Conditioned pilots for ISI channels

Future Network and MobileSummit 2013 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2013
ISBN: 978-1-905824-37-3

Conditioned Pilots for ISI Channels

Nicolò MAZZALI, Giulio COLAVOLPE
SPADiC Lab, Dep. of Information Engineering, University of Parma
viale delle Scienze 181/A, 43124 Parma, Italy
Email: nicolo.mazzali@unipr.it, giulio@unipr.it

Abstract: One of the proposals to increase the spectral efficiency of the DVB-S2 standard is based on time-frequency packing. This technique causes intersymbol and interchannel interferences to arise, requiring a significant growth of the number of pilots used to carry out frequency and phase synchronization. Therefore, a new pilot design will be introduced and suited optimal and suboptimal reduced-complexity algorithms derived. We will show that the proposed pilot strategy may outperform the classical one in terms of bit error rate and spectral efficiency.

Keywords: Pilots, spectral efficiency, time-packing, factor graphs, reduced-complexity algorithms.

1. Introduction

A satisfactory frequency and phase synchronization is a mandatory requirement for all kinds of practical wireless systems. Carrier synchronization is often performed through the aid of some pilot symbols periodically inserted in the transmitted data stream (see, e.g., the DVB-S2 standard [1]). To gain an insight (far from being exhaustive), the reader is referred to [2, 3, 4, 5] and references therein. As far as pilot symbols are concerned, their optimal position inside the data packet has been object of a thorough study in [6] where it has been shown that, under mild conditions, equally-spaced single pilots are one of the possible optimal configurations in the sense that they minimize the Cramér-Rao bound for channel estimation. Moreover, in [3] it is shown that arranging pilots in clusters induces a substantial performance penalty on a channel with additive white Gaussian noise (AWGN) and Wiener phase noise.

For any kind of communication system, one of the merit figures that must be reckoned with during the system design process is certainly the spectral efficiency (SE). In a multi-user scenario it has been shown that for linear modulations it is possible to increase the SE of the system simply giving up the orthogonality condition among users and packing them in the time and frequency domains [7, 8, 9]. This procedure causes known inter-symbol interference (ISI) and inter-channel interference (ICI) to arise. The presence of ISI implies that phase and frequency synchronization must be performed through clusters of pilots. These clusters must be at least longer than the channel memory in order to force the channel state. This pilot insertion obviously induces an energy loss and a spectral efficiency degradation. Moreover, since multiple clusters distributed all over the data packet allow a more reliable estimation than concentrated pilots [6], the resulting penalties may be significant.

In this paper, we propose a new design of the pilot symbols aiming at minimizing the overhead and guaranteeing good performance on ISI linear[1]channels. The main idea is to give up on pilot clusters and use instead equally-spaced, time-varying, data-dependent, isolated pilots, allowing a dramatic reduction of the overhead and of the consequent wasted energy and bandwidth. The value assumed by each pilot depends

[1]Nonlinear effetcs, present in satellite channels, will be addressed in future works.

on the L previous and the L following data symbols, where L is the size of the channel memory. This dependence causes an increase in the number of possible states and an expansion of the optimal detector trellis, but permits the receiver to observe, at sample epochs corresponding to pilots, a known value that can be exploited during the synchronization process. Since the optimal detector for ISI channels has a complexity which grows exponentially with the size of the channel memory, in case of severe ISI it becomes infeasible and a reduced-complexity solution is then proposed. We chose to adopt the Ungerboeck ISI model [10] because the implementation of the whitening filter (needed by the Forney ISI model [11]) is critical in several practical scenarios [12], and for applications when the detector is designed to cope only with a portion of the existing interference, a receiver working on the matched filter output results to be more robust to the unmanaged interference [8, 13]. Concerning the pilot definition, the adoption of the Forney model for pilots entails a dangerous increase in the pilot mean squared value (MSV), which translates into an energy loss that cannot be avoided, as it will be shown later. Since the Ungerboeck model appears to greatly reduce this MSV increase, and since it allows to get rid of the whitening filter, we chose to focus our analysis on the Ungerboeck model for pilots as well. Moreover, if isolated pilots are employed, the noise samples corrupting the useful part of the sampled received signal result to be approximately uncorrelated even though the Ungerboeck pilot model is adopted, provided that the spacing between two consecutive pilots is large enough.

