Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2006, Article ID 29075, Pages 1–13
DOI 10.1155/WCN/2006/29075

Equalization of Sparse Intersymbol-Interference Channels Revisited

Jan Mietzner,[1] Sabah Badri-Hoeher,[1] Ingmar Land,[2] and Peter A. Hoeher[1]

[1] *Information and Coding Theory Lab (ICT), Faculty of Engineering, University of Kiel, Kaiserstrasse 2, 24143 Kiel, Germany*
[2] *Department of Communication Technology, Digital Communications Division, Aalborg University, Frederik Bajers Vej 7, A3, Aalborg East 9220, Denmark*

Received 18 April 2005; Revised 12 January 2006; Accepted 28 February 2006

Recommended for Publication by Brian Sadler

Sparse intersymbol-interference (ISI) channels are encountered in a variety of communication systems, especially in high-data-rate systems. These channels have a large memory length, but only a small number of significant channel coefficients. In this paper, equalization of sparse ISI channels is revisited with focus on trellis-based techniques. Due to the large channel memory length, the complexity of maximum-likelihood sequence estimation by means of the Viterbi algorithm is normally prohibitive. In the first part of the paper, a unified framework based on factor graphs is presented for complexity reduction without loss of optimality. In this new context, two known reduced-complexity trellis-based techniques are recapitulated. In the second part of the paper a simple alternative approach is investigated to tackle general sparse ISI channels. It is shown that the use of a linear filter at the receiver renders the application of standard reduced-state trellis-based equalization techniques feasible without significant loss of optimality.

1. INTRODUCTION

Sparse intersymbol-interference (ISI) channels are encountered in a wide range of communication systems, such as aeronautical/satellite communication systems or high-data-rate mobile radio systems (especially in hilly terrain, where the delay spread is large). For mobile radio applications, fading channels are of particular interest [1]. The equivalent discrete-time channel impulse response (CIR) of a sparse ISI channel has a large channel memory length, but only a small number of significant channel coefficients.

Due to the large memory length, equalization of sparse ISI channels with a reasonable complexity is a demanding task. The topics of linear and decision-feedback equalization (DFE) for sparse ISI channels are, for example, addressed in [2], where the sparse structure of the channel is explicitly utilized for the design of the corresponding finite-impulse-response (FIR) filter(s). DFE for sparse channels is also considered in [3–6].

Trellis-based equalization for sparse channels is addressed in [7–10]. The complexity in terms of trellis states of an optimal trellis-based equalizer algorithm, based on the Viterbi algorithm (VA) [11] or the Bahl-Cocke-Jelinek-Raviv algorithm (BCJRA)[1] [12], is normally prohibitive for sparse ISI channels, because it grows exponentially with the channel memory length. However, reduced-complexity algorithms can be derived by exploiting the sparseness of the channel. In [7], it is observed that given a sparse channel, there is only a comparably small number of possible branch metrics within each trellis segment. By avoiding to compute the same branch metric several times, the computational complexity is reduced significantly without loss of optimality. However, the complexity in terms of trellis states remains the same. As an alternative, another equalizer concept called multitrellis Viterbi algorithm (M-VA) is proposed in [7] which is based on multiple parallel *irregular* trellises (i.e., time-variant trellises). The M-VA is claimed to be optimal while having a significantly reduced computational complexity and number of trellis states.

[1] The VA is optimal in the sense of maximum-likelihood sequence estimation (MLSE) and the BCJRA in the sense of maximum a posteriori (MAP) symbol-by-symbol estimation. The VA and the BCJRA operate on the same trellis diagram. Therefore, all statements concerning complexity issues apply both for the VA and the BCJRA.

A particularly simple solution to reduce the complexity of the conventional VA without loss of optimality can be found in [8, 9]: the parallel-trellis Viterbi algorithm (P-VA) is based on multiple parallel *regular* trellises. However, it can only be applied for sparse channels with a so-called *zero-pad* structure, where the nonzero channel coefficients are placed on a regular grid. In order to tackle more general sparse channels with a CIR close to a zero-pad channel, it is proposed in [8, 9] to exchange tentative decisions between the parallel trellises and thus cancel residual ISI. This modified version of the P-VA is, however, suboptimal and is denoted as *sub*-P-VA in the sequel.

A generalization of the P-VA and the sub-P-VA can be found in [10], where corresponding algorithms based on the BCJRA are presented. These are in the sequel denoted as parallel-trellis BCJR algorithms (P-BCJRA and sub-P-BCJRA, resp.). Some interesting enhancements of the (sub-)-P-BCJRA are also discussed in [10]. Specifically, it is shown that the performance of the sub-P-BCJRA can be improved by means of minimum-phase prefiltering [13–15].

Alternatives to trellis-based equalization are the tree-based LISS algorithm [16, 17] and the joint Gaussian (JG) approach in [18]. A factor-graph approach [19] for sparse channels, based on the sum-product algorithm, is presented in [20]. Turbo equalization [21] for sparse channels is addressed in [22]. In particular, an efficient trellis-based soft-input soft-output (SISO) equalizer algorithm is considered, which combines ideas of the M-VA and the sub-P-BCJRA. A non-trellis-based equalizer algorithm for fast-fading sparse ISI channels, based on the symbol-by-symbol MAP criterion, is presented in [23].

This paper focuses on trellis-based equalization techniques for sparse ISI channels. In Section 2, a unified framework for complexity reduction without loss of optimality is presented. It is based on factor graphs [19] and might be useful in order to derive new reduced-complexity algorithms for specific sparse ISI channels (see also [20]). Based on this framework, the M-VA and the P-VA are recapitulated. It is shown that the M-VA is, in fact, clearly *suboptimal*. Moreover, it is illustrated why the optimal P-VA can only be applied for zero-pad channels. As a result, there is no *optimal* reduced-complexity trellis-based equalization technique for *general* sparse ISI channels available in the literature. Moreover, since the sub-P-VA requires a CIR structure close to a zero-pad channel, it is of rather limited practical relevance, especially in the case of fading channels.

Little effort has yet been made, in order to compare the performance of the above algorithms with that of standard (suboptimal) reduced-complexity receivers not specifically designed for sparse channels. In Section 3, a simple alternative to the sub-P-VA/sub-P-BCJRA is therefore investigated. Specifically, the idea in [10] to employ prefiltering at the receiver is picked up. It is demonstrated that the use of a linear minimum-phase filter [13–15] renders the application of efficient reduced-state trellis-based equalizer algorithms such as [24, 25] feasible, without significant loss of optimality. As an alternative receiver structure, the use of a linear channel shortening filter [26] is investigated, in conjunction with a conventional VA operating on a shortened channel memory.

The considered receiver structures are notably simple: the employed equalizer algorithms are standard, that is, not specifically designed for sparse channels. (The sparse channel structure is normally lost after prefiltering.) Solely the linear filters are adjusted to the current CIR, which is particularly favorable with regard to fading channels. Moreover, the filter coefficients can be computed using standard techniques available in the literature. In order to illustrate the efficiency of the considered receiver structure, numerical results are presented in Section 4 for various types of sparse ISI channels. Using a minimum-phase filter in conjunction with a delayed decision-feedback sequence estimation (DDFSE) equalizer [25], bit error rates can be achieved that deviate only 1–2 dB from the matched filter bound (at a bit error rate of 10^{-3}). To the authors' best knowledge, similar performance studies for prefiltering in the case of sparse ISI channels have not yet been presented in the literature.

2. COMPLEXITY REDUCTION WITHOUT LOSS OF OPTIMALITY

A *general* sparse ISI channel is characterized by a comparably large channel memory length L, but has only a small number of significant channel coefficients h_g, $g = 0, \ldots, G$ $(G \ll L)$, according to

$$\mathbf{h} := \Big[h_0 \overbrace{\underbrace{0 \cdots 0}_{f_0 \text{ zeros}}\ h_1\ \underbrace{0 \cdots 0}_{f_1 \text{ zeros}}\ h_2\ \cdots\ h_{G-1}\ \underbrace{0 \cdots 0}_{f_{G-1} \text{ zeros}}\ h_G}^{\text{Channel memory length } L} \Big]^T, \tag{1}$$

where the numbers f_i are nonnegative integers and $L = \sum_{i=0}^{G-1}(f_i+1)$. A sparse ISI channel, for which $f_0 = f_1 = \cdots = f_{G-1} =: f$ holds, is called a *zero-pad* channel [8, 9]. (In a more relaxed definition, one would allow for coefficients that are not exactly zero, but still negligible.)

