}
&\sum_{j=1}^N \beta^j(s_{k+1}) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\Big(y_{k+1} : \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{x}_{k+1|k+2:T}^j(s_{k+1}), \\
&\qquad \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{P}_{k+1|k+2:T}^j(s_{k+1})\mathbf{h}_{k+1}(s_k, s_{k+1})^* + N_0 \Big) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\Big(\mathbf{x}_k : \mathbf{x}_{k|k+1:T}^j(s_k, s_{k+1}), \mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1}) \Big).
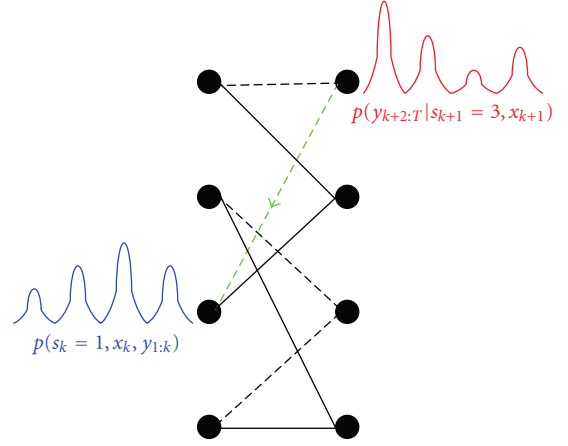\end{aligned}
\tag{37}
$$

FIGURE 5: Illustration of the computation of smoothed Gaussian mixtures (with $N = 4$) on a 4-state trellis.

Therefore, (31) can be rewritten as follows:

$$
\begin{aligned}
&\sum_{i=1}^N \sum_{j=1}^N \alpha^i(s_k)\beta^j(s_{k+1})p(s_{k+1} \mid s_k) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\Big(y_{k+1} : \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{x}_{k+1|k+2:T}^j(s_{k+1}), \\
&\qquad \mathbf{h}_{k+1}(s_k, s_{k+1})^T \mathbf{P}_{k+1|k+2:T}^j(s_{k+1})\mathbf{h}_{k+1}(s_k, s_{k+1})^* + N_0 \Big) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\Big(\mathbf{x}_k : \mathbf{x}_{k|k}^i(s_k), \mathbf{P}_{k|k}^i(s_k) \Big) \\
&\quad \times \mathcal{N}_{\mathbb{C}}\Big(\mathbf{x}_k : \mathbf{x}_{k|k+1:T}^j(s_k, s_{k+1}), \mathbf{P}_{k|k+1:T}^j(s_k, s_{k+1}) \Big).
\end{aligned}
\tag{38}
$$

After straightforward algebraic manipulations on the product of two Gaussian densities, the desired result (32) is obtained. □

Figure 5 illustrates how the Gaussian mixture $p(s_k = 1, s_{k+1} = 3, \mathbf{x}_k, y_{1:T})$ is computed on a 4-state trellis. The components of the Gaussian mixtures $p(y_{k+2:T} \mid s_{k+1} = 3, \mathbf{x}_{k+1})$ undergo a Kalman correction and backward prediction given the hypothesized data symbol on the valid trellis branch ($s_k = 1, s_{k+1} = 3$). The resulting Gaussian mixture is multiplied with the Gaussian mixture $p(s_k = 1, \mathbf{x}_k, y_{1:k})$ computed in the forward pass and by the scalar $p(s_{k+1} = 3 \mid s_k = 1)$, so as to obtain $p(s_k = 1, s_{k+1} = 3, \mathbf{x}_k, y_{1:T})$.