2. System model

We consider a packet transmission on a linear channel where each packet contains a sequence of K symbols $\{a_k\}$ belonging to an M-ary alphabet $\mathcal{A}$ and a sequence of pilots $\{b_k\}$, which may not belong to $\mathcal{A}$, inserted every $P-1$ information symbols.[2] Focusing our investigation on linear modulations, the low-pass equivalent of the received signal reads

$$r(t) = s(t) + w(t) = \sum_{k=-\infty}^{+\infty} (a_k + b_k)\, p(t - kT) + w(t) \tag{1}$$

where T is the symbol period, $p(t)$ the shaping pulse (typically, a pulse with root-raised-cosine-shaped spectrum, denoted by RRC in the following), and $w(t)$ is a complex circularly-symmetric white Gaussian process with mean zero and variance $\sigma^2 = N_0$ per component. When $k = mP$, with $m \in \mathbb{N}_+$, the symbol a_{mP} is fictitious and only the pilot b_{mP} is transmitted. Conversely, for all other values of the time index k, b_k is fictitious and only the information symbol a_k is transmitted. The ISI coefficients are assumed to be known and in finite number. In order to limit the trellis expansion (as will be explained in the following), we consider only values of P higher than the duration of the channel memory. This implies that two consecutive pilots never interfere on each other. The resulting sufficient statistics [10] $\{r_k\}$ become

$$r_k = s_k + n_k = \sum_{\ell=-L}^{L} (a_{k-\ell} + b_{k-\ell})\, g_\ell + n_k \tag{2}$$

where $\{n_k\}$ are samples of a complex circularly-symmetric colored Gaussian process with mean zero and autocorrelation function $R_n(m) = 2\sigma^2 g_m$, and $\{g_\ell\}$ are the $2L+1$

[2] The number of pilots is thus $\lfloor K/(P-1) \rfloor$, where $\lfloor x \rfloor$ denotes the maximum integer lower than x.

Ungerboeck ISI coefficients.

We want the useful part of the received samples to have a constant and known value c at pilot epochs. In other words, for $k = mP$ we define the pilot as

$$b_{mP} = \frac{1}{g_0}\left(c - \sum_{\substack{\ell=-L \\ \ell\neq 0}}^{L} g_\ell a_{mP-\ell}\right) . \tag{3}$$

In order to limit the energy increase, we choose to impose

$$|c|^2 = \underset{|c|^2\geq\kappa}{\operatorname{argmin}} \left|\mathcal{E}_b - \mathcal{E}_a\right| \tag{4}$$

where $\mathcal{E}_a$ and $\mathcal{E}_b$ denote the MSV of the information symbols and of the pilots, respectively. The positive real constant κ is chosen in order to have observed samples with enough power at pilot epoch to perform reliable estimation.

An alternative design can be obtained by replacing the Ungerboeck pilot definition in (3) with the Forney pilot definition

$$b_{mP} = \frac{1}{f_0}\left(c - \sum_{\ell=1}^{L} f_\ell a_{mP-\ell}\right) \tag{5}$$

where $\{f_\ell\}$ are the $L+1$ Forney ISI coefficients.

3. Optimal algorithm on expanded trellis

In the following, we denote by $\mathbf{a} = (a_0, \ldots, a_{mP-1}, 0, a_{mP+1}, \ldots, a_{K-1})^T$ the vector of the data symbols, and similarly by $\boldsymbol{\sigma}$ and $\mathbf{r}$ the vectors of states[3] and received samples, respectively. The optimal maximum a posteriori probability (MAP) symbol detection strategy is

$$\hat{a}_k = \underset{a_k}{\operatorname{argmax}}\, P\left(a_k|\mathbf{r}\right) \tag{6}$$

where $P(a_k|\mathbf{r})$[4] may be evaluated by marginalizing the joint PMF $P(\mathbf{a}|\mathbf{r})$. This latter PMF can be obtained as $P(\mathbf{a}|\mathbf{r}) \propto P(\mathbf{a})p(\mathbf{r}|\mathbf{a})$, where $\propto$ denotes a proportionality relation, and the last term can be factored as [10, 14]

$$p\left(\mathbf{r}|\mathbf{a}\right) \propto \prod_{k=0}^{K-1} H_k\left(\mathbf{a}\right) . \tag{7}$$