Throughout this paper, the complex baseband notation is used. The kth transmitted data symbol is denoted as $x[k]$, where k is the time index. A hypothesis for $x[k]$ is denoted as $\tilde{x}[k]$ and the corresponding hard decision as $\hat{x}[k]$. In the case of fading, we will assume a block-fading channel model for simplicity (block length $N \gg L$). The equivalent discrete-time channel model (for a single block of data symbols) is given by

$$y[k] = h_0 x[k] + \sum_{g=1}^{G} h_g x[k - d_g] + n[k], \tag{2}$$

where $y[k]$ denotes the kth received sample and $n[k]$ the kth sample of a complex additive white Gaussian noise (AWGN) process with zero mean and variance σ_n^2. Moreover,

$$d_g := \sum_{i=1}^{g} (f_{i-1} + 1), \quad 1 \le g \le G \tag{3}$$

denotes the position of channel coefficient h_g within the channel vector $\mathbf{h}$ $(d_G := L)$.

In the following, the channel vector $\mathbf{h}$ is assumed to be known at the receiver. Moreover, an M-ary alphabet for the data symbols is assumed. The complexity in terms of trellis states of the conventional Viterbi/BCJR algorithm is given by $\mathcal{O}(M^L)$ and is therefore normally prohibitive. Given a zero-pad channel, the conventional trellis diagram with $M^L = M^{(f+1)G}$ states can be decomposed into $(f+1)$ parallel regular trellises (without loss of optimality), each having only M^G states (P-VA) [8, 9]. As will be shown in the sequel, such a decomposition is *not* possible for general sparse channels.

2.1. Application of the parallel-trellis Viterbi algorithm

In order to decompose a given trellis diagram into multiple parallel trellises, the following question is of central interest. Which symbol decisions $\hat{x}[k]$, $0 \le k \le N-1$, are influenced by a certain symbol hypothesis $\tilde{x}[k_0]$, where k_0 denotes a specific time index? Suppose, a certain decision $\hat{x}[k_1]$ is *not* influenced by the hypothesis $\tilde{x}[k_0]$. Furthermore, let the set $\hat{\mathbb{X}}_{k_0} := \{\hat{x}[k] \mid \hat{x}[k] \text{depends on } \tilde{x}[k_0]\}$ contain all decisions $\hat{x}[k]$, $0 \le k \le N-1$, influenced by $\tilde{x}[k_0]$ and the set $\hat{\mathbb{X}}_{k_1}$ all decisions influenced by $\tilde{x}[k_1]$. If these two sets are disjoint, that is, $\hat{\mathbb{X}}_{k_0} \cap \hat{\mathbb{X}}_{k_1} = \emptyset$, the hypotheses $\tilde{x}[k_0]$ and $\tilde{x}[k_1]$ can be accommodated in *separate* trellis diagrams without loss of optimality. In other words, a decomposition of the overall trellis diagram into (at least two) parallel regular trellises is possible.

This fact is illustrated in Figure 1 for two example CIRs ($L = 8$ and $G = 2$ in both cases):

$$\begin{aligned}\mathbf{h}^{(1)} &:= \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & 0 & h_1 & 0 & h_2 \end{bmatrix}^T, \\ \mathbf{h}^{(2)} &:= \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & 0 & 0 & h_1 & h_2 \end{bmatrix}^T.\end{aligned} \tag{4}$$

Consider a particular symbol hypothesis $\tilde{x}[k_0]$. For simplicity it is assumed that hard decisions $\hat{x}[k]$ are already available for all time indices $k < k_0$. Moreover, it is assumed that the hypothesis $\tilde{x}[k_0]$ does not influence any decision $\hat{x}[k]$ with $k > k_0 + DL$, where $D = 2$ is considered in the example. (This corresponds to the assumption that a VA with a decision delay of DL symbol durations is optimal in the sense of MLSE.) The diagrams in Figure 1 may be interpreted as factor graphs [19] and illustrate the dependencies between hypothesis $\tilde{x}[k_0]$ and all decisions $\hat{x}[k]$, $k_0 \le k \le k_0 + DL$.

To start with, consider first the CIR $\mathbf{h}^{(1)}$ (cf. Figure 1(a)). It can be seen from (2) that only the received samples $y[k_0]$, $y[k_0+6]$, and $y[k_0+8]$ are directly influenced by the data symbol $x[k_0]$. Therefore, there is a dependency between hypothesis $\tilde{x}[k_0]$ and the decisions $\hat{x}[k_0]$, $\hat{x}[k_0+6]$, and $\hat{x}[k_0+8]$. The received sample $y[k_0+8]$, for example, is also influenced by the data symbol $x[k_0+2]$. Correspondingly, there is also a dependency between $\tilde{x}[k_0]$ and the decision $\hat{x}[k_0+2]$. The data symbols $x[k_0+6]$ and $x[k_0+8]$ again influence the received samples $y[k_0+12]$, $y[k_0+14]$, and $y[k_0+16]$, and so on. Including all dependencies, one obtains the second graph of Figure 1(a).

As can be seen, there is a dependency between $\tilde{x}[k_0]$ and all decisions $\hat{x}[k_0+2\nu]$, where $\nu = 0, 1, \ldots, \lfloor DL/2 \rfloor$, that is, symbol decisions for even and odd time indices are independent. Consequently, in this example it is possible to decompose the conventional trellis diagram into two parallel regular trellises, one comprising all even time indices and the other one comprising all odd time indices. While the conventional trellis diagram has M^8 trellis states, there are only M^4 states in each of the two parallel trellises. (Moreover, a single trellis segment in the parallel trellises spans two consecutive time indices.) This result is in accordance with [8, 9], since the CIR $\mathbf{h}^{(1)}$ in fact constitutes a zero-pad channel with CIR $\begin{bmatrix} h'_0 & 0 & h'_1 & 0 & h'_2 & 0 & h'_3 & 0 & h'_4 \end{bmatrix}^T$, where $G' = 4$, $f' = 1$, and $h'_1 = h'_2 = 0$. Generally spoken, a decomposition of a given trellis diagram into multiple parallel regular trellises is possible, if all nonzero channel coefficients of the sparse ISI channel are on a zero-pad grid with $f \ge 1$. Only in this case can the optimal P-VA be applied; otherwise one has to resort to the sub-P-VA or to alternative solutions such as the M-VA.

The computational complexities of the conventional VA and the P-VA, in terms of the overall number of branch metrics computed for a single decision $\hat{x}[k_0]$, are stated in Table 1. If there are only $(G+1)$ non-zero channel coefficients, the conventional VA can be modified such that it avoids to compute the same branch metric several times [7], which leads to a computational complexity of only $\mathcal{O}(M^{G+1})$. However, the number of trellis states is not reduced. As opposed to this, the P-VA offers both a reduced computational complexity and a reduced number of trellis states.

The second CIR $\mathbf{h}^{(2)}$ constitutes an example, where a decomposition of the conventional trellis diagram into multiple parallel regular trellises is *not* possible (at least not without loss of optimality). As can be seen in Figure 1(b), symbol hypothesis $\tilde{x}[k_0]$ influences *all* other symbol decisions $\hat{x}[k]$, $k_0 \le k \le k_0+DL$. Still, a decomposition into multiple parallel *irregular* trellises is possible, as proposed in [7] for the M-VA. By this means, sparse ISI channels with a general structure can be tackled.

2.2. Suboptimality of the multitrellis Viterbi algorithm

The basic idea of the M-VA is to construct an irregular trellis diagram for each individual symbol decision $\hat{x}[k_0]$, $0 \le k_0 \le N-1$. The trellis diagram for time index k_0 is based on all time indices $k = k_0 + n_1 d_1 + n_2 d_2 + \cdots + n_G d_G$, where $n_1, \ldots, n_G$ are nonnegative integers and the values of $d_1, \ldots, d_G$ are given by the sparse CIR under consideration (cf. (2) and (3)). (Similarly to Figure 1(a), it is assumed that symbol decisions are already available for all time indices $k < k_0$.) In order to obtain a trellis diagram of finite length, only those integer values n_g are taken into account for which $k \le DL$ results, that is, a certain predefined decision delay DL is required ($D > 0$ integer). The symbol decision for time index k_0 finally results from searching the maximum-likelihood path within the corresponding irregular trellis diagram (using the VA).