Since we are interested in soft-output equalization, we must compute smoothed bit-by-bit marginal probabilities. Let $B_k^{(m)}$ be the set of state transitions (s_{k-1}, s_k) such that the information bit $b_k = m$, with $m = -1, 1$. Taking (32) at instant $k-1$ and marginalizing out the vector $\mathbf{x}_{k-1}$, we obtain

$$
p(b_k = m \mid y_{1:T}) \propto \sum_{(s_{k-1}, s_k) \in B_k^{(m)}} \sum_{i=1}^N \sum_{j=1}^N \sigma^{i,j}(s_{k-1}, s_k).
\tag{39}
$$

The hard decision can then be written as follows:

$$
\hat{b}_k = \arg\max_{m \in \{-1, +1\}} p(b_k = m \mid y_{1:T}).
\tag{40}
$$

Similarly, the *a posteriori* pdf of the channel vector is obtained as a Gaussian mixture by marginalizing out all possible ISI state transitions

$$p(\mathbf{x}_k \mid y_{1:T}) \propto \sum_{(s_k, s_{k+1})} \sum_{i=1}^N \sum_{j=1}^N \sigma^{i,j}(s_k, s_{k+1})$$

$$\times \mathcal{N}_{\mathbb{C}}\left(\mathbf{x}_k : \mathbf{x}_{k|1:T}^{i,j}(s_k, s_{k+1}), \mathbf{P}_{k|1:T}^{i,j}(s_k, s_{k+1})\right). \tag{41}$$

Under the minimum mean square error (MMSE) criterion, the forward filtered channel vector estimated at instant k is obtained by marginalizing out the ISI state variable

$$\hat{\mathbf{x}}_{k|k} = \int \mathbf{x}_k p(\mathbf{x}_k \mid y_{1:k}) d\mathbf{x}_k \propto \sum_{s_k} \sum_{i=1}^N \alpha^i(s_k) \mathbf{x}_{k|k}^i(s_k). \tag{42}$$

Similarly, the MMSE smoothed estimate of the channel vector at instant k is obtained by marginalizing out all possible ISI state transitions

$$\hat{\mathbf{x}}_{k|1:T} = \int \mathbf{x}_k p(\mathbf{x}_k \mid y_{1:T}) d\mathbf{x}_k$$

$$\propto \sum_{(s_k, s_{k+1})} \sum_{i=1}^N \sum_{j=1}^N \sigma^{i,j}(s_k, s_{k+1}) \mathbf{x}_{k|1:T}^{i,j}(s_k, s_{k+1}). \tag{43}$$

3.5. Complexity Evaluation. It is well known that the complexity of one recursion of the Kalman filter is $\mathcal{O}(d^3)$ [16], where d denotes the dimension of the continuous-valued state estimate. However, in our forward and backward filters, the complexity of one recursion of the Kalman filter reduces to $\mathcal{O}(d^2)$ due to the block diagonal form of $\mathbf{F}$ and the fact that the matrix inversion reduces to a division by a scalar. Thus, the overall complexity per information bit of the forward and backward filter is $\mathcal{O}(2^L(L+1)^2 N)$. The complexity per information bit of the smoothing pass can be evaluated as $\mathcal{O}(2^L(L+1)^3 N^2)$, due to the matrix inversions.

4. Numerical Results

4.1. Comparison with Existing Methods. We consider a memory-2 Rayleigh fading channel simulated with the method introduced in [17]. The standard deviations of the resulting three complex processes (c_k^0, c_k^1, c_k^2) are set at $(0.407, 0.815, 0.407)$. The block size is $B = 100$ bits. As illustrated in Figure 1, a tail of length 2 bits is used, which enables the proposed algorithm to start with the correct initial and final ISI state when processing each frame. This is necessary to remove the $\pm\pi$ phase ambiguity inherent to BPSK modulation. We assume that each data block is affected by an independent channel realization. E_b denotes the average energy per bit.