In the following we decompose the time index as $k = mP+j$, with $j \in [0, P-1]$. When $j \in [L+2, P-1]$ or $j \in [1, L]$, factor $H_k(\mathbf{a})$ reads

$$H_k(\mathbf{a}) = \exp\left\{\frac{1}{\sigma^2}\Re\left[r_k a_k^* - \frac{1}{2}\left|a_k\right|^2 g_0 - \sum_{\substack{\ell=1 \\ \ell\neq j}}^{L} a_k^* a_{k-\ell} g_\ell\right]\right\} . \tag{8}$$

When $j \in [L+2, P-1]$, the state results to be $\sigma_k = (a_{k-1}, \ldots a_{k-L})$. On the contrary, when $j \in [1, L]$, it reads $\sigma_k = (a_{k-1}, \ldots, a_{mP+1}, a_{mP-1}, \ldots a_{mP-L})$. It is worth noting

[3] State σ_k is defined as the set of past symbols needed to completely describe the system, along with symbol a_k, at discrete-time k.

[4] By $p(.)$ we denote the probability density function (PDF) of a continuous random variable (RV), while by $P(.)$ we denote the probability mass function (PMF) of a discrete RV.

that symbols older than a_{k-L} are not directly used in (8), anyway they are present in the state because a pilot is in the system memory. When $j = 0$, $H_{mP}(\mathbf{a}) = 1$ and the state is $\sigma_{mP} = (a_{mP-1}, \ldots, a_{mP-L})$. Finally, when $j = L + 1$, factor $H_k(\mathbf{a})$ becomes

$$H_k(\mathbf{a}) = \exp\left\{\frac{1}{\sigma^2}\Re\left[r_k a_k^* - \frac{1}{2}|a_k|^2 g_0 - \sum_{\ell=1}^{L} a_k^* a_{k-\ell} g_\ell\right]\right\} \cdot$$
$$\cdot \exp\left\{\frac{1}{\sigma^2}\Re\left[r_{mP} b_{mP}^* - \frac{1}{2}|b_{mP}|^2 g_0 - \sum_{\substack{\ell=-L \\ \ell\neq 0}}^{L} b_{mP}^* a_{mP-\ell} g_\ell\right]\right\} \tag{9}$$

and the state is $\sigma_{mP+L+1} = (a_{mP+L}, \ldots, a_{mP+1}, a_{mP-1}, \ldots a_{mP-L})$. We can now replace the symbol vector $\mathbf{a}$ in all factors $\{H_k(\mathbf{a})\}$ with the couple (a_k, σ_k).

Since the a posteriori probability (APP) $P(a_k|\mathbf{r})$, needed for the MAP symbol detection strategy in (6), can be obtained also marginalizing the joint PMF $P(\mathbf{a}, \boldsymbol{\sigma}|\mathbf{r})$, we choose to use the factorization $P(\mathbf{a}, \boldsymbol{\sigma}|\mathbf{r}) \propto p(\mathbf{r}|\mathbf{a}, \boldsymbol{\sigma}) P(\boldsymbol{\sigma}|\mathbf{a}) P(\mathbf{a})$, where each term can be further factored as

$$P(\mathbf{a}) = \prod_{k=0}^{K-1} P(a_k)$$
$$P(\boldsymbol{\sigma}|\mathbf{a}) = P(\sigma_0) \prod_{k=1}^{K-1} P(\sigma_k|\sigma_{k-1}, a_{k-1}) = P(\sigma_0) \prod_{k=1}^{K-1} I(\sigma_k, \sigma_{k-1}, a_{k-1})$$
$$p(\mathbf{r}|\mathbf{a}, \boldsymbol{\sigma}) \propto \prod_{k=0}^{K-1} H_k(a_k, \sigma_k) \tag{10}$$

being $I(.)$ an indicator function equal to one when σ_k, σ_{k-1}, and a_{k-1} satisfy the trellis constraint, and equal to zero otherwise. From (10), it is possible to derive the corresponding factor graph (FG), represented in Fig. 1. Applying the sum-product algorithm (SPA) to the FG, we will be able to compute the marginal APPs needed for the MAP symbol detection strategy in (6). In Fig. 1, we denote by $G_k = G_k(a_k, \sigma_k, \sigma_{k+1}) =$

