ELSEVIER

Signal Processing 80 (2000) 1999–2015

SIGNAL PROCESSING

www.elsevier.nl/locate/sigpro

Burst mode equalization: optimal approach and suboptimal continuous-processing approximation☆

Elisabeth de Carvalho, Dirk T.M. Slock*

Institut EURECOM, 2229 route des Crêtes, B.P. 193, 06904 Sophia Antipolis Cedex, France

Received 5 January 1999; received in revised form 15 August 1999

Abstract

We consider a transmission by burst where the data are organized and sent by bursts. At each end of the burst of data, a sequence of symbols is assumed known, and the channel considered as constant over the burst duration. The optimal structure of the burst mode equalizers is derived. The class of linear and decision feedback equalizers is considered, as well the class of ISI cancelers that use past but also future decisions: for each class of equalizers the MMSE, the unbiased MMSE and the MMSE zero forcing versions are derived. Unlike in the continuous-processing mode, the optimal burst mode filters are time varying. The performance of the different equalizers are evaluated and compared to each other in terms of SNR and probability of error: these measures depend on the position of the estimated symbol and on the presence of known symbols. Finally, we show that, by choosing correctly the number and position of the known symbols, (time-invariant) continuous-processing filters applied to burst mode can be organized to give sufficiently good performance, so that optimal (time-varying) burst processing implementation can be avoided.

Zusammenfassung

Wir betrachten die Burst-Übertragung, wobei die Daten in Bursts organisiert und gesendet werden. Am Ende jedes Datenburst wird eine Symbolsequenz als bekannt und der Kanal wird als konstant über der Burstdauer vorausgesetzt. Die optimale Struktur des Burst-Modus-Entzerrers wird hergeleitet. Es wird sowohl die Klasse der linearen, entscheidungsrückgekoppelten Entzerrer als auch die Klasse der lCl-Canceler betrachtet die sowohl vergangene als auch zukünftige Entscheidungen benutzen. Für jede Entzerrerklasse werden der MMSE, der biasfrcic MMSE und MMSE nullerzwingende Versionen abgeleitet. Im Gegensatz zum kontinuierlichen Verarbeitungsmodus sind die optimalen Burstmodus-Filter zeitvariant. Die Leistung der verschiedenen Entzerrer wird bewertet und untereinander in Hinblick auf SNR und Fehlerwahrscheinlichkeit verglichen. Diese Maße hängen von der Position des geschätzten Symbols und der Präsens unbekannter Symbole ab. Schließlich zeigen wir, dass durch die richtige Wahl der Nummer und Position der bekannten Symbole, (zeitinvariante) kontinuierlich verarbeitende Filter angewandt auf den Burstmodus derartig organisiert werden können, dass hinreichend gute Ergebnisse erzielt werden können, so dass optimale (zeitvariante) burstverarbeitende Implementierungen umgangen werden können.

☆The work of Elisabeth de Carvalho was supported by Laboratories d'Electronique Philips under contract Cifre 297/95. The work of Dirk Slock was supported in part by Laboratories d'Electronoque Philips under contract LEP 95FAR008.

* Corresponding author. Tel.: + 33 4 93 00 26 26; fax: + 33 4 93 00 26 27.

E-mail addresses: carvalho@eurecom.fr (E. de Carvalho), slock@eurecom.fr (D.T.M. Slock).

0165-1684/00/$ - see front matter
PII: S0165-1684(00)00066-9

Résumé

Nous considérons une transmission par rafales où les données sont organisées et envoyées en rafales. A chaque fin de la rafale de données, une séquence de symboles est supposée connue, et le canal est considéré constant sur la durée de la rafale. Nous dérivons la structure optimale de l'égaliseur de mode rafale. Nous considérons la classe des égaliseurs linéaires et à retour de décision, ainsi des éliminateurs ISI qui utilisent les décisions passées mais aussi futures: pour chaque classe d'égaliseur, la MMSE, la MMSE non biaisée et la variante MMSE Zero Forcing sont dérivées. Contrairement au mode de traitement continu, les filtres de mode en rafales optimaux sont variables dans le temps. Nous évaluons les performances des différents égaliseurs et nous les comparons entre eux en terme de rapport signal à bruit et de probabilité d'erreur: ces mesures dépendent de la position du symbole estimé et de la présence de symboles connus. Finalement, nous montrons que, en choisissant correctement le nombre et la position des symboles connus, des filtres de traitement continus (invariants dans le temps) appliqués au mode en rafale peuvent être organisés pour donner des performances suffisamment bonnes, de sorte que une implémentation de traitement de rafales optimale (variable dans le temps) peut être évitée.

Keywords: Burst mode equalization; Multichannel; Linear equalizer; Decision feedback equalizer; ISI canceler; Non-causal decision feedback equalizer; Unbiased MMSE equalizer

Nomenclature

$(\cdot)^*$	complex conjugate	$A_{i,j}$	element i,j of the matrix A
$(\cdot)^{\mathrm{T}}$	transpose	$\hat{\theta}$	estimate of parameter θ
$(\cdot)^{\mathrm{H}}$	Hermitian transpose	E	mathematical expectation
$(\cdot)^{-1}$	inverse	I	identity matrix of adequate dimension
A_i	element i of the vector A	w.r.t.	with respect to

1. Introduction

In most of the actual mobile communication systems, the data are divided and transmitted in bursts. In general, the bursts are separated by guard intervals, which avoid interburst interference, and contain known symbols, like synchronization bits or a training sequence to estimate the channel. This is typically the case of GSM (global system for mobile communications), where the channel is assumed constant over the duration of a burst and is estimated by a middamble training sequence, and the Viterbi algorithm is applied to estimate the transmitted data symbols.