We compare the bit error rate (BER) of our method with two kinds of blind equalizers. The first kind of blind equalizers consists of Baud-rate linear filters optimized with the constant modulus algorithm (CMA) [2] or with the first cost function (FCF) introduced in [18]. These equalizers
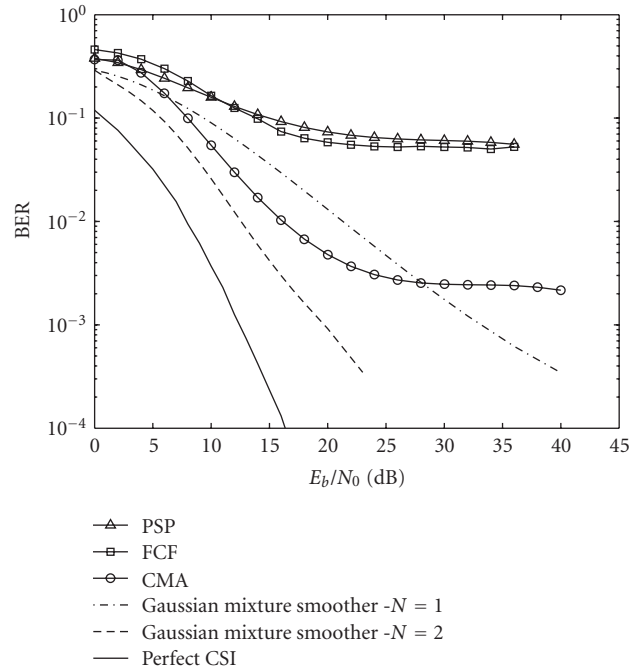

FIGURE 6: BER performances of blind equalization at $B_d T = 0$.

are iterated 50 times back and forth on each data block, in order to aleviate the slow convergence problem [19]. There is also the issue of the ambiguities inherent to these blind equalizers. Differential encoding of the transmitted data is used to solve the phase ambiguity problem. Also, a length-10 known preamble is used to resolve the delay ambiguity. The lengths of the CMA and FCF equalizers were optimized by hand, to 5 and 15 coefficients, respectively. The second kind of blind equalizer is the per-survivor processing (PSP) algorithm [20], which is similar in spirit to the proposed method, since it is a trellis-based algorithm operating on the conventional 4-state ISI trellis [7] and using Kalman filtering for channel estimation. However, since the path pruning strategy employs the Viterbi algorithm [7], it is not an SISO method.

Figure 6 illustrates the BER of the different equalizers as a function of E_b/N_0 for a static Rayleigh fading channel ($B_d T = 0$). The CMA and FCF equalizers reach an error floor due to residual ISI. The same phenonenon is observed for the PSP algorithm and is due to the misacquisition problem analysed in [21]. Our Gaussian mixture smoother in the degenerate case, where the channel estimation is performed with only $N = 1$ Gaussian per discrete ISI state, is clearly outperformed by the proposed algorithm with $N = 2$, which attains performances close to perfect channel state information (CSI). Note that our Gaussian mixture smoother with $N = 1$ is essentially the same as [3], but adapted to block transmissions, since fixed-lag smoothing has been replaced by fixed-interval smoothing.