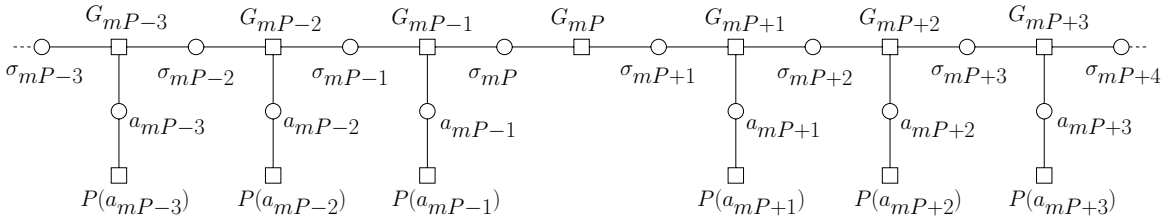

Figure 1: Factor graph for the optimal algorithm for $P > 3$.

$H_k(a_k, \sigma_k) I(\sigma_k, \sigma_{k+1}, a_k)$. The optimal MAP symbol detector is therefore the classic Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [15] running on a time-varying trellis. It can be shown that the factorization (10) and the corresponding FG in Fig. 1 are independent of the models (Forney's or Ungerboeck's) adopted for pilots and by the detector. On the contrary, the trellis structure depend on the assumed pilot model. If pilots are defined according to (5), the number of states varies with k from M^L to M^{2L-1}, whereas with design (3) the cardinality of the state set reaches M^{2L}.

4. Suboptimal algorithm on reduced trellis

The suboptimal algorithm here derived can be easily extended to other combinations of ISI and pilot models, all having the same FG and the same trellis structure. In the following, we denote by ϱ_k a hidden variable playing a role similar to that played by state σ_k in the derivation of the optimal algorithm[5]. To avoid the trellis expansion, we define a new factorization

$$p(\mathbf{r}|\mathbf{a}) \propto \prod_{k=0}^{K-1} H_k(\mathbf{a}) = \prod_{k=0}^{K-1} C_k(\mathbf{a})$$

where all terms depending on b_{mP} are now taken into account in factor $C_{mP}(\mathbf{a})$. When $j \in [1, P-1]$, the new factor $C_k(\mathbf{a})$ reads

$$C_k(\mathbf{a}) = \exp\left\{\frac{1}{\sigma^2}\Re\left[r_k a_k^* - \frac{1}{2}|a_k|^2 g_0 - \sum_{\substack{\ell=1 \\ \ell\neq j}}^{L} a_k^* a_{k-\ell} g_\ell\right]\right\} \tag{11}$$

and when $j = 0$, $C_{mP}(\mathbf{a})$ becomes

$$C_{mP}(\mathbf{a}) = \exp\left\{\frac{1}{\sigma^2}\Re\left[r_{mP} b_{mP}^* - \frac{1}{2}|b_{mP}|^2 g_0 - \sum_{\substack{\ell=-L \\ \ell\neq 0}}^{L} b_{mP}^* a_{mP-\ell} g_\ell\right]\right\}. \tag{12}$$

When $j \in [L+1, P-1]$ (respectively, when $j \in [1, L]$), factor (11) depends on the L (respectively, $L-1$) previous symbols. They can be grouped forming the hidden variable $\varrho_k = (a_{k-1}, \ldots, a_{k-L})$ (respectively, $\varrho_k = (a_{k-1}, \ldots, a_{mP+1}, a_{mP-1}, \ldots, a_{k-L})$). When $j = 0$, $C_{mP}(\mathbf{a})$ depends not only on the L previous symbols, grouped in the ***present*** hidden variable $\varrho_{mP} = (a_{mP-1}, \ldots, a_{mP-L})$, but also on the L next symbols, which may be grouped in the ***future*** hidden variable[6] $\varrho_{mP+L+1} = (a_{mP+L}, \ldots, a_{mP+1})$. The hidden variable can take on only M^{L-1} (when $j \in [1, L]$) or M^L (when $j \notin [1, L]$) different values. Hence, we can replace the symbol vector $\mathbf{a}$ in (11) and (12) with the couple (a_k, ϱ_k), when $j \in [1, P-1]$, and with the couple $(\varrho_{mP}, \varrho_{mP+L+1})$, when $j = 0$.