We propose a scenario where a sequence of known symbols is attached to each end of the burst of information symbols. This scheme is proven to include the GSM case. The channel is assumed constant during the transmission of a burst. As we are operating with a finite amount of data, the usual time-invariant continuous-processing equalizers are not optimal anymore. We propose a derivation of the optimal burst mode equalizers, which are time varying. Three classes of equalizers are considered: the usual linear and decision feedback equalizers, as well as the ISI canceler. This last equalizer uses past but also future decisions and was proposed in its continuous-processing version in [10,7], and in its burst mode version in [11,6] where it is called non-causal decision feedback equalizer (NCDFE); in this paper, we use the term NCDFE to designate the ISI canceler. In a burst mode implementation of the NCDFE, a classical linear or decision feedback equalizer is used to give a first estimation of the symbols. Based on these decisions, the NCDFE computes new estimates and its output can be used to do other iterations. This NCDFE is potentially the most powerful

equalizer: indeed when no errors occur at the input of the non-causal feedback, all the ISI is removed and the matched filter bound (MFB) is attained.

These three classes of equalizers are derived according to three different criteria MMSE, unbiased MMSE and MMSE zero forcing (MMSE-ZF) corresponding to increasingly strong constraints: the first criterion is unconstrained, the second one is the element-wise best linear unbiased estimate (BLUE), and the third one is the block-wise BLUE. These three criteria will then give increasing MSEs. The MMSE equalizer gives biased estimates of the symbols: the Unbiased MMSE equalizer is the best equalizer in the MMSE sense, giving unbiased estimates. Although possessing a higher SNR than its MMSE counterpart, for non constant modulus symbol constellations, the unbiased MMSE equalizer gives a better error probability, because the decision device is built for unbiased symbol estimates. The unbiased MMSE DFE equalizer was introduced in [2]; we propose here a generalization to the other classes of equalizers.

All the equalizers are derived in a multichannel framework, which typically happens when there is an antenna array at the reception of a wireless system. We prove that the optimal processing consists in first removing the contribution of the known symbols, then applying the burst mode multichannel matched filter; the following filters depend on the specific equalizer considered. The performance of the different equalizers is evaluated in terms of SNR, studied according to the position of the unknown symbols in the burst and the presence of the known symbols.

In [8], burst mode MMSE and ZF equalizers are derived but for single channels: the ZF equalizer exists then only if there are at least a number of known symbols equal to the channel memory. In the multichannel context considered here, even with no known symbols, ZF equalizers exists and in fact a whole class of ZF equalizers: we will present the special class of MMSE-ZF versions of the equalizers which gives, among all the ZF equalizers, the lowest MSE. Furthermore, in [8], continuous-time matched filtering is done, matched to the overall channel, which is unrealizable, followed by the symbol rate burst mode processing. We follow the more realistic fractionnally spaced approach in which simple continuous-time lowpass filtering is followed by oversampling. Another interest of the multichannel model, which we will not detail in the paper, is that it allows better performance as the number of subchannels increases, and in particular, gives better performance than a single channel.

Ref. [8] presents complexity computations of the burst mode filters, which appear more complex than the time-invariant filter continuous processing mode. We propose to compare the performance obtained by applying the optimal time-varying burst mode filters with the performance obtained by applying the time-invariant continuous-processing filters to burst mode, which is not done in [8]. Refs. [4,3] proposed to enable time-invariant processing (with cyclic convolution though) by introducing cyclic prefixes. We propose to minimize the suboptimality of continuous processing by considering the influence of the pre and postamble lengths on the degradation between time-invariant filters and the optimal processing: the best situation happens when the lengths of these pre and postamble equals the channel memory. In [1], Al-Dhahir presents such a comparison, but considering the single-channel MMSE DFE only. His treatments of the known symbol is not correct however. He estimates the unknown symbols in terms of the received data only, whereas the correct treatment consists in estimating the unknown symbols in terms of the received data and also of the known symbols present in the burst. In his attempt to compare time-invariant and optimal processing fairly, he averages SNR in both cases over different amounts of symbols, estimating the known symbols also in the time-invariant processing, whereas the number of unknown symbols (to be equalized) is the same in both cases. So the comparison appears unfair. Furthermore, he summarizes the performance into one SNR average number over the burst: as will be seen in the paper, it appears important to consider on the contrary the performance as a function of symbol position.

The paper is organized as follows: in Section 2, the multichannel model is presented as well as our burst transmission model; in Section 3, the structure of the burst mode equalizers is derived and their performance studied in Section 4; at last in

Section 5, we explain how continuous processing is applied to burst mode in order to minimize suboptimality. This paper is an extented version of the conference presentation [12].

2. Problem formulation

2.1. Multichannel model

We consider a FIR multichannel model. This model applies to different cases: oversampling w.r.t. the symbol rate of a single received signal [13–15] or the separation into the real (in phase) and imaginary (in quadrature) parts of the demodulated received signal if the symbol constellation is real [9,17]. In the context of mobile digital communications, a third possibility appears in the form of multiple received signals from an array of sensors. These three sources for multiple channels can also be combined.