Figure 7 illustrates the channel mean square error (MSE). The forward filtered and the smoothed channel estimates are computed according to (42) and (43), respectively. We note that smoothing provides a dramatic improvement over
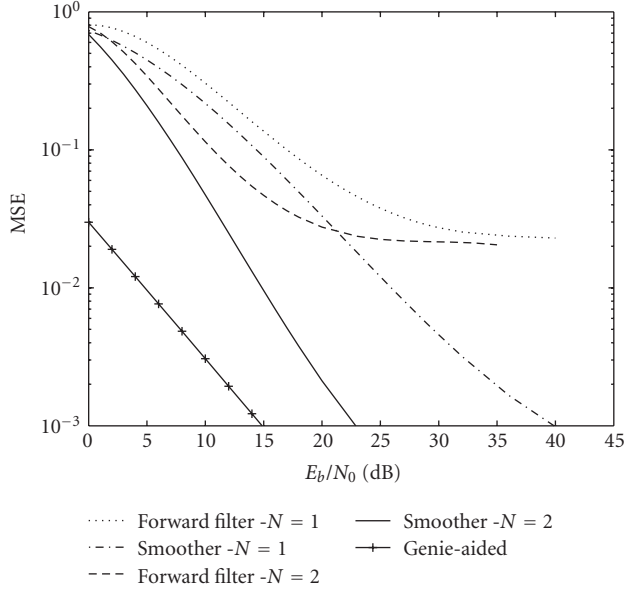
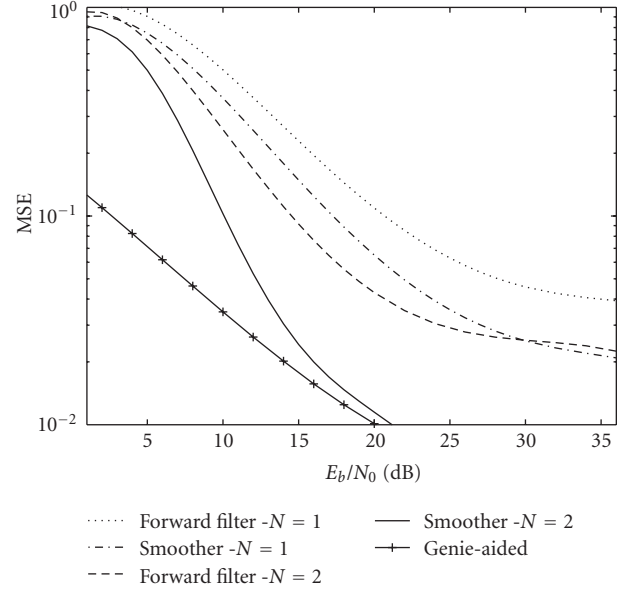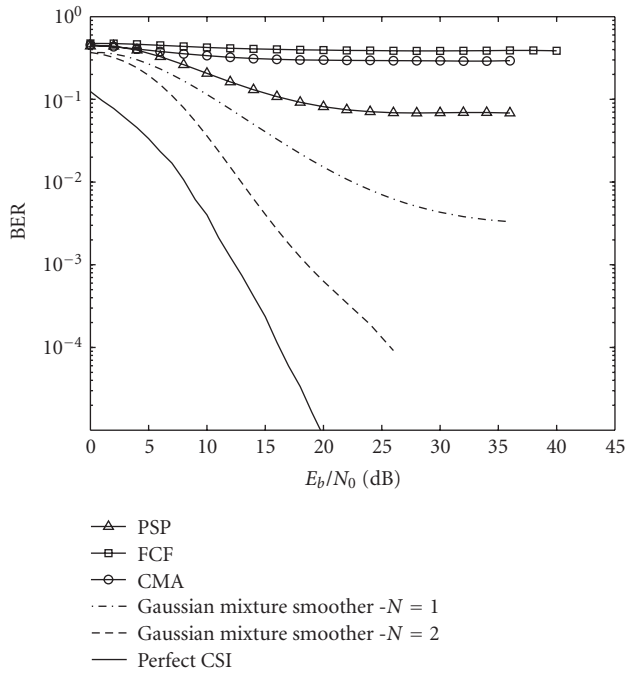FIGURE 7: Channel MSE of blind Gaussian mixture-based equalization at $B_dT = 0$.

FIGURE 8: BER performances of blind equalization at $B_dT = 10^{-2}$.

forward filtering alone both for $N = 1$ and $N = 2$. We interpret this result by the fact that the smoothing pass, by exploiting the knowledge of future observations, is able to correct errors and phase ambiguities due to fading events or occasionally high noise in the forward pass. As a reference, the channel MSE attained by a genie-aided Kalman smoother (with known symbols) is also shown.

We also study a fast-fading scenario with $B_dT = 10^{-2}$, in order to study the robustness of our algorithm against

FIGURE 9: Channel MSE of blind Gaussian mixture-based equalization at $B_dT = 10^{-2}$.

TABLE 1: Saleh-Valenzuela model parameters.

Parameter and units	Notation and numerical value
Intercluster arrival rate (1/s)	$\Lambda = 5/T$
Intercluster decay constant (s)	$\Gamma = 0.6T$
Intracluster arrival rate (1/s)	$\lambda = 80/T$
Intracluster decay constant (s)	$\gamma = 0.35T$

a large Doppler spread. Figures 8 and 9 illustrate the BER and the channel MSE, respectively. Considering the BER performances, we observe that the CMA, FCF, and PSP equalizers as well as the Gaussian mixture smoother with $N = 1$ reach an error floor, while the performances of the Gaussian mixture smoother with $N = 2$ are still satisfactory. Again, the channel MSE after the smoothing pass is much smaller than after forward filtering alone.