Since the APPs $\{P(a_k|\mathbf{r})\}$ needed for the MAP symbol detection strategy in (6) can be obtained also marginalizing the joint PMF $P(\mathbf{a}, \boldsymbol{\varrho}|\mathbf{r})$, we choose to use the new factorization $P(\mathbf{a}, \boldsymbol{\varrho}|\mathbf{r}) \propto P(\mathbf{a})P(\boldsymbol{\varrho}|\mathbf{a})p(\mathbf{r}|\mathbf{a}, \boldsymbol{\varrho})$. The first two terms can be factored as in (10) (replacing σ_k with ϱ_k), whereas the last term can be factored as

$$p(\mathbf{r}|\mathbf{a}, \boldsymbol{\varrho}) = \prod_{k=0}^{K-1} p(r_k|a_k, \varrho_k) \propto \prod_{\substack{k=0 \\ k\neq mP}}^{K-1} C_k(a_k, \varrho_k) \prod_{m=1}^{\left\lfloor \frac{K-1}{P} \right\rfloor} C_{mP}(\varrho_{mP}, \varrho_{mP+L+1}). \tag{13}$$

From (13) it is possible to derive the FG of the suboptimal algorithm, presented in

[5]We define ϱ_k as a set of past symbols, different from a_k, needed to compute factor $C_k(\mathbf{a})$. We will see later that ϱ_k is not a proper state since the couple (a_k, ϱ_k) is not enough to perfectly describe the system in any discrete-time k.

[6]The future hidden variable definition, found considering $C_{mP}(\mathbf{a})$ in (12), is identical to the present hidden variable definition that can be obtained considering $C_{mP+L+1}(\mathbf{a})$ in (11). Therefore, the hidden variable is well defined for every discrete-time k and no conflicts arise.

Fig. 2, where we defined

$$D_k = \begin{cases} D_k\left(a_k, \varrho_k, \varrho_{k+1}\right) & = C_k\left(a_k, \varrho_k\right) I\left(\varrho_k, \varrho_{k+1}, a_k\right) \quad \text{if } k \neq mP \\ D_k\left(\varrho_k, \varrho_{k+1}, \varrho_{k+L+1}\right) & = C_k\left(\varrho_k, \varrho_{k+L+1}\right) I\left(\varrho_k, \varrho_{k+1}, a_k\right) \quad \text{if } k = mP\,. \end{cases}$$

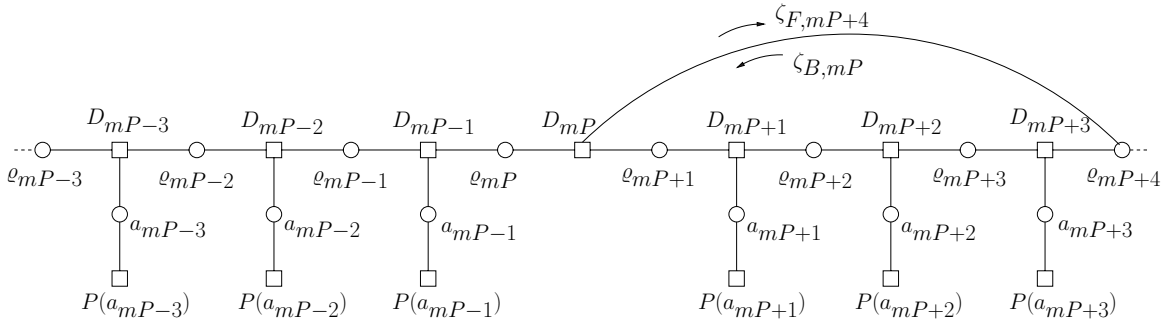

Figure 2: Factor graph for the suboptimal algorithm, with $L = 3$ and $P \geq 4$.

The FG on which the suboptimal algorithm runs and the hidden variable dimension are independent of the pilot design. The dependence of node D_{mP} on future hidden variable ϱ_{mP+L+1} introduces cycles in the FG, as shown in Fig. 2. The presence of cycles yields an approximated computation of the symbol APPs, but since in the cases of practical interest (i.e., ISI channels with $L > 1$) the girth of the graph is $2(L+1) > 4$, their convergence to the exact APPs is expected [16].