As an example, the irregular trellis structure resulting for the CIR $\mathbf{h}^{(1)}$ is depicted in Figure 2 (for $D = 2$ and binary transmission). The replicas $\tilde{y}[k] = h_0\tilde{x}[k] + \sum_g h_g\tilde{x}[k-d_g]$

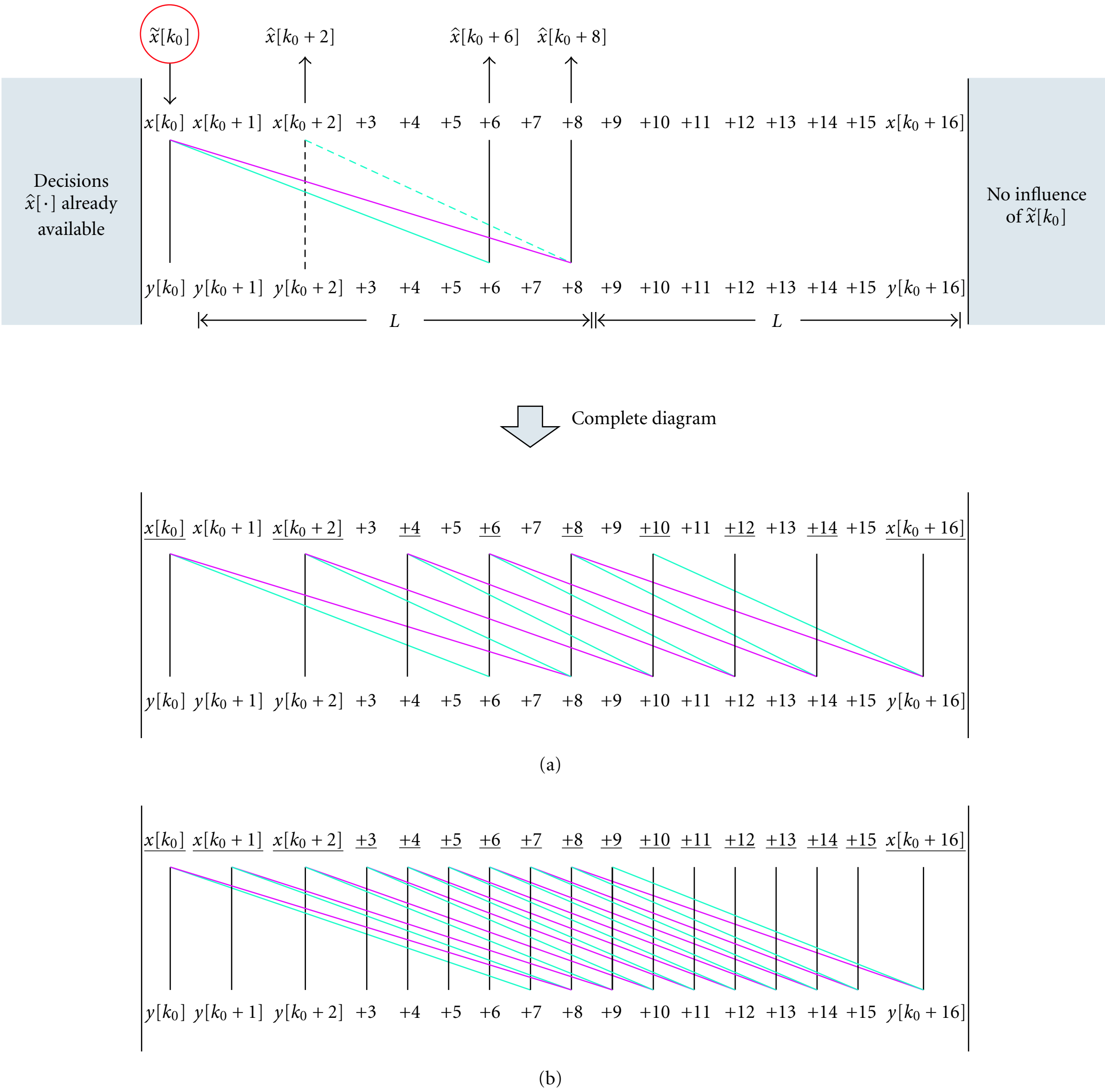

FIGURE 1: Dependencies between symbol hypothesis $\tilde{x}[k_0]$ and subsequent decisions $\hat{x}[k]$ for two different example channels. (a) CIR $\mathbf{h}^{(1)} = [h_0\ 0\ 0\ 0\ 0\ 0\ h_1\ 0\ h_2]^T$ and (b) CIR $\mathbf{h}^{(2)} = [h_0\ 0\ 0\ 0\ 0\ 0\ 0\ h_1\ h_2]^T$.

TABLE 1: Computational complexity in terms of the overall number of branch metrics computed for each symbol decision: conventional Viterbi algorithm (VA) and parallel-trellis VA (P-VA). In the case of the P-VA, it was assumed that all channel coefficients on the zero-pad grid are unequal to zero.

Conventional VA, any CIR with memory length L [and $(G+1)$ nonzero coefficients]	P-VA, zero-pad CIR with $(G+1)$ nonzero coefficients
$\mathcal{O}(M^{L+1})$ $[\mathcal{O}(M^{G+1})]$	$\mathcal{O}((f+1)\cdot M^{G+1})$

(and the associated symbol hypotheses $\tilde{x}[\cdot]$) required for the calculation of the branch metrics $|y[k]-\tilde{y}[k]|^2$ are also included (see [7] for further details). It should be noted that for some trellis branches multiple branch metrics have to be calculated. For example, for the replica $\tilde{y}[k_0+8]$, the hypotheses $\tilde{x}[k_0+8]$, $\tilde{x}[k_0+2]$, and $\tilde{x}[k_0]$ are required. Since hypothesis $\tilde{x}[k_0+2]$ is not accommodated in the corresponding trellis states, all M possibilities have to be checked in order to find the best branch metric.

The computational complexity of the M-VA depends on the channel memory length of the given CIR, the number of nonzero channel coefficients, the parameters $d_1,\ldots,d_G$, and on the choice of the parameter D. It is therefore difficult to find general rules. In Table 2, the computational complexity of the M-VA is stated for the example CIR $\mathbf{h}^{(1)}$ and different decision delays DL ($D = 1,2,3$). The corresponding

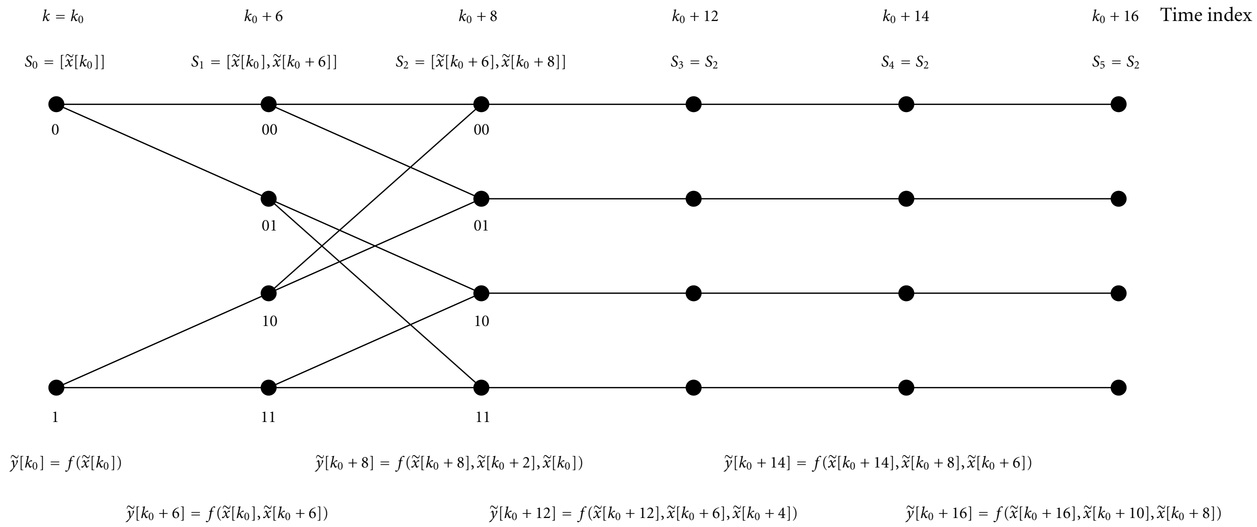

Figure 2: Irregular trellis structure of the M-VA resulting for a single symbol decision $\hat{x}[k_0]$ ($D = 2$, binary transmission, example CIR $\mathbf{h}^{(1)} = [h_0\ 0\ 0\ 0\ 0\ 0\ h_1\ 0\ h_2]^T$).

Table 2: Computational complexity in terms of the overall number of branch metrics computed for each symbol decision: multitrellis VA (M-VA) with different decision delays DL (example CIR $\mathbf{h}^{(1)} = [h_0\ 0\ 0\ 0\ 0\ 0\ h_1\ 0\ h_2]^T$).

M-VA ($D = 1$)	M-VA ($D = 2$)	M-VA ($D = 3$)
$\mathcal{O}(M^4)$	$\mathcal{O}(2M^4+M^3+M^2+M)$	$\mathcal{O}(4M^5+3M^4+M^2+M)$

complexity of the conventional VA and the P-VA is given by $\mathcal{O}(M^9)$ and $\mathcal{O}(2M^5)$, respectively.

Taking a closer look at the trellis diagram in Figure 2, it can be seen that a significant part of the dependencies shown in Figure 1(a) is neglected by the M-VA. This is illustrated in Figure 3. As a result, the M-VA is clearly suboptimal, although it was claimed to be optimal in the sense of MLSE [7]. Moreover, as will be shown in Section 4, for a good performance, the required decision delay DL (and thus the computational complexity) tends to be quite large.[2]

2.3. Drawbacks of the suboptimal parallel-trellis Viterbi algorithm

With regard to sparse channels having a general structure, the sub-P-VA constitutes an alternative to the M-VA. The main principle of the sub-P-VA is as follows. Given a general sparse ISI channel, one first tries to find an underlying zero-pad channel with a structure as close as possible to the CIR under consideration. Based on this, the multiple parallel (regular) trellises are defined. Finally, in order to cancel residual ISI, tentative (soft) decisions are exchanged between the parallel trellises [8–10].