To further develop the case of oversampling, consider linear digital modulation over a linear channel with additive noise so that the cyclostationary received signal can be written as

$$\boldsymbol{y}(t) = \sum_k \boldsymbol{h}(t - kT)a(k) + \boldsymbol{v}(t), \tag{1}$$

where the $a(k)$ are the transmitted symbols, T is the symbol period and $h(t)$ is the channel impulse response. The channel is assumed to be FIR with duration NT (approximately). If the received signal is oversampled at the rate m/T (or if m different received signals are captured by m sensors every T seconds, or a combination of both), the discrete input–output relationship can be written as

$$\boldsymbol{y}(k) = \sum_{i=0}^{N-1} \boldsymbol{h}(i)a(k-i) + \boldsymbol{v}(k) = \boldsymbol{H}A_N(k) + \boldsymbol{v}_k,$$

$$\boldsymbol{y}(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_m(k) \end{bmatrix}, \quad \boldsymbol{v}(k) = \begin{bmatrix} v_1(k) \\ \vdots \\ v_m(k) \end{bmatrix}, \quad \boldsymbol{h}(k) = \begin{bmatrix} h_1(k) \\ \vdots \\ h_m(k) \end{bmatrix},$$

$$\boldsymbol{H} = [\boldsymbol{h}(N-1) \cdots \boldsymbol{h}(0)],$$

$$A_N(k) = [a^{\mathrm{H}}(k-N+1) \cdots a^{\mathrm{H}}(k)]^{\mathrm{H}}, \tag{2}$$

where the subscript j in the second equation denotes the jth channel. In the case of oversampling, $y_j(k)$, $j = 1, \ldots, m$, represent the m phases of the polyphase representation of the oversampled signal: $y_j(k) = y(t_0 + (k + j/m)T)$. In this polyphase representation, we get a discrete-time circuit in which the sampling rate is the symbol rate. Its output is a vector signal corresponding to a SIMO (single input multiple output) or vector channel consisting of m SISO discrete-time channels where m is the sum of the oversampling factors used for the possibly multiple antenna signals. Let $\boldsymbol{H}(z) = \sum_{i=0}^{N-1} \boldsymbol{h}(i)z^{-i} = [\mathrm{H}_1^{\mathrm{H}}(z) \cdots \mathrm{H}_m^{\mathrm{H}}(z)]^{\mathrm{H}}$ be the SIMO channel transfer function. Consider additive-independent white Gaussian noise $\boldsymbol{v}(k)$ with $r_{vv}(k-i) = \mathrm{E}\boldsymbol{v}(k)\boldsymbol{v}^{\mathrm{H}}(i) = \sigma_v^2 I_m \, \delta_{ki}$. Assume we receive M samples:

$$\boldsymbol{Y}_M(k) = \mathscr{T}_M(\boldsymbol{H})\, A_{M+N-1}(k) + \boldsymbol{V}_M(k) \tag{3}$$

where $\boldsymbol{Y}_M(k) = [\boldsymbol{y}^{\mathrm{H}}(k-M+1) \cdots \boldsymbol{y}^{\mathrm{H}}(k)]^{\mathrm{H}}$ and similarly for $\boldsymbol{V}_M(k)$; $\mathscr{T}_M(\boldsymbol{H})$ is a block Toeplitz matrix with M block rows and $[\boldsymbol{H}\ 0_{m\times(M-1)}]$ as first block row.

The channel length is assumed to be N which implies $\boldsymbol{h}(0) \neq 0$ and $\boldsymbol{h}(N-1) \neq 0$ whereas the impulse response is zero outside of the indicated range.

Multichannels present a certain number of advantages w.r.t. single channels. The performance of the equalizers is better for multichannels than for single channels: as the number of subchannels increases, performance gets better. In the case of reception by multiple antennas, this can be intuitively explained: each time a subchannel is added, diversity is added. In fact, if the number of independent subchannels increases, the multichannel tends to an allpass transfer function, in which case the SNR at the output of a simple MMSE-ZF linear equalizer becomes equal to the matched filter bound. This improvement also manifests itself in the oversampling case (oversampling at a rate larger than the Nysquist frequency does not improve performance any longer however). Furthermore, as detailed in Section 2.3, under certain conditions on the channel, ZF equalizers exist for multichannels but not for single channels for a number of known symbols present in the burst lower than the channel memory.

2.2. Burst transmission

We consider a transmission by burst in which detection is done burst by burst. We assume that the channel is time-invariant during the transmission of a burst. In the input burst denoted A, a pre and post-amble sequence of known symbols of variable length is attached to the burst of data symbols: n_1 known symbols at the beginning, grouped in the vector A_{K_1}, and n_2 at the end, grouped in the vector A_{K_2}: see Fig. 1. The total length of the input burst is $L + n_1 + n_2$; we want to detect the L central unknown symbols, grouped in the vector A_U. For that purpose, we consider as observation data $\boldsymbol{Y}$, the channel outputs that contain only symbols of burst A (the symbols to be detected or the known symbols of the burst), and not outputs containing symbols of neighboring bursts: see Fig. 1. The input–output relationship (3) between the observation data $\boldsymbol{Y}$ and A is written in simplified notation as

$$\boldsymbol{Y} = \mathscr{T} A + \boldsymbol{V}. \tag{4}$$

More data could be considered also, but this possibility will not be explored in this paper.

In the following, we consider the decomposition:

$$\boldsymbol{Y} = \mathscr{T} A + \boldsymbol{V} = \mathscr{T}_{K_1} A_{K_1} + \mathscr{T}_U A_U + \mathscr{T}_{K_2} A_{K_2} + \boldsymbol{V} = \mathscr{T}_K A_K + \mathscr{T}_U A_U + \boldsymbol{V}, \tag{5}$$

where $\mathscr{T}_i A_i$ represents the contribution of the symbols in A_i. $A_K = [A_{K_1}^{\mathrm{H}} A_{K_2}^{\mathrm{H}}]^{\mathrm{H}}$ groups all the known symbols. As will be seen, the optimal process consists first in removing the contribution of the known symbols, all the filters will then be applied to the processing data $\boldsymbol{Y}_U$:

$$\boldsymbol{Y}_U = \mathscr{T}_U A_U + \boldsymbol{V} = \boldsymbol{Y} - \mathscr{T}_K A_K. \tag{6}$$

It has to be noted that the derivations of the paper are valid for any position for the known symbols.