Since the graph has cycles, the SPA does not have a natural termination but a proper schedule must be defined. We denote $\zeta_{F,mP+L+1}$ and $\zeta_{B,mP}$ as the messages going forward and backward (respectively) on the upper branch of the FG and connecting the function node D_{mP} to the hidden variable ϱ_{mP+L+1}. Since the main structure of the FG in Fig. 2 is almost identical to the Wiberg graph of the BCJR algorithm [16], the SPA applied to the FG [16] will produce a slightly modified instance of the BCJR algorithm. The adopted schedule is therefore the following:

1. forward recursion of the BCJR algorithm; during the forward recursion, when $k = mP + L + 1$ the message $\zeta_{F,mP+L+1}$ is computed;

2. backward recursion of the BCJR algorithm; during the backward recursion, when $k = mP$ the message $\zeta_{B,mP}$ is computed;

3. update of the messages $\zeta_{F,mP+L+1}$;

4. completion of the BCJR algorithm considering also the contribution of messages $\zeta_{F,mP+L+1}$.

Since we consider serially concatenated schemes, we propose to perform a single detector iteration and then to pass the extrinsic information produced by the detector as a priori information to the decoder, in order to perform iterative detection and decoding.

5. Numerical results

In all simulations concerning the bit error rate (BER), we use packets of 2000 information bits, a spread interleaver, a non-systematic non-recursive convolutional code with

rate 1/2, polynomial generators $[5, 7]_8$ and four states, a Gray mapper with symbol MSV $\mathcal{E}_a = 1$, a binary phase-shift keying (BPSK) modulator (i.e. $M = 2$), a RRC pulse with roll-off $\alpha = 0.2$, a maximum of 20 iterations between detector and decoder, and pilot insertion with period $P = 21$. In order to determine the ISI coefficients to be assumed, we compute the induced ISI taps of a time-packed signaling system with a RRC pulse, roll-off $\alpha = 0.2$ and $\tau = 0.5$, where τ is the time compression factor[7] [17]. Since these coefficients would be too many (theoretically infinite) for the implementation of the optimal detectors, we keep only the first $L+1 = 7$ taps of the Forney model. Moreover, in all the suboptimal detectors we employ $\sigma^2 = N_0 + N_I$, where N_0 is the one-sided power spectral density of the AWGN and N_I is a parameter, independent of N_0, optimized via numerical simulation aiming at minimizing the BER.[8]

In order to do some comparisons, we consider also a system employing pilots defined according to (5), a system without pilots, and two systems which entail pilot insertion in blocks of pilots. For both Ungerboeck (3) and Forney (5) pilots, we also considered detectors based on Forney [11] and Ungerboeck [10] ISI models.[9] So as to be fair, we need to make comparisons among systems having the same synchronization capability. The bottle-neck of the synchronization is the carrier estimation, and being it dependent on the spacing between pilots [5], we keep constant the number of symbols between two consecutive pilot insertions. For the systems with pilot blocks, in order to reduce the overhead, we consider only blocks of size $N_p = L+1$. This is the minimum size allowing to have one known observed sample to be exploited for synchronization, the previous L pilots being necessary to force the state of the channel. For the first system with block pilots, we consider random pilots belonging to the symbol alphabet $\mathcal{A}$. Since the energy loss caused by the pilot insertion may be important, in the second considered system with block pilots we choose to set the L state forcing pilots to 0, and to $\mathcal{E}_a$ the $(L + 1)$-th pilot used for synchronization. The resulting MSVs and mean energies per symbol are shown in Table 1. Since the constraint $\mathcal{E}_a = \mathcal{E}_b$ cannot be satisfied, we

Table 1: Pilot MSVs and mean energies per symbol relative to different systems.

	For. pilots (5)	Ung. pilots (3)	bl. pilots $\in \mathcal{A}$/no pilots	bl. pilots $\notin \mathcal{A}$
$\mathcal{E}_b$	131.29	2.026	1	0.14
E_s	7.204	1.049	1	0.96

arbitrarily set $c = 1$.

System performance will not be evaluated only in terms of BER but also in terms of spectral efficiency. This latter can be computed as $\eta = I/(BT)$ (bits/s/Hz) where I is the information rate in bits per channel use and BT is the bandwidth normalized to the symbol period. The information rate of all systems is evaluated with the simulation-based technique described in [18] resorting to the corresponding optimal MAP symbol detector. For what concerns the normalized bandwidth, it is known and equal to $BT = \tau(1 + \alpha)$ for all the considered systems.