For a good performance, however, the given CIR should at least be close to a zero-pad structure, that is, there should only be some small nonzero coefficients in between the main coefficients. Given a fading channel, the sub-P-VA seems to be of limited practical relevance: the algorithm has to be redesigned for each new channel realization, because the position of the main channel coefficients might change. Moreover, the amount of required decision feedback between the parallel trellises can be quite large, because in a practical system there are normally no channel coefficients that are exactly zero.

2.4. A simple alternative

The above discussion has shown that trellis-based equalization of general sparse ISI channels is quite a demanding task: the optimal P-VA (or the P-BCJRA) can only be applied for zero-pad channels. For general sparse channels, there is no optimal reduced-complexity trellis-based equalization technique available in the literature. Indeed, the suboptimal M-VA or the sub-P-VA can be applied for general sparse channels. However, the complexity of the M-VA tends to be quite large, and for a good performance of the sub-P-VA the CIR should be close to a zero-pad structure.

In this context the question arises, whether it is really useful to explicitly utilize the sparse channel structure for trellis-based equalization, especially in the case of a fading

[2] If all dependencies shown in Figure 1(a) were taken into account in order to construct the irregular trellis diagrams, the complexity of the M-VA would actually exceed that of the conventional VA. Even then the M-VA would—strictly speaking—not be optimal in the sense of MLSE, due to the finite decision delay DL. (In the case of the P-VA the finite decision delay is, in fact, not required. It has only been introduced here for illustrative purposes.)

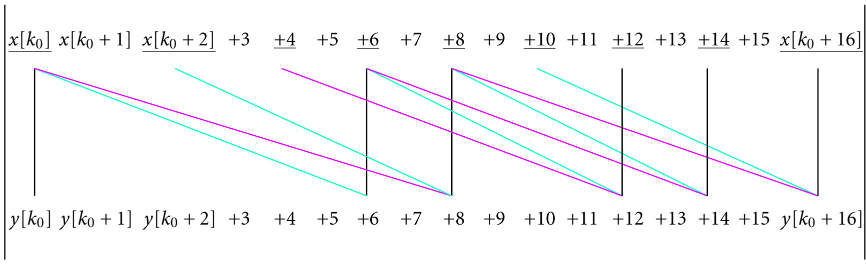

FIGURE 3: Dependencies between the individual symbol hypotheses $\tilde{x}[k]$ that are taken into account by the M-VA ($D = 2$, example CIR $\mathbf{h}^{(1)} = [h_0\ 0\ 0\ 0\ 0\ 0\ h_1\ 0\ h_2]^T$).

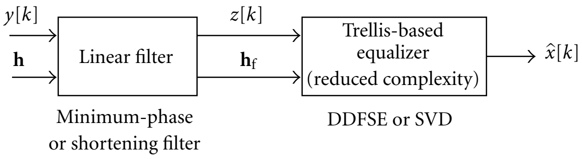

FIGURE 4: Receiver structure under consideration.

channel.[3] How efficient are standard trellis-based equalization techniques (designed for conventional, non-sparse ISI channels) in conjunction with prefiltering, when applied to (general) sparse ISI channels? This question is addressed in the following section.

3. PREFILTERING FOR SPARSE CHANNELS

The receiver structure considered in the sequel is illustrated in Figure 4, where $z[k]$ denotes the kth received sample after prefiltering and $\mathbf{h}_f$ the filtered CIR.

Two types of linear filters are considered here, namely, a minimum-phase filter [13–15] and a channel shortening filter [26]. In the case of the minimum-phase filter, a DDFSE equalizer [25] is employed. (As will be discussed in Section 3.5, the sparse channel structure is normally lost after prefiltering, which suggests the use of a standard trellis-based equalizer designed for non-sparse channels.) As an alternative receiver structure, the channel shortening filter is used in conjunction with a conventional Viterbi equalizer. The Viterbi equalizer operates on a shortened CIR with memory length $L_s \ll L$, which is in the following indicated by the term shortened Viterbi detector (SVD). The SVD equalizer is no longer optimal in the sense of MLSE. The considered receiver structures are notably simple, because solely the linear filters are adjusted to the current CIR, which is particularly favorable with regard to fading channels. The filter coefficients can be computed efficiently using standard techniques available in the literature. Moreover, the receiver structures offer a flexible complexity-performance trade-off.

To start with, the two prefiltering approaches and the equalizer concepts are briefly recapitulated. Then, the overall complexities of the receiver structures under consideration are discussed as well as the channel structure after prefiltering. Numerical results for various examples will be presented in Section 4, so as to demonstrate the efficiency of the considered receiver structures.

3.1. Minimum-phase filter

Consider a static ISI channel with CIR $\mathbf{h} := [h_0, h_1, \ldots, h_L]^T$ and let $H(z)$ denote the z-transform of $\mathbf{h}$. Furthermore, let $\mathbf{h}_{\min} := [h_{\min,0}, h_{\min,1}, \ldots, h_{\min,L}]^T$ denote the equivalent minimum-phase CIR of $\mathbf{h}$ and $H_{\min}(z)$ the corresponding z-transform. In the z-domain, all zeros of $H_{\min}(z)$ are either inside or on the unit circle [27, Chapter 3.4]. In the time domain, $\mathbf{h}_{\min}$ is characterized by an energy concentration in the first channel coefficients [13, 14] (especially if the zeros of $H(z)$ are not too close to the unit circle). The z-transform $H_{\min}(z)$ is obtained by reflecting those zeros of $H(z)$, that are outside the unit circle, into the unit circle, whereas all other zeros are retained for $H_{\min}(z)$. The ideal linear filter, which transforms $\mathbf{h}$ into its minimum-phase equivalent, has allpass characteristic [14], that is, it does not color the noise. A good overview of possible practical realizations can be found in [14]. In this paper, we use an approach that is based on an implicit spectral factorization based on the Kalman filter [13, 15], so as to approximate the ideal linear minimum-phase filter by a finite-impulse-response (FIR) filter of length $L_F < \infty$. (It should be noted that some performance degradation has to be expected, when using a practical filter with a finite length [10].) The resulting filter approximates a discrete-time whitened matched filter (WMF). The computational complexity of calculating the filter coefficients is $\mathcal{O}(L_F L^2)$, that is, it is only linear with respect to the filter length. By this means, comparably large filter lengths are feasible.

3.2. Channel shortening filter

In this approach, a linear filter is used to transform a given CIR $\mathbf{h} := [h_0, h_1, \ldots, h_L]^T$ into a shortened CIR $\mathbf{h}_s := [h_{s,0}, h_{s,1}, \ldots, h_{s,L_s}]^T$, where $L_s < L$ denotes the desired

[3] In contrast to this, utilizing the sparse channel structure for linear or decision-feedback equalization indeed leads to efficient reduced-complexity techniques [2–6]. Also, linear or decision-feedback schemes might be more suitable for adaptive equalization of sparse channels than trellis-based techniques.

channel memory length. Several methods to design a linear channel shortening filter (CSF) can be found in the literature, see, for example, [28] for an overview. In this paper, a method described in [26] is used, which is based on the feed-forward filter (FFF) of a minimum mean-squared error (MMSE) DFE. The filter design is as follows: for the feedback filter (FBF) of the MMSE-DFE, a fixed filter length of $(L_s + 1)$ is chosen. Under this constraint, the FFF of the DFE is then optimized with respect to the MMSE criterion, where the length L_F of the FFF can be chosen irrespective of L_s. The optimized FFF finally constitutes a linear finite-length CSF: the mean-squared error between the shortened CIR $\mathbf{h}_s$ after the FFF and the coefficients of the FBF is minimized, that is, all channel coefficients $h_{s,l}$ with $l < 0$ or $l > L_s$ are optimally suppressed in the MMSE sense. Correspondingly, a subsequent SVD equalizer will only take the desired channel coefficients $h_{s,l}$, $0 \le l \le L_s$, into account. As opposed to the minimum-phase filter, an arbitrary power distribution results among the desired coefficients. Moreover, the CSF does not approximate an all-pass filter, that is, depending on the given CIR $\mathbf{h}$ the CSF can lead to colored noise. The computational complexity of calculating the filter coefficients is $\mathcal{O}(L_F^3)$ [26].