4.2. Performances on a Realistic Channel Model. A realistic model for quasi-static indoor multipath propagation is given by the Saleh-Valenzuela channel model [22], whose impulse response is given by

$$h(t) = \sum_i h_i \delta(t - \tau_i). \qquad (44)$$

The τ_i represent the propagation delays of the rays, which arrive in clusters. These clusters and the rays within each cluster form Poisson arrival processes, with different but fixed rates. The h_i are independent complex zero-mean Gaussian coefficients whose variances decay exponentially with the cluster and ray delay. In our simulations, the parameters of the Saleh-Valenzuela channel model are given in Table 1. Therefore, assuming rectangular transmit and receive pulse shaping as well as perfect symbol and frame
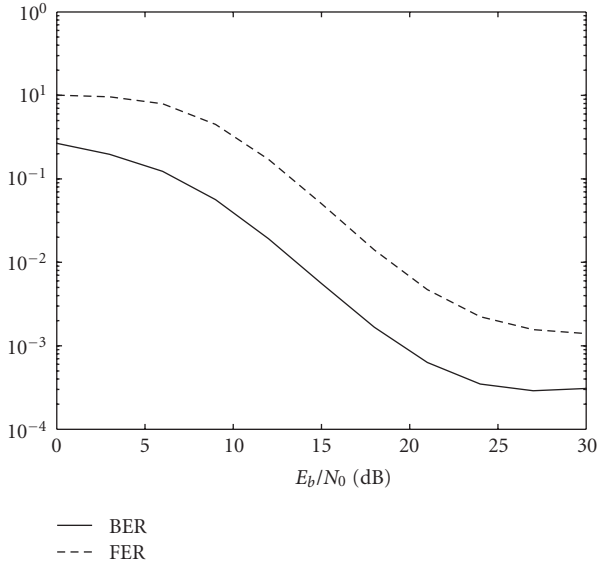
FIGURE 10: BER and FER performances of blind SISO equalization on a quasi-static Saleh-Valenzuela channel model.

synchronization, we obtain the equivalent discrete channel model described in Figure 2 with

$$
c_k^0 = \sum_{i:0 \le \tau_i < T} h_i \left(1 - \frac{\tau_i}{T} \right),
$$

$$
c_k^1 = \sum_{i:0 \le \tau_i < T} h_i \frac{\tau_i}{T} + \sum_{i:T \le \tau_i < 2T} h_i \left(1 - \frac{\tau_i - T}{T} \right), \quad (45)
$$

$$
c_k^2 = \sum_{i:T \le \tau_i < 2T} h_i \frac{\tau_i - T}{T} + \sum_{i:2T \le \tau_i < 3T} h_i \left(1 - \frac{\tau_i - 2T}{T} \right).
$$

Here, the equivalent discrete channel model is truncated to $L = 2$, thus, all existing rays with propagation delay $\tau_i \ge 2T$ will generate unmodeled ISI. Using the fact that the h_i are statistically independent, we obtain

$$
E\left[c_k^0 c_k^{1*} \right] = \sum_{i:0 \le \tau_i < T} E\left[|h_i|^2 \right] \left(1 - \frac{\tau_i}{T} \right) \frac{\tau_i}{T},
$$

$$
E\left[c_k^1 c_k^{2*} \right] = \sum_{i:T \le \tau_i < 2T} E\left[|h_i|^2 \right] \left(1 - \frac{\tau_i - T}{T} \right) \frac{\tau_i - T}{T}, \quad (46)
$$

$$
E\left[c_k^0 c_k^{2*} \right] = 0.
$$

In general, the taps of the discrete-time channel model are correlated by the transmit and receive pulse shapes and the correlation depends on the power delay profile (PDP), which is of course unknown to the receiver. Therefore, the standard hypothesis of statistical independence of the channel coefficients $c_k^i, i = 0, \ldots, L$ in Section 2, corresponds to neglecting the correlations in (46).