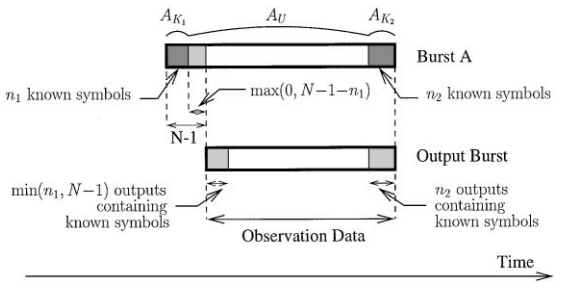

Fig. 1. Burst transmission.

2.3. Some useful properties of $\mathscr{T}$ and $\mathscr{T}_U$

If $M \geqslant N$ (with notations of Eq. (3)), the convolution matrix $\mathscr{T} = \mathscr{T}_M(\boldsymbol{H})$ is a tall matrix and has full-column rank if and only if the multichannel $\boldsymbol{H}(z)$ is irreducible [16]. A multichannel is said to be irreducible if it has no zeros, i.e. $\forall z$, $\boldsymbol{H}(z) \neq 0$. This is also equivalent to say that the subchannels of $\boldsymbol{H}(z)$ have no zero in common.

When $\boldsymbol{H}(z)$ is irreducible, $\mathscr{T}$ admits then a left inverse F, such that $F\mathscr{T} = I$: this result will be used for the derivation of ZF equalizers. In the case of a single channel, $\mathscr{T}$ has more columns than rows and admits no left inverses.

Now, assume that the multichannel $\boldsymbol{H}(z)$ is not irreducible and admits N_z zeros, then the column rank of $\mathscr{T}$ is $M + N - 1 - N_z$ ($M + N - 1$ is the number of columns of $\mathscr{T}$). If there are N_z known symbols in the burst, then $\mathscr{T}_U$ has full column rank [5] and admits a left inverse. For a single channel (which can be seen as a limiting case of a multichannel for which all the zeros are in common), provided that there are $N - 1$ zeros in the burst, $\mathscr{T}_U$ has full column rank and admits then a left inverse.

3. Burst-mode equalizers

In this section, we derive the expressions for the different equalizers in burst mode. Linear equalizers (LE), classical decision feedback equalizers (DFE) and the non-causal DFE (NCDFE) are considered for the minimum mean squared rrror (MMSE), the unbiased MMSE ($\mathscr{U}$MMSE) and the MMSE zero-forcing (MMSE ZF) criteria.

The different equalizers are linear estimators of the input symbols.

- Linear equalizers give linear estimates based on the received data $\boldsymbol{Y}$ and the known symbols A_K.

- DFEs give linear estimates based on $\boldsymbol{Y}$, A_K, as well as the decisions on the past input symbols.
- The NCDFE gives linear estimates given the $\boldsymbol{Y}$, A_K and the decisions on the past and future input symbols.

We shall assume those past (and future) decisions to be error-free.

The different equalizers are solutions of the MSE criterion

$$\min_F \|A_U - F\boldsymbol{Y}'\|^2, \tag{7}$$

where F is a matrix filled out with filter coefficients and $\boldsymbol{Y}'$ groups the whole observation set (e.g. $\boldsymbol{Y}$ and A_K for the LE), under different constraints:

- MMSE: no constraints.
- $\mathscr{U}$MMSE: element-wise best linear unbiased estimate (BLUE).
- MMSE zero-forcing: burst-wise BLUE.

In burst mode, the equalizer filters are time varying. We define the MSE of the ith symbol as

$$\mathrm{MSE}_i = (E(\hat{A}_U - A_U)(\hat{A}_U - A_U)^{\mathrm{H}})_{ii} \tag{8}$$

where $\hat{A}_U$ is the vector estimate of the unknown input symbols and the signal-to-noise ratio (SNR) of the ith symbol:

$$\mathrm{SNR}_i = \frac{\sigma_a^2}{\mathrm{MSE}_i}. \tag{9}$$

3.1. Linear equalizers

3.1.1. The MMSE linear equalizer

The MMSE LE gives the unconstrained MMSE estimate of the unknown symbols A_U based on the observations:

$$\boldsymbol{Y}'^{\mathrm{H}} = [\boldsymbol{Y}^{\mathrm{H}} \quad A_K^{\mathrm{H}}]^{\mathrm{H}}. \tag{10}$$

The linear MMSE estimate of A_U is

$$\hat{A}_{U,\,\mathrm{MMSE\ LE}} = R_{A_U\boldsymbol{Y}'}R_{\boldsymbol{Y}'\boldsymbol{Y}'}^{-1}\boldsymbol{Y}' = R_{A_U\boldsymbol{Y}_U}R_{\boldsymbol{Y}_U\boldsymbol{Y}_U}^{-1}\boldsymbol{Y}_U. \tag{11}$$

The last equality, proved in Appendix A, shows that linear estimation in terms of $\boldsymbol{Y}'$ is the same as in terms of $\boldsymbol{Y}_U$: the optimal processing can be seen as eliminating first the contributions of known symbols from the observation data $\boldsymbol{Y}$ to get $\boldsymbol{Y}_U$ and then applying the MMSE equalizer determined on the basis of $\boldsymbol{Y}_U$. For the other equalizers, the previous result is also true but will not be restated.

When a sequence of known symbols of length larger than the channel memory is present in the middle of the burst, like in GSM, the processing data $\boldsymbol{Y}_U$ can then be decomposed into two independent parts, and all the equalization process can be done on the two parts independently. This situation becomes equivalent to our proposed scenario of known symbols at each end of the burst.