3.3. Equalizer concepts

The main difference between the conventional Viterbi equalizer used for MLSE detection and suboptimal reduced-state equalizers, such as the SVD equalizer or the DDFSE equalizer, concerns the number of trellis states and the calculation of the branch metrics. (The accumulated branch metrics constitute the basis on which the Viterbi equalizer—or a reduced-state version thereof—selects the most probable data sequence.) In the case of the Viterbi equalizer (and white Gaussian noise), the optimal branch metrics $\mu_k(y[k], \tilde{y}[k])$ at time instant k are given by the squared Euclidean distance between the kth received sample $y[k]$ and all possible hypotheses (replicas) $\tilde{y}[k]$:

$$\begin{aligned}\mu_k(y[k], \tilde{y}[k]) &:= |y[k] - \tilde{y}[k]|^2 \\ &= \left| y[k] - h_0\tilde{x}[k] - \sum_{l=1}^{L} h_l \tilde{x}[k-l] \right|^2.\end{aligned} \tag{5}$$

The number of trellis states is given by the number of possible hypotheses $\tilde{x}[k-l]$ $(l = 1, \ldots, L)$, which is M^L. As opposed to this, the SVD equalizer operates on a shortened channel memory length $L_s < L$, that is, the number of trellis states is M^{L_s}. (The branch metric computation is the same as in (5), where L is replaced by L_s.)

The DDFSE equalizer is obtained from the conventional Viterbi equalizer by applying the principle of parallel decision feedback [25]: the number of trellis states is reduced to M^K, $K < L$, by replacing the hypotheses $\tilde{x}[k-l]$, $l = K+1, \ldots, L$, by tentative decisions:

$$\begin{aligned}\mu_k(y[k], \tilde{y}[k]) = \bigg| & y[k] - h_0\tilde{x}[k] \\ & - \sum_{l=1}^{K} h_l \tilde{x}[k-l] - \sum_{l=K+1}^{L} h_l \hat{x}[k-l] \bigg|^2.\end{aligned} \tag{6}$$

Note that in the special case $K = L$, the DDFSE equalizer is equivalent to the Viterbi equalizer, whereas in the special case $K = 1$ it is equivalent to a DFE. It should be noted that due to the parallel decision feedback, the complexity of the DDFSE equalizer is slightly larger than that of the SVD equalizer, given the same value for K and L_s.

3.4. Computational complexity of the considered receiver structures

In the sequel, three different receiver structures are considered (cf. Figure 4):

(i) a full-state Viterbi equalizer (MLSE, memory length L, no prefiltering),
(ii) a DDFSE equalizer with memory length $K < L$ and minimum-phase filter (WMF),
(iii) an SVD equalizer with memory length $L_s < L$ and channel shortening filter (CSF).

(In the case of MLSE, minimum-phase prefiltering has no impact on the bit-error-rate performance [15].)

The computational complexity of these three receiver structures is summarized in Table 3. In order to obtain a complexity similar to that of the sub-P-VA/sub-P-BCJRA equalizer, the parameters K, L_s should be chosen such that[4]

$$K, L_s \le \log_M(f+1) + G, \tag{7}$$

where the parameters f and G are associated with the underlying zero-pad channel selected for the sub-P-VA/sub-P-BCJRA.

3.5. Channel structure after prefiltering

The sparse structure of a given CIR $\mathbf{h}$ is normally lost after prefiltering. This is obvious in the case of the shortening filter, since an arbitrary power distribution results among the desired (L_s+1) channel coefficients. However, the sparse structure is—in general—also lost when applying the minimum-phase filter.

An exception is the zero-pad channel, where the sparse CIR structure is *always* preserved after minimum-phase prefiltering. Let $\mathbf{h} := [h_0 \ h_1 \ \cdots \ h_G]^T$ denote a (non-sparse) CIR with z-transform $Z\{\mathbf{h}\} = H(z)$ and equivalent minimum-phase z-transform $H_{\min}(z)$, and let $\mathbf{h}_{\mathrm{ZP}}$ denote the corresponding zero-pad CIR with memory length $(f+1)G$ and z-transform $H_{\mathrm{ZP}}(z)$, which results from inserting f zeros in between the coefficients of $\mathbf{h}$. Furthermore, let $z_{0,1}, \ldots, z_{0,G}$

[4] Equation (7) constitutes only a rule-of-thumb: on the one hand, it does not take the prefilter computation into account that is required for the considered receiver structures. On the other hand, it also neglects the exchange of tentative decisions required for the sub-P-VA/sub-P-BCJRA equalizer. In order to obtain a similar complexity in both cases, the parameter K of the DDFSE equalizer (or L_s for the SVD equalizer) should be chosen such that the number of branch metrics computed per symbol decision is not larger than for the sub-P-VA/sub-P-BCJRA equalizer, that is, M^{K+1} should be smaller or equal to $(f+1)M^{G+1}$ (cf. Tables 1 and 3).

Table 3: Computational complexity of the considered receiver structures. Delayed decision-feedback sequence estimation (DDFSE) with whitened matched filter (WMF), and shortened Viterbi detection (SVD) with channel shortening filter (CSF). For the equalizer algorithms, the overall number of branch metrics computed for each symbol decision is stated and for the linear filters the approximate computational complexity of calculating the filter coefficients.

	DDFSE + WMF, (memory length $K < L$)	SVD + CSF, (memory length $L_s < L$)	Conventional VA, (memory length L)
Equalizer	$\mathcal{O}(M^{K+1})$	$\mathcal{O}(M^{L_s+1})$	$\mathcal{O}(M^{L+1})$
Prefilter	$\mathcal{O}(L_F L^2)$	$\mathcal{O}(L_F^3)$	—

denote the zeros of $H(z)$. An insertion of f zeros in the time domain corresponds to a transform $z \mapsto z^{f+1}$ in the z-domain, that is, $H_{\text{ZP}}(z) = H(z^{f+1})$. This means, the $(f+1)G$ zeros of $H_{\text{ZP}}(z)$ are given by the $(f+1)$ complex roots of $z_{0,1}, \ldots, z_{0,G}$, respectively. Consider a certain zero $z_{0,g} := r_{0,g} \cdot \exp(j\varphi_{0,g})$ of $H(z)$ that is outside the unit circle ($r_{0,g} > 1$). This zero will lead to $(f+1)$ zeros

$$z_{0,g}^{(\lambda)} := r_{0,g}^{1/(f+1)} \cdot \exp\left(j\frac{2\pi\lambda + \varphi_{0,g}}{f+1}\right) \tag{8}$$

of $H_{\text{ZP}}(z)$ ($\lambda = 0, \ldots, f$) that are located on a circle of radius $r_{0,g}^{1/(f+1)} > 1$ that is also outside the unit circle. By means of (ideal) minimum-phase prefiltering, these zeros are reflected into the unit circle, that is, the corresponding zeros of $H_{\text{ZP, min}}(z)$ are given by $1/z_{0,g}^{(\lambda)*}$, where $(\cdot)^*$ denotes complex conjugation.

Correspondingly, the sparse CIR structure is retained after minimum-phase prefiltering (with the same zero-pad grid). The zeros of $H_{\text{ZP, min}}(z)$ are the $(f+1)$ roots of the zeros of $H_{\min}(z)$, and the nonzero coefficients of $\mathbf{h}_{\text{ZP, min}}$ are given by the minimum-phase CIR $\mathbf{h}_{\min} = Z^{-1}\{H_{\min}(z)\}$. If the zeros of $H(z)$ are not too close to the unit circle, $\mathbf{h}_{\min}$ is characterized by a significant energy concentration in the first channel coefficients. In this case, the effective channel memory length of $\mathbf{h}_{\text{ZP}}$ is significantly reduced by minimum-phase prefiltering, namely, by some multiples of $(f+1)$ (cf. (1)).

4. NUMERICAL RESULTS

In the sequel, the efficiency of the receiver structures considered in Section 3 is illustrated by means of numerical results obtained by Monte-Carlo simulations over 10 000 data blocks. In all cases, the channel coefficients were perfectly known at the receiver. Channel coding was not taken into account.

4.1. Static channel impulse response

To start with, a *static* sparse ISI channel is considered, and the bit-error-rate (BER) performance of the receiver structures considered in Section 3 is compared with that of the M-VA equalizer [7]. As an example, the CIR $\mathbf{h}^{(1)}$ from Section 2 is

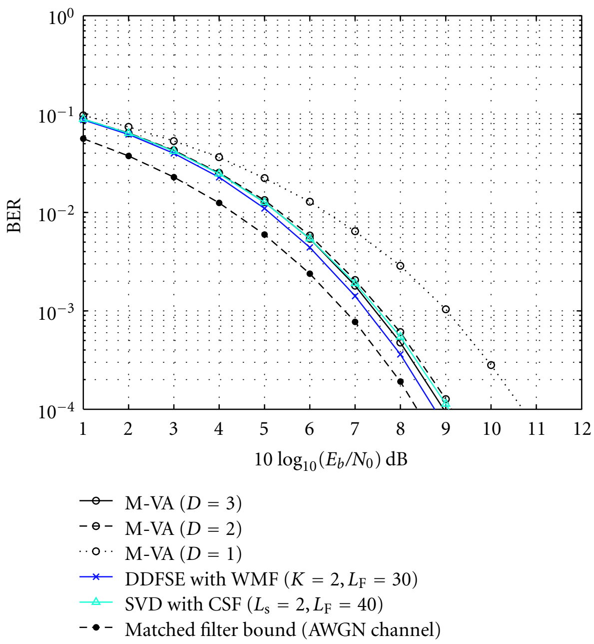

Figure 5: BER performance of the considered receiver structures compared to the M-VA equalizer [7] (static sparse ISI channel).

considered with $h_0 = 0.2076$, $h_1 = 0.87$, and $h_2 = 0.4472$ ($\|\mathbf{h}^{(1)}\| = 1$), that is, $\mathbf{h}^{(1)}$ is nonminimum phase.