[7] τ is defined as the ratio between the used symbol interval and symbol interval for which the Nyquist condition for ISI absence is respected.

[8] N_I is a sort of extra noise variance considered only by the detector. It reduces the confidence of the BCJR algorithm in the computed messages, assuming the channel more noisy than it actually is. Therefore, it contributes to take into account the suboptimality caused by the cycles in the FG.

[9] The derivation of the optimal and suboptimal detection algorithms based on Forney ISI model is not shown for lack of space.

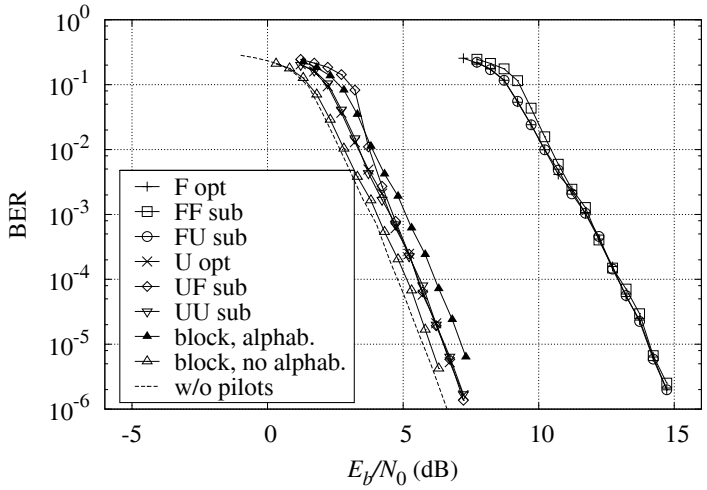

Figure 3: BER curves of the optimal and suboptimal detectors for Forney and Ungerboeck pilots, compared with curves of the systems with pilot blocks and without pilots.

The BER performance of all the investigated systems are shown in Fig. 3. In the legend we denote as "FU" the system using Forney pilots and Ungerboeck detection, "UU" the system completely based on the Ungerboeck model, "UF" the system with Ungerboeck pilots and Forney detection, and "FF" the system based on the Forney model. For both the types of pilots, the suboptimal detector based on the Ungerboeck model (marked as FU and UU) performs as the optimal one, while the detector based on the Forney model (marked as FF and UF) shows worse performance at low SNR. This behavior may be ascribed to the higher sensitivity of Forney detection to the suboptimality of the detection algorithm, as pointed out in [8].

All systems with Forney pilots present an impressive energy loss with respect to the system without pilots due to two different contributions. The first one is an obvious insertion loss due to the presence of pilots (which is also present in the systems with Ungerboeck pilots), while the second and predominant is a penalty due to the difference between the symbol and pilot MSVs. Since for the chosen channel $\mathcal{E}_a < \mathcal{E}_b^{(U)} \ll \mathcal{E}_b^{(F)}$ as shown in Table 1, the transmitter has to employ a lot of energy to transmit a Forney pilot. However, systems with Ungerboeck pilots outperform the traditional system with block pilots belonging to $\mathcal{A}$. On the other hand, the loss with respect to the system with block pilots not belonging to $\mathcal{A}$ is caused by the different mean energy per symbol E_s, as can be shown in Table 1.

The SE of systems with block pilots is heavily reduced, as reported in Fig. 4. It can be seen that the proposed systems greatly outperform systems with block pilots. If compared to the SE of the system without pilots, a loss due to the pilot insertion can be noticed. This loss may be reduced only by increasing the spacing between the consecutive pilots, which entails a reduction in the synchronization capability of the receiver.