From Eq. (11),

$$\begin{aligned}\hat{A}_{U,\mathrm{MMSE\ LE}} &= \sigma_a^2\mathscr{T}_U^{\mathrm{H}}(\sigma_a^2\mathscr{T}_U(h)\mathscr{T}_U^{\mathrm{H}}(h) + \sigma_v^2 I)^{-1}\boldsymbol{Y}_U \\ &= \left(\mathscr{T}_U^{\mathrm{H}}\mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2}I\right)^{-1}\mathscr{T}_U^{\mathrm{H}}\boldsymbol{Y}_U. \end{aligned}\tag{12}$$

The last equality is obtained via the matrix inversion lemma. We will denote

$$R = \mathscr{T}_U^{\mathrm{H}}\mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2}I. \tag{13}$$

Due to the presence of the regularizing term $(\sigma_v^2/\sigma_a^2)I$, the matrix R is invertible and the MMSE LE is always defined.

In the continuous-processing case, the MMSE equalizer gives the output:

$$\hat{a}_{\mathrm{MMSE\ LE}}(k) = \left(\boldsymbol{H}^{\dagger}(q)\boldsymbol{H}(q) + \frac{\sigma_v^2}{\sigma_a^2}I\right)^{-1}\boldsymbol{H}^{\dagger}(q)\boldsymbol{y}(k), \tag{14}$$

where $\boldsymbol{H}^{\dagger}(z) = \boldsymbol{H}^{\mathrm{H}}(1/z^*)$ and $q^{-1}\boldsymbol{y}(k) = \boldsymbol{y}(k-1)$. By analogy with the continuous-processing case, we can find interpretations for the expression (12) in filtering terms:

- $\mathscr{T}_U^{\mathrm{H}}$ represents the multichannel matched filter, matched to the filter $\mathscr{T}_U$. When the length of the two sequences of known symbols equals or is larger than the memory of the channel $N-1$, $\mathscr{T}_U^{\mathrm{H}}$ is Toeplitz, banded and upper triangular, which implies that the filtering is time-invariant, FIR and anticausal. When the length of the sequences are shorter however, the filter is time-varying at the edges.
- R is the FIR denominator of an IIR filter, R^{-1} is non-causal.

Fig. 2 shows the MMSE LE structure.

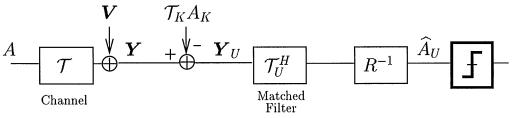

Fig. 2. Structure of the MMSE and the MMSE-ZF LE.

The UDL decomposition of $R = L^{\mathrm{H}}DL$ can be used to do a fast implementation of the MMSE LE as mentioned in [8]. The Schur algorithm can indeed be used to compute these factors. $R^{-1} = L^{-1}D^{-1}L^{-\mathrm{H}}$: the output of $L^{-\mathrm{H}}$, $Z' = L^{-\mathrm{H}}Z \Rightarrow LZ' = Z$, can be solved by back-substitution. The same kind of remark is valid for the output of L^{-1}. So it becomes not necessary to inverse R and the complexity is of order MN.

For the burst mode MMSE LE:

$$\mathrm{SNR}_i(\mathrm{MMSE\ LE}) = \frac{\sigma_a^2}{\sigma_v^2 (R^{-1})_{ii}}. \tag{15}$$

The SNR depends on the position of the symbol in the burst and we will see the influence of the known symbols on the SNR according to the position of the symbols to be estimated. This remark will also be valid for the other equalizers.

3.1.2. The general unbiased MMSE problem

A MMSE equalizer produces a biased estimate of the symbol $a(i)$: the MMSE equalizer output can indeed be written as $\alpha(i)a(i) + n(i)$, where $n(i)$ and $a(i)$ are uncorrelated ($n(i)$ contains symbols different from $a(i)$ and noise terms). This bias increases the probability of error [2], as the decision devices are made for unbiased data. The purpose of the unbiased MMSE equalizer is to correct this bias. We then derive the best equalizer, in the MMSE sense, that gives unbiased symbol estimates: we will see that its SNR gets reduced w.r.t. the MMSE, but that the error probability gets increased. Note that ZF equalizers are unbiased equalizers: they minimize the MSE under the unbiasness constraint but also the zero ISI constraint; the $\mathcal{U}$MMSE are derived under the unbiasness constraint only. So ZF and $\mathcal{U}$MMSE equalizers are different except when there is no ISI at the output of the $\mathcal{U}$MMSE, which will be the case for the NCDFE.

In terms of estimation theory, the unbiased MMSE equalizer is the element-wise BLUE. We give and prove here results that will be valid for all the Unbiased MMSE equalizers (LE, DFE, NCDFE).

Consider the estimation of symbol $a(i)$. $\boldsymbol{Y}'$ contains all the information available for estimation, $\boldsymbol{Y}_U$ and $\bar{A}$: $\bar{A}$ denotes here the past decisions w.r.t. $a(i)$ for the DFE, the past and future decisions for the NCDFE, and is zero for the LE. Let us decompose the processing data $\boldsymbol{Y}_U$ onto the contribution of $a(i)$ and of the other symbols $\bar{A}_{U,i}$.