The BER performance for binary antipodal transmission ($x[k] \in \{\pm 1\}$, $M = 2$) of the M-VA equalizer, the DDFSE equalizer with WMF, and the SVD equalizer with CSF is displayed in Figure 5, as a function of E_b/N_0 in dB, where E_b denotes the average energy per bit and N_0 the single-sided noise power density ($E_b/N_0 := 1/\sigma_n^2$). Due to the given channel memory length, the complexity of MLSE detection is prohibitive. As a reference curve, however, the matched filter bound (MFB) is included, which constitutes a lower bound on the BER of MLSE detection [29]. The filter lengths for the WMF and the CSF were chosen sufficiently large (in this case $L_F = 30$ for the WMF and $L_F = 40$ for the CSF), that is, a further increase of the filter lengths gives only marginal performance improvements. (According to a rule of thumb, the filter length for the WMF should be chosen as $L_F \geq 2.5(L+1)$ [15].) Since the channel is static, the filters have to be computed only once. The memory length of the DDFSE equalizer/the SVD equalizer was chosen as $K, L_s = 2$, that is, there were only four trellis states. For the M-VA equalizer, different decision delays DL were considered ($D = 1, 2, 3$).

As can be seen, the performance of the DDFSE equalizer with WMF and the SVD equalizer with CSF is quite close to the MFB. (At a BER of 10^{-3}, the gap is less than 1 dB.) When a decision delay of $2L$ or $3L$ is chosen for the M-VA equalizer, a similar performance is achieved. Note, however, that the complexity is well above that of the DDFSE equalizer with WMF/the SVD equalizer with CSF (cf. Table 2). When the decision delay is reduced to L, a significant performance loss has to be accepted for the M-VA, and still the complexity is larger than for the DDFSE equalizer with WMF/the SVD equalizer with CSF. (However, no prefilter coefficients have to be computed.)

In Figure 6, the BER performance of the considered receiver structures is compared with the sub-P-BCJRA equalizer [10]. As an example, the CIR

$$\mathbf{h} = \begin{bmatrix} h_0 & 0 & 0 & 0 & h_1 & 0 & 0 & h_2 & 0 \cdots 0 & h_3 \end{bmatrix}^T \qquad (L = 15) \tag{9}$$

with $h_0 = 0.87$ and $h_1 = h_2 = h_3 = 0.29$ from [10] was taken ($\|\mathbf{h}\| = 1$), which is nonminimum phase and has a general sparse structure (i.e., not a zero-pad structure). When the parameters K and L_s for the DDFSE and the SVD equalizer, respectively, are chosen as $K, L_s = 4$, the overall receiver complexity is approximately the same as for the sub-P-BCJRA equalizer. In this case, the DDFSE equalizer in conjunction with the WMF achieves a similar BER performance as the sub-P-BCJRA equalizer. At a BER of 10^{-3}, the loss with respect to the MFB is only about 1 dB.[5] At the expense of a small loss (0.5 dB at the same BER), the complexity of the DDFSE equalizer can be further reduced to $K = 3$. The BER performance of the SVD equalizer in conjunction with the CSF is worse than that of the DDFSE equalizer with WMF: at a BER of 10^{-3}, the gap to the MFB is about 2.1 dB for $L_s = 4$ and 4.2 dB for $L_s = 3$. (Obviously, the considered CIR is more difficult to equalize than the one in Figure 5, since both the channel memory length and the number of nonzero channel coefficients is larger.)

4.2. *Fading channel impulse response*

Next, we consider the case of a sparse Rayleigh fading channel model, that is, the channel coefficients h_g ($g = 0, \ldots, G$) in (1) are now zero-mean complex Gaussian random variables

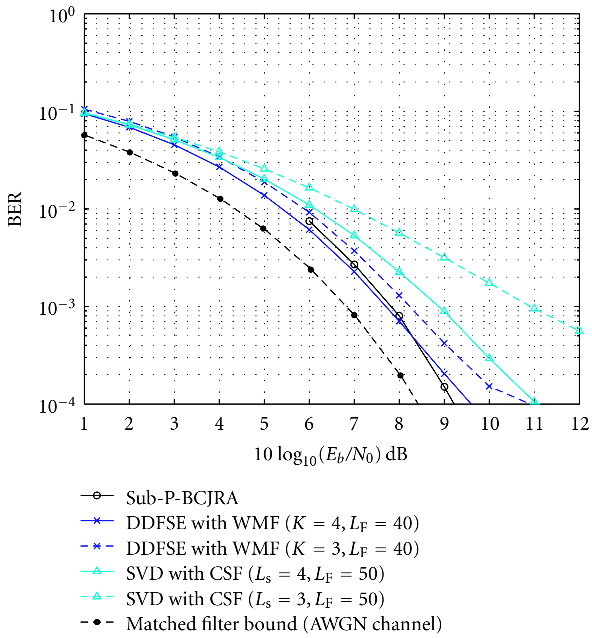

FIGURE 6: BER performance of the considered receiver structures compared to the sub-P-BCJRA equalizer [10] (static sparse ISI channel).

with variance $E\{|h_g|^2\} =: \sigma_{h,g}^2$. It is assumed in the following that the individual channel coefficients are statistically independent. Moreover, block fading is considered for simplicity (block length $N \gg L$). As an example, we consider a CIR with $G = 3$ and a power profile

$$\mathbf{p} := \begin{bmatrix} \sigma_{h,0}^2 & \underbrace{0 \cdots 0}_{f \text{ zeros}} & \sigma_{h,1}^2 & 0 & 0 & 0 & \sigma_{h,2}^2 & \sigma_{h,3}^2 \end{bmatrix}^T. \tag{10}$$

Note that this CIR again does not have a zero-pad structure. By choosing different values for the parameter f, different channel memory lengths $L = f + 6$ can be studied. To start with, consider a power profile with equal variances $\sigma_{h,0}^2 = \cdots = \sigma_{h,3}^2 = 0.25$ and a memory length of $L = 12$. Figure 7 shows the power profiles that result after prefiltering with the WMF and the CSF, respectively, for large values of E_b/N_0. The filter length was $L_F = 36$ in both cases. As can be seen, after prefiltering with the WMF the sparse structure of the power profile is lost (cf. Section 3.5). Significant variances $E\{|h_{\min,l}|^2\}$ occur, for example, at $l = 1$, $l = 4$, and $l = 5$. The power profile after the WMF exhibits a considerable energy concentration in the first channel coefficient, whereas the variances $E\{|h_{\min}, l|^2\}$ for $l = 7$, $l = 11$, and $l = 12$ are smaller than for the original CIR. As will be seen, this significantly improves the performance of the subsequent DDFSE equalizer. For the CSF, a desired channel memory length of $L_s = 5$ was chosen. After prefiltering with the CSF, the variances $E\{|h_{s,l}|^2\}$ for $l < 0$ and $l > L_s$ are

[5] It should be noted that for large values of E_b/N_0 the performance of the DDFSE equalizer with WMF is (slightly) inferior to that of the sub-P-BCJRA, which is mainly due to residual ISI: the convolution of the original CIR with the WMF generates non-zero channel coefficients h_l with $l > L$, which we did not take into account so as to limit the overall complexity of the DDFSE equalizer. However, since most practical systems employ channel coding, uncoded BERs of $10^{-2} \cdots 10^{-3}$ are of primary interest, that is, E_b/N_0 is typically smaller than 8 dB in coded systems (cf. Figure 6).

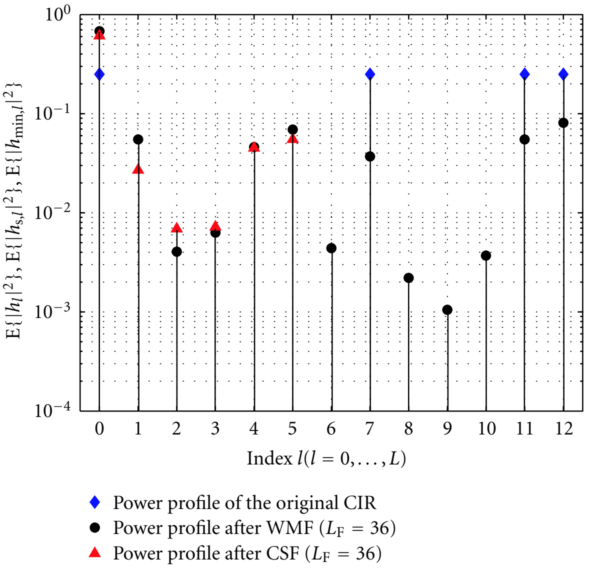

Figure 7: Power profiles after prefiltering with the WMF/CSF, resulting for large values of E_b/N_0. Sparse Rayleigh fading channel with $L = 12$ ($G = 3$) and equal variances $\sigma_{h,g}^2$ of the nonzero channel coefficients.

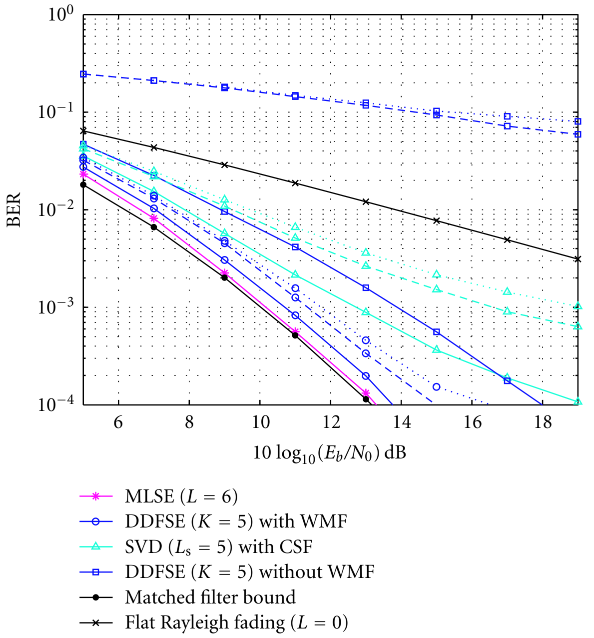

Figure 8: BER performance of the considered receiver structures: sparse Rayleigh fading channel with equal variances $\sigma_{h,g}^2$ of the nonzero channel coefficients; three different channel memory lengths L are considered (solid lines: $L = 6$, dashed lines: $L = 12$, dotted lines: $L = 20$).

virtually zero.[6] Correspondingly, a subsequent SVD equalizer with memory length $L_s = 5$ will not excessively suffer from residual ISI.

Figure 8 shows the BER performance of the considered receiver structures (binary transmission), again for equal variances $\sigma_{h,0}^2 = \cdots = \sigma_{h,3}^2 = 0.25$ and three different channel memory lengths L (solid lines: $L = 6$, dashed lines: $L = 12$, dotted lines: $L = 20$). The filter lengths have been chosen as $L_F = 20$ ($L = 6$), $L_F = 36$ ($L = 12$), and $L_F = 60$ ($L = 20$), both for the WMF and the CSF. As reference curves, the BER for flat Rayleigh fading ($L = 0$) is included as well as the MFB. For binary antipodal transmission, the MFB can generally be calculated as [29, Chapter 14.5]

$$\bar{P}_b = \frac{1}{2}\sum_{g=0}^{G}\left(\prod_{\substack{g'=0\\ \gamma_{g'}\neq\gamma_g}}^{G}\frac{\gamma_g}{\gamma_g-\gamma_{g'}}\right)\left(1-\sqrt{\frac{\gamma_g}{1+\gamma_g}}\right), \tag{11}$$

where $\gamma_g := \sigma_{h,g}^2/\sigma_n^2$ ($g = 0, \ldots, G$) and $\sigma_{h,0}^2 + \cdots + \sigma_{h,G}^2 := 1$. (Note that the MFB does not depend on the channel memory length L as long as the variances $\sigma_{h,g}^2$ remain unchanged.)

In the case $L = 6$, MLSE detection is still feasible. As can be seen in Figure 8, its performance is very close to the MFB. The DDFSE equalizer with $K = 5$ in conjunction with the WMF achieves a BER performance that is close to MLSE detection (the loss at a BER of 10^{-3} is only about 0.6 dB). Even when the channel memory length is increased to $L = 20$, the BER curve of the DDFSE equalizer with WMF deviates only 2 dB from the MFB (at the same BER). However, when the DDFSE equalizer is used without WMF, a significant performance loss occurs already for $L = 6$. Considering the case $L = 12$, it can be seen that the influence of the WMF (cf. Figure 7) makes a huge difference: the BER increases by several decades when the WMF is not used. Similar to the case of the static sparse ISI channels, the performance of the SVD equalizer ($L_s = 5$) with CSF is worse than that of the DDFSE equalizer with WMF, especially for large channel memory lengths L. Still, a significant gain compared to flat Rayleigh fading is achieved, that is, a good portion of the inherent diversity (due to independently fading channel coefficients) is captured.

Finally, in Figure 9 the case of unequal variances $\sigma_{h,g}^2$ is considered ($L = 12$; solid lines: energy concentration in the last channel coefficient; dashed lines: energy concentration in the first channel coefficient). In both cases, the performance of the DDFSE equalizer with WMF is quite close to the respective MFB (the difference is about 1.3–1.7 dB at a BER of 10^{-3}). As can be seen, the benefit of the WMF is smaller (but still significant) when the power profile of the original CIR already has an energy concentration in the first channel coefficient.

[6] As discussed in Section 3.2, the CSF is designed such that a given CIR is optimally shortened in the sense of the MMSE criterion. Since large values of E_b/N_0 are considered here, the MMSE solution and the zero-forcing (ZF) solution become equivalent, that is, the channel coefficients with $l < 0$ and $l > L_s$ are virtually nulled.

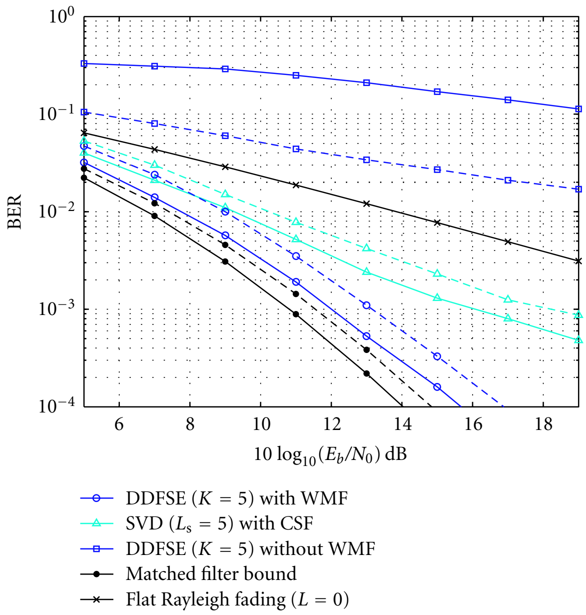

FIGURE 9: BER performance of the considered receiver structures: sparse Rayleigh fading channel with unequal variances $\sigma^2_{h,g}$ of the nonzero channel coefficients ($L = 12$; solid lines: $\sigma^2_{h,g} = \{0.1, 0.1, 0.3, 0.5\}$; dashed lines: $\sigma^2_{h,g} = \{0.7, 0.1, 0.1, 0.1\}$).

4.3. *Final remarks*

It should be noted that minimum-phase prefiltering of sparse ISI channels is also beneficial when using a tree-based equalization algorithm, such as the LISS algorithm [16, 17]. In order to obtain a small overall complexity, the metrics of two competing paths that deviate closely to the root of the tree should differ as much as possible. This is achieved by means of minimum-phase prefiltering, due to the energy concentration in the first coefficients of the filtered CIR.

An alternative to minimum-phase prefiltering could be to design a linear filter which transforms a given general sparse CIR into one with a zero-pad structure. This would then enable the use of the (optimal) P-VA after the linear filter. However, it was shown in [9] that no complexity reduction can be achieved by this approach.

Finally, it should be noted that the receiver structures considered in Section 3 are well suited for turbo equalization, where a soft-input soft-output (SISO) equalizer and a SISO channel decoder exchange soft information in an iterative fashion [21, 22]. For example, the soft values provided by soft-output versions of the DDFSE equalizer (e.g., based on the BCJRA) are known to be of good quality [13].

5. CONCLUSIONS

In this paper, trellis-based equalization of sparse intersymbol-interference channels has been revisited. Due to the large channel memory length of sparse channels, efficient equalization with an acceptable complexity is a demanding task. Based on a unified framework for complexity reduction without loss of optimality, two known trellis-based equalization techniques for sparse channels were recapitulated. It was demonstrated, in which cases a decomposition of the conventional trellis diagram into multiple parallel regular trellises is possible. Moreover, it was shown that the second equalization technique, designed for general sparse channels, is clearly suboptimal (although claimed otherwise). In order to tackle general sparse channels, receiver structures consisting of a linear filter and a reduced-complexity equalizer were studied. The employed equalizer algorithms were standard (i.e., not specifically designed for sparse channels), which is particularly favorable with regard to fading channels: only the filter coefficients have to be adjusted to the current channel impulse response, and they can be computed efficiently using standard techniques available in the literature. By means of numerical results, it was demonstrated that the considered receiver structures are able to compete with techniques specifically designed for sparse channels: using a delayed decision-feedback equalizer in conjunction with a whitened matched filter, bit error rates were achieved that deviate only 1–2 dB from the matched filter bound (at a bit error rate of 10^{-3}).

ACKNOWLEDGMENTS

The authors would like to thank Dr. Wolfgang Gerstacker (Chair of Mobile Communications, University of Erlangen-Nuremberg, Germany), Ragnar Thobaben (Institute for Circuits and System Theory, University of Kiel, Germany), Professor Jörg Kliewer (Coding Research Group, University of Notre Dame, Indiana, USA), and Dr. Nigel McGinty (Signals Analysis Group, Department of Defence, Edinburgh, South Australia, Australia) for fruitful discussions and helpful suggestions, especially concerning the prefiltering part. Moreover, many thanks to the anonymous reviewers for their valuable comments.

REFERENCES

[1] T. S. Rappaport, *Wireless Communications—Principles and Practice*, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.

[2] F. K. H. Lee and P. J. McLane, "Design of nonuniformly spaced tapped-delay-line equalizers for sparse multipath channels," *IEEE Transactions on Communications*, vol. 52, no. 4, pp. 530–535, 2004.

[3] I. J. Fevrier, S. B. Gelfand, and M. P. Fitz, "Reduced complexity decision feedback equalization for multipath channels with large delay spreads," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 927–937, 1999.

[4] S. F. Cotter and B. D. Rao, "The adaptive matching pursuit algorithm for estimation and equalization of sparse time-varying channels," in *Proceedings of the 34th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1772–1776, Pacific Grove, Calif, USA, November 2000.

[5] E. F. Haratsch, A. J. Blanksby, and K. Azadet, "Reduced-state sequence estimation with tap-selectable decision-feedback," in *Proceedings of IEEE International Conference on Communications (ICC '00)*, vol. 1, pp. 372–376, New Orleans, La, USA, June 2000.

[6] A. A. Rontogiannis and K. Berberidis, "Efficient decision feedback equalization for sparse wireless channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 570–581, 2003.

[7] N. Benvenuto and R. Marchesani, "The Viterbi algorithm for sparse channels," *IEEE Transactions on Communications*, vol. 44, no. 3, pp. 287–289, 1996.

[8] N. C. McGinty, R. A. Kennedy, and P. A. Hoeher, "Parallel trellis Viterbi algorithm for sparse channels," *IEEE Communications Letters*, vol. 2, no. 5, pp. 143–145, 1998, see also: N. C. McGinty, R. A. Kennedy, and P. A. Hoeher, "Equalization of sparse ISI channels using parallel trellises," in *Proceedings of 7th Communication Theory Mini-Conference* in conjunction with *IEEE Globecom '98*, pp. 65–70, 1998.

[9] N. C. McGinty, "Reduced complexity equalization for data communication," Ph.D. dissertation, Canberra, Australia, 1997.

[10] F. K. H. Lee and P. J. McLane, "Iterative parallel-trellis MAP equalizers with nonuniformly-spaced prefilters for sparse multipath channels," in *Proceedings of 56th IEEE Vehicular Technology Conference (VTC '02)*, vol. 4, pp. 2201–2205, Vancouver, BC, Canada, September 2002.

[11] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 363–378, 1972.

[12] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, 1974.

[13] S. Badri-Hoeher, "Digitale Empfängeralgorithmen für TDMA-Mobilfunksysteme mit besonderer Berücksichtigung des EDGE-Systems," Ph.D. dissertation, Erlangen, Germany, 2001.

[14] W. H. Gerstacker, F. Obernosterer, R. Meyer, and J. B. Huber, "On prefilter computation for reduced-state equalization," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 793–800, 2002.

[15] S. Badri-Hoeher and P. A. Hoeher, "Fast computation of a discrete-time whitened matched filter based on Kalman filtering," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2417–2424, 2004.

[16] J. Hagenauer, "A soft-in/soft-out list sequential (LISS) decoder for turbo schemes," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '03)*, p. 382, Kanagawa, Japan, June-July 2003.

[17] C. Kuhn, "A bidirectional list-sequential (BI-LISS) equalizer for turbo schemes," in *Proceedings of the 14th IST Mobile & Wireless Communications Summit*, Dresden, Germany, June 2005, paper no. 306.

[18] L. Liu, W. K. Leung, and L. Ping, "Simple iterative chip-by-chip multiuser detection for CDMA systems," in *Proceedings of 57th IEEE Vehicular Technology Conference (VTC '03)*, vol. 3, pp. 2157–2161, Orlando, Fla, USA, October 2003.

[19] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 28–41, 2004.

[20] G. Colavolpe and G. Germi, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 818–825, 2005.

[21] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *European Transactions on Telecommunications and Related Technologies*, vol. 6, no. 5, pp. 507–511, 1995.

[22] J. Park and S. B. Gelfand, "Turbo equalizations for sparse channels," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC '04)*, vol. 4, pp. 2301–2306, Atlanta, Ga, USA, March 2004.

[23] R. Cusani and J. Mattila, "Equalization of digital radio channels with large multipath delay for cellular land mobile applications," *IEEE Transactions on Communications*, vol. 47, no. 3, pp. 348–351, 1999.

[24] M. V. Eyuboğlu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Transactions on Communications*, vol. 36, no. 1, pp. 13–20, 1988.

[25] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Transactions on Communications*, vol. 37, no. 5, pp. 428–436, 1989.

[26] K.-D. Kammeyer, "Time truncation of channel impulse responses by linear filtering: a method to reduce the complexity of Viterbi equalization," *AEÜ International Journal of Electronics and Communications*, vol. 48, no. 5, pp. 237–243, 1994.

[27] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, NJ, USA, 4th edition, 2002.

[28] N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: a unified approach," *IEEE Transactions on Information Theory*, vol. 42, no. 3, pp. 903–915, 1996.

[29] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 4th edition, 2001.

Jan Mietzner studied electrical engineering and information engineering at the Faculty of Engineering, University of Kiel, Germany, with focus on digital communications. During his studies, he spent six months in the year 2000 with the Global Wireless Systems Research Group, Lucent Technologies, Bell Labs UK, in Swindon, England. He received the Dipl.-Ing. degree from the University of Kiel in 2001. For his diploma thesis on space-time coding he received the Professor Dr. Werner Petersen-Award. Since August 2001, he has been working toward his Ph.D. degree as a Research Assistant at the Information and Coding Theory Lab (ICT), University of Kiel. His research interests concern physical layer aspects of future wireless communications systems, especially multiple-antenna techniques and space-time coding.

Sabah Badri-Hoeher received an M.S. degree ("licence en physique") from the University of Casablanca in 1991, and Dipl.-Ing. and Dr.-Ing.(Ph.D.) degrees in electrical engineering from the University of Paderborn and the University of Erlangen-Nuremberg, Germany, in 1996 and 2001, respectively. From October 1996 to July 2004, she was with the Fraunhofer Institute for Integrated Circuits (IIS-A) in Erlangen, Germany. Since January 2003, she has been with the Faculty of Engineering at the University of Kiel, Germany. Her research interests are in the general area of communications technology. She received the Fraunhofer-Award in 1999.

Ingmar Land is Assistant Professor for communication theory at Aalborg University, Denmark. His research topics are channel coding, iterative decoding, and information theory. He received his Dr.-Ing. degree at the University of Kiel, Germany, in 2004, where he was employed as Research and Teaching Assistant. He studied electrical engineering at the Universities of Ulm and Erlangen-Nuremberg, Germany, where he received his Dipl.-Ing. degree in 1999.

Peter A. Hoeher received Dipl.-Ing. and Dr.-Ing. (Ph.D.) degrees in electrical engineering from the Technical University of Aachen, Germany, and the University of Kaiserslautern, Germany, in 1986 and 1990, respectively. From October 1986 to September 1998, he was with the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. From December 1991 to November 1992, he was on leave at AT&T Bell Laboratories, Murray Hill, NJ. In October 1998 he joined the University of Kiel, Germany, where he is currently a Professor of electrical engineering. His research interests are in the general area of communication theory with applications in wireless communications and underwater communications, including digital modulation techniques, channel coding, iterative processing, equalization, multiuser detection, interference cancellation, and channel estimation-subjects on which he has published more than 130 papers and filed 12 patents. Dr. Hoeher received the Hugo-Denkmeier-Award '90. Between 1999 and 2004 he served as an Associated Editor for IEEE Transactions on Communications. He is a frequent consultant for the telecommunications industry.