6. Conclusions

We have proposed a new design for pilot symbols to be used for synchronization over channels with known ISI. Our pilots are time-varying, data-dependent, isolated, and

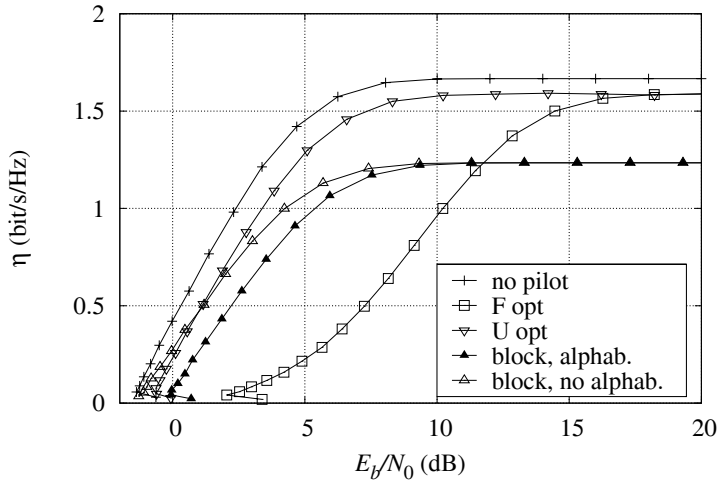

Figure 4: Spectral efficiencies of the systems with Forney pilots, Ungerboeck pilots, block pilots, and without pilots.

properly defined in order to make the detector receive, at pilot epochs, a known and constant value. For these pilots we derived the optimal MAP symbol detection algorithm, which turned out to run over a time-varying trellis with an extremely high number of states. Hence, we proposed a suboptimal reduced-complexity algorithm, whose BER performance is as good as the optimal one when Ungerboeck detection is adopted (a small penalty at low SNR may be seen when Forney detection is used). The remarkable complexity reduction has been obtained without resorting to any kind of modification of the joint PMF, but just rearranging factors in a proper way. With respect to the traditional pilots inserted in blocks, the proposed detectors gain in terms of BER and SE, and the choice of the design (Forney's or Ungerboeck's) entails great differences in the performance of the systems. Particularly, those with Ungerboeck pilots outperform the system with block pilots both in terms of BER and SE, and exhibit the appealing absence of whitening filters, often critical to design.

References

[1] ETSI EN 301 307 Digital Video Broadcasting (DVB); V1.1.2 (2006-06), Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other Broadband satellite applications, 2006. Available on ETSI web site (http://www.etsi.org).

[2] H. L. Van Trees, *Detection, Estimation, and Modulation Theory - Part I.* John Wiley & Sons, 1968.

[3] A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," *IEEE Trans. Commun.*, vol. 59, pp. 1966 –1974, july 2011.

[4] A. Barbieri and G. Colavolpe, "On the Cramer-Rao bound for carrier frequency estimation in the presence of phase noise," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 575–582, Feb. 2007.

[5] A. Barbieri and G. Colavolpe, "On pilot-symbol-assisted carrier synchronization for DVB-S2 systems," *IEEE Trans. Broadcast.*, vol. 53, pp. 685–692, Sept. 2007.

[6] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 3055 – 3069, dec 2002.

[7] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes," *IEEE Trans. Commun.*, vol. 57, pp. 2951–2959, Oct. 2009.

[8] A. Modenini, G. Colavolpe, and N. Alagha, "How to significantly improve the spectral efficiency of linear modulations through time-frequency packing and advanced processing," in *Proc. IEEE Intern. Conf. Commun.*, pp. 3299–3304, June 2012.

[9] A. Piemontese, A. Modenini, G. Colavolpe, and N. Alagha, "Improving the spectral efficiency of nonlinear satellite systems through time-frequency packing and advanced processing." submitted to IEEE Trans. on Commun., available on *http://arxiv.org/abs/1301.4184*, Jan. 2013.

[10] G. Ungerboeck, "Adaptive maximum likelihood receiver for carrier-modulated data-transmission systems," *IEEE Trans. Commun.*, vol. com-22, pp. 624–636, May 1974.

[11] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, pp. 284–287, May 1972.

[12] A. Hafeez and W. E. Stark, "Decision feedback sequence estimation for unwhitened ISI channels with applications to multiuser detection," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1785–1795, Dec. 1998.

[13] F. Rusek and A. Prlja, "Optimal channel shortening for MIMO and ISI channels," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 810–818, Feb. 2012.

[14] G. Colavolpe and A. Barbieri, "On MAP symbol detection for ISI channels using the Ungerboeck observation model," *IEEE Commun. Letters*, vol. 9, pp. 720–722, Aug. 2005.

[15] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[16] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, pp. 498–519, Feb. 2001.

[17] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Trans. Inform. Theory*, vol. 55, pp. 764 –775, Feb. 2009.

[18] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3498–3508, Aug. 2006.