$$\boldsymbol{Y}_U = \mathcal{T}_U A_U + \boldsymbol{V} = \mathcal{T}_{U,i} a(i) + \bar{\mathcal{T}}_{U,i} \bar{A}_{U,i} + \boldsymbol{V}, \tag{16}$$

$$\boldsymbol{Y}' = \begin{bmatrix} \mathcal{T}_{U,i} \\ 0 \end{bmatrix} a(i) + \begin{bmatrix} \bar{\mathcal{T}}_{U,i} \bar{A}_{U,i} + \boldsymbol{V} \\ \bar{A} \end{bmatrix}$$

$$= \mathcal{T}'_{U,i} a(i) + \boldsymbol{V}'. \tag{17}$$

The BLUE theory for this linear model gives us as estimate for $a(i)$:

$$\hat{a}_{\mathrm{BLUE}}(i) = (\mathcal{T}'^{\mathrm{H}}_{U,i} R'^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} \mathcal{T}'_{U,i})^{-1} \mathcal{T}'^{\mathrm{H}}_{U,i} R'^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} \boldsymbol{Y}' \tag{18}$$

which can also be written as

$$\hat{a}_{\mathrm{BLUE}}(i) = \sigma_a^2 (R_{a(i)\boldsymbol{Y}'} R^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} R_{a(i)\boldsymbol{Y}'})^{-1} \underbrace{R_{a(i)\boldsymbol{Y}'} R^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} \boldsymbol{Y}'}_{\hat{a}_{\mathrm{MMSE}}(i)}. \tag{19}$$

The unbiased estimate for the whole burst is then

$$\hat{A}_{U,\mathcal{U}\mathrm{MMSE}} = \underbrace{\sigma_a^2 (\mathrm{diag}(R_{A_U\boldsymbol{Y}'} R^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} R_{\boldsymbol{Y}'A_U}))^{-1}}_{\mathcal{D}} \underbrace{R_{A_U\boldsymbol{Y}'} R^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} \boldsymbol{Y}'}_{\hat{A}_{U,\mathrm{MMSE}}}, \tag{20}$$

$$\hat{A}_{U,\mathcal{U}\mathrm{MMSE}} = \mathcal{D} \hat{A}_{U,\mathrm{MMSE}}. \tag{21}$$

The unbiased MMSE equalizer is simply a scaled version of the MMSE equalizer.

The SNR of the $\mathcal{U}$MMSE is related to the SNR of the MMSE:

$$\mathrm{SNR}_i(\mathcal{U}\mathrm{MMSE}) = \mathrm{SNR}_i(\mathrm{MMSE}) - 1. \tag{22}$$

The proof can be found in Appendix B.

3.1.3. The unbiased MMSE linear equalizer

For the specific case of the MMSE LE:

$$R_{A_U Y'} R_{Y'Y'}^{-1} R_{Y' A_U} = \left(\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2} I \right)^{-1} \mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U$$

$$\Rightarrow \quad \mathscr{D} = \left(\mathrm{diag}\left(\left(\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2} I \right)^{-1} \mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U \right) \right)^{-1}. \tag{23}$$

$\mathscr{D}$ can be further rearranged, and we get

$$\hat{A}_{U,\mathscr{U}\mathrm{MMSE\ LE}} = \left(I - \frac{\sigma_v^2}{\sigma_a^2} \mathrm{diag}[(\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2} I)^{-1}] \right)^{-1} \hat{A}_{U,\mathrm{MMSE\ LE}}. \tag{24}$$

As $\mathscr{D}$ is invertible $\hat{A}_{U,\mathscr{U}\mathrm{MMSE\ LE}}$ is always defined. Fig. 3 shows the $\mathscr{U}$MMSE structure.

In the continuous-processing case, the output of the $\mathscr{U}$MMSE LE has for expression,

$$\hat{a}_{\mathscr{U}\mathrm{MMSE\ LE}}(k) = \left(1 - \frac{\sigma_v^2}{\sigma_a^2} \oint \frac{\mathrm{d}z}{z} \left(\boldsymbol{H}^{\dagger}(z) \boldsymbol{H}(z) + \frac{\sigma_a^2}{\sigma_v^2} \right) \right)^{-1} \hat{a}_{\mathrm{MMSE\ LE}}(k). \tag{25}$$

3.1.4. The MMSE-ZF linear equalizer

A ZF equalizer has for property to leave the signal part of the received data undistorted: a block ZF equalizer F verifies

$$F \mathscr{T}_U = I. \tag{26}$$

As seen in Section 2.3, in the monochannel case, the existence of the ZF equalizer is conditioned to the presence of known symbols. When there are no known symbols, $\mathscr{T}_U = \mathscr{T}$ admits no left inverse. For a number of known symbols of exactly $N-1$, the channel memory, $\mathscr{T}_U$ is square and there is a unique ZF equalizer which is also the MMSE ZF equalizer. For a number of known symbols of more than $N-1$, $\mathscr{T}_U$ is strictly tall and full-column rank and ZF equalizers exist.

For a multichannel and also for a single channel, $\mathscr{T}_U$ has full column rank if $M \geqslant N$ and if there are as many known symbols as the number of zeros. These will be the conditions for a ZF equalizer to exist.

When $\mathscr{T}_U$ is strictly tall and has full column rank, it admits several left inverses. Indeed, let $\mathscr{T}_U^{\perp}$ be a matrix which columns are orthogonal to those of $\mathscr{T}_U$, then $\mathscr{T}_U^{\perp \mathrm{H}} \mathscr{T}_U = 0$. $F = (\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U)^{-1} \mathscr{T}_U^{\mathrm{H}}$, the Moore–Penrose pseudo-inverse of $\mathscr{T}_U$ verifies $F \mathscr{T}_U = I$, but also $F = (\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U)^{-1} \mathscr{T}_U^{\mathrm{H}} + C \mathscr{T}_U^{\perp \mathrm{H}}$, where C is any $M \times M$ matrix.

We shall here concentrate on the MMSE-ZF LE, which give the lowest MSE among all the ZF LE equalizers. The MMSE-ZF LE corresponds to the block-wise BLUE based on $\boldsymbol{Y}_U$. Given the linear model: $\boldsymbol{Y}_U = \mathscr{T}_U A_U + \boldsymbol{V}$, the BLUE is given by

$$\hat{A}_{U,\mathrm{BLUE}} = (\mathscr{T}_U^{\mathrm{H}} R_{Y_U Y_U}^{-1} \mathscr{T}_U)^{-1} \mathscr{T}_U^{\mathrm{H}} R_{Y_U Y_U}^{-1} \boldsymbol{Y} = (\mathscr{T}_U^{\mathrm{H}} R_{VV}^{-1} \mathscr{T}_U)^{-1} \mathscr{T}_U^{\mathrm{H}} R_{VV}^{-1} \boldsymbol{Y}_U. \tag{27}$$

So the MMSE-ZF LE is

$$\hat{A}_{U,\mathrm{MMSE\text{-}ZF\ LE}} = (\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U)^{-1} \mathscr{T}_U^{\mathrm{H}} \boldsymbol{Y}_U. \tag{28}$$

Consider now the UDL decomposition of $\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U = L^{\mathrm{H}} D L$:

$$(\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U)^{-1} = L^{-1} D^{-1} L^{-\mathrm{H}}. \tag{29}$$

After the matched filter, the optimal process consists in whitening the noise by the filter $L^{-\mathrm{H}}$. We will find these two optimal steps (matched filtering and noise whitening) for all the ZF equalizers. If we denote now $R = \mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U$, the process is the same as for the MMSE LE (see Fig. 2). The remarks on the fast implementation are also valid here.

The output burst mode SNR is

$$\mathrm{SNR}_i(\mathrm{MMSE\text{-}ZF\ LE}) = \frac{\sigma_a^2}{\sigma_v^2 ((\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U)^{-1})_{ii}}. \tag{30}$$

In the continuous-processing case, the MMSE-ZF LE output is

$$\hat{a}_{\mathrm{MMSE\text{-}ZF\ LE}}(k) = (\boldsymbol{H}^{\dagger}(q) \boldsymbol{H}(q))^{-1} \boldsymbol{H}^{\dagger}(q) \boldsymbol{y}(k). \tag{31}$$

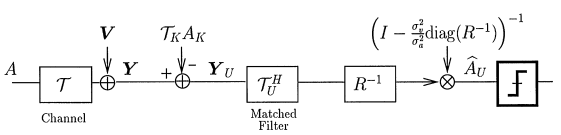

Fig. 3. Structure of the $\mathscr{U}$MMSE LE.

3.2. Decision feedback equalizers

3.2.1. The MMSE decision feedback equalizer

The decision feedback equalizers consider the linear estimation of symbol $a(i)$ based on the processing data $\boldsymbol{Y}_U$ and the past decisions w.r.t. $a(i)$ assumed known that we denote A_i^{p}:

$$\hat{a}_{\text{MMSE DFE}}(i) = F_i \boldsymbol{Y}_U - B_i A_i^{\mathrm{p}}, \tag{32}$$

where F_i is the forward filter and B_i the feedback filter. Let $\boldsymbol{Y}' = [\boldsymbol{Y}_U^{\mathrm{H}} \; A_i^{\mathrm{p}^{\mathrm{H}}}]^{\mathrm{H}}$, and let us decompose $\boldsymbol{Y}_U$ onto the contribution of A_i^{p} the past symbols and A_i^{f} grouping $a(i)$ and the future symbols:

$$\boldsymbol{Y}_U = \mathscr{T}_U^{\mathrm{p}} A_i^{\mathrm{p}} + \mathscr{T}_U^{\mathrm{f}} A_i^{\mathrm{f}} + \boldsymbol{V}, \tag{33}$$

$$\begin{aligned}[F_i \;\; -B_i] &= R_{a(i)\boldsymbol{Y}'} R_{\boldsymbol{Y}'\boldsymbol{Y}'}^{-1} \\ &= [\sigma_a^2 \mathscr{T}_{U,i}^{\mathrm{H}} (\sigma_a^2 \mathscr{T}_U^{\mathrm{f}} \mathscr{T}_U^{\mathrm{f}^{\mathrm{H}}} + \sigma_v^2 I)^{-1} \\ &\quad - \sigma_a^2 \mathscr{T}_{U,i}^{\mathrm{H}} (\sigma_a^2 \mathscr{T}_U^{\mathrm{f}} \mathscr{T}_U^{\mathrm{f}^{\mathrm{H}}} + \sigma_v^2 I)^{-1} \mathscr{T}_U^{\mathrm{p}}].\end{aligned} \tag{34}$$

Consider the UDL factorization of $R = \mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U + (\sigma_v^2/\sigma_a^2) I = L^{\mathrm{H}} D L$. After some manipulations, it can be proven that F_i is the ith row of $D^{-1} L^{-\mathrm{H}} \mathscr{T}_U^{\mathrm{H}}$ and that B_i the ith row of $L - I$. A proof for this result is provided in Appendix C.

The symbol estimate is then

$$\begin{aligned}&\hat{A}_{U,\text{MMSE DFE}} \\ &\quad = D^{-1} L^{-\mathrm{H}} \mathscr{T}_U^{\mathrm{H}} \boldsymbol{Y}_U - (L - I)\mathrm{dec}(\hat{A}_{U,\text{MMSE DFE}}).\end{aligned} \tag{35}$$

The MMSE DFE is always defined like the MMSE LE. The forward filter consists in the cascade of the multichannel matched filter and an anticausal filter $D^{-1} L^{-\mathrm{H}}$. $L - I$ is a strictly causal filter, so that the feedback operation involves only past decisions. Fig. 4 shows the structure of the MMSE DFE. As for the LEs, a fast implementation of the DFE using the UDL decomposition is also possible here [8]: the resulting complexity is of order MN.

The SNR is

$$\mathrm{SNR}_i(\text{MMSE DFE}) = \frac{\sigma_a^2}{\sigma_v^2 (D^{-1})_{ii}}. \tag{36}$$

In the continuous-processing case:

$$\begin{aligned}&\hat{a}_{\text{MMSE DFE}}(k) \\ &\quad = \frac{\mathbf{H}^{\dagger}(q)}{d G^{\dagger}(q)} \boldsymbol{y}_k - (G(q) - 1)\mathrm{dec}(\hat{a}_{\text{MMSE DFE}}(k))\end{aligned} \tag{37}$$

where $\boldsymbol{H}^{\dagger}(q)\boldsymbol{H}(q) + \sigma_v^2/\sigma_a^2 = G^{\dagger}(q) d G(q)$, $G(q)$ is causal and $G(\infty) = 1$.

3.2.2. The unbiased MMSE decision feedback equalizer

Using the results of Section 3.1.2, we can prove that the output of the unbiased MMSE DFE is

$$\hat{A}_{U,\mathscr{U}\text{MMSE DFE}} = \left(I - \frac{\sigma_v^2}{\sigma_a^2} D^{-1} \right)^{-1} \hat{A}_{U,\text{MMSE DFE}}. \tag{38}$$

Fig. 5 shows this structure. The burst output SNR is decreased by 1 with respect to the MMSE DFE.

The continuous-processing equalizer output is

$$\begin{aligned}\hat{a}_{\mathscr{U}\text{MMSE DFE}}(k) = \left(1 - \frac{\sigma_v^2}{\sigma_a^2} \exp\left(- \oint \frac{\mathrm{d}z}{z} \ln(\boldsymbol{H}^{\dagger}(z)\boldsymbol{H}(z) \right.\right. \\ \left.\left. + \frac{\sigma_a^2}{\sigma_v^2} \right) \right)^{-1} \hat{a}_{\text{MMSE DFE}}(k).\end{aligned} \tag{39}$$

3.2.3. The MMSE-ZF decision feedback equalizer

As for the ZF LE, there is a whole class of ZF equalizers, and we derive here the ZF MMSE DFE equalizer. Consider the UDL factorization of $\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U = L^{\mathrm{H}} D L$. Then the forward and feedback

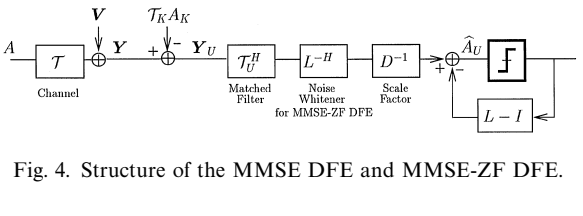

Fig. 4. Structure of the MMSE DFE and MMSE-ZF DFE.

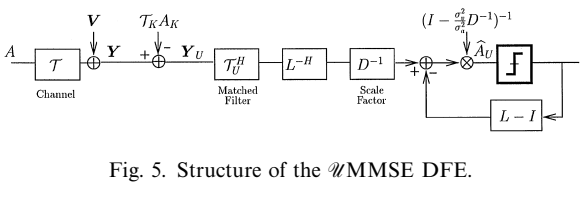

Fig. 5. Structure of the $\mathscr{U}$MMSE DFE.

filters are proven in Appendix D to be

$$F = L^{-H}D^{-1}\mathcal{T}_U^H, \qquad B = L - I, \tag{40}$$

we have the same structure as the MMSE DFE. The same equalizability conditions as for MMSE-ZF LE hold here also.

Let us end this section by noting that the expression of the MMSE and MMSE-ZF DFEs can be recovered from the LEs: for the MMSE LE, we consider the UDL factorization of $R = L^H DL$ and for the MMSE ZF LE, the UDL factorization of $\mathcal{T}_U^H\mathcal{T}_U = L^H DL$. The ouput of these two equalizers can then be written as

$$\begin{aligned}\hat{A}_U &= L^{-1}D^{-1}L^{-H}\mathcal{T}_U^H\boldsymbol{Y}_U \\ &= D^{-1}L^{-H}\mathcal{T}_U^H\boldsymbol{Y} - (L-I)\hat{A}_U.\end{aligned} \tag{41}$$

The DFE operation consists of taking $(L-I)\mathrm{dec}(\hat{A}_U)$ instead of $(L-I)\hat{A}_U$, where dec(.) is the decision operation.

3.3. Non-causal decision feedback equalizers

3.3.1. The MMSE NCDFE

The NCDFE considers the linear estimation of symbol $a(i)$ based on the processing data $\boldsymbol{Y}_U$ and the past and future decisions w.r.t. $a(i)$ assumed known that we denote $\bar{A}_{U,i}$. The burst mode equalizer is implemented in an iterative way. At the first iteration, the past and future decisions come from another classical LE or DFE. The output the NCDFE can then be used to reinitialized the NCDFE, and other iterations can be done. As for the DFE, we consider the past and future decisions as correct:

$$\hat{a}_{\text{MMSE NCDFE}}(i) = F_i\boldsymbol{Y}_U - B_i\bar{A}_{U,i}, \tag{42}$$

where F_i is the forward filter and B_i the feedback filter. Let $\boldsymbol{Y}' = [\boldsymbol{Y}_U^H \ \bar{A}_{U,i}^H]^H$,

$$\boldsymbol{Y}_U = \mathcal{T}_{U,i}a(i) + \bar{\mathcal{T}}_{U,i}\bar{A}_{U,i}, \tag{