In Figure 10, we test the Gaussian mixture smoother with $N = 2$, for a Saleh-Valenzuela channel model with parameters given in Table 1. We observe that the bit error rate (BER) and the frame error rate (FER) exhibit an error

floor. This floor can be interpreted as the result of ignoring the correlation between the channel coefficients and to unmodeled ISI due to the fact that the equivalent discrete channel model is truncated to a memory-2 channel.

5. Conclusions

In this paper, we introduced a new blind receiver for soft-output equalization. The proposed method, adapted from an earlier work on fixed-interval blind MIMO demodulation, is well suited for block transmissions. In essence, the algorithm propagates Gaussian mixtures forward and backward in the conventional ISI trellis, in order to perform joint ISI state and channel estimation. Numerical simulations showed that the proposed method outperforms several well-known blind equalization schemes and works well even in fast-fading scenarios.

Future extensions of this work include the application of blind Gaussian mixture-based equalization to Rician fading and higher-order modulations. Since the proposed algorithm is soft output in nature, its application to turbo equalization will also be investigated.

References

[1] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Transactions on Communications*, vol. 23, no. 6, pp. 679–682, 1975.

[2] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, 1980.

[3] R. A. Iltis, J. J. Shynk, and K. Giridhar, "Bayesian algorithms for blind equalization using parallel adaptive filtering," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1017–1032, 1994.

[4] G. Lee, S. B. Gelfand, and M. P. Fitz, "Bayesian decision feedback techniques for deconvolution," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 1, pp. 155–166, 1995.

[5] A. Anastasopoulos and K. M. Chugg, "Adaptive soft-input soft-output algorithms for iterative detection with parametric uncertainty," *IEEE Transactions on Communications*, vol. 48, no. 10, pp. 1638–1649, 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 Transactions on Communications*, vol. 49, no. 12, pp. 2106–2114, 2001.

[7] G. D. Forney Jr., "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.

[8] J. Gunther, D. Keller, and T. Moon, "A generalized BCJR algorithm and its use in iterative blind channel identification," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 661–664, 2007.

[9] F. Lehmann, "Blind soft-output decoding of space-time trellis coded transmissions over time-varying Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 2088–2099, 2009.

[10] F. Lehmann, "Blind soft-output equalization of block-oriented wireless communications," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09)*, pp. 2801–2804, Tapei, Taiwan, April 2009.

[11] Y. Zheng, Z. Lin, and Y. Ma, "Recursive blind identification of non-Gaussian time-varying AR model and application to blind equalisation of time-varying channel," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 148, no. 4, pp. 275–282, 2001.

[12] G. L. Stüber, *Principles of Mobile Communications*, Kluwer Academic Publishers, Norwell, Mass, USA, 1999.

[13] F. Lehmann, "Blind estimation and detection of space-time trellis coded transmissions over the Rayleigh fading MIMO channel," *IEEE Transactions on Communications*, vol. 56, no. 3, pp. 334–338, 2008.

[14] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 2002.

[15] G. Kitagawa, "The two-filter formula for smoothing and an implementation of the Gaussian-sum smoother," *Annals of the Institute of Statistical Mathematics*, vol. 46, no. 4, pp. 605–623, 1994.

[16] F. Daum, "Nonlinear filters: beyond the kalman filter," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 57–69, 2005.

[17] Y. Li and X. Huang, "The simulation of independent Rayleigh faders," *IEEE Transactions on Communications*, vol. 50, no. 9, pp. 1503–1514, 2002.

[18] J. Sala-Alvarez and G. Vázquez-Grau, "Statistical reference criteria for adaptive signal processing in digital communications," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 14–31, 1997.

[19] B.-J. Kim and D. C. Cox, "Blind equalization for short burst wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 4, pp. 1235–1247, 2000.

[20] R. Raheli, A. Polydoros, and C. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environments," *IEEE Transactions on Communications*, vol. 43, no. 2, pp. 354–364, 1995.

[21] K. M. Chugg, "Blind acquisition characteristics of PSP-based sequence detectors," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1518–1529, 1998.

[22] A. A. M. Saleh and R. A. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, 1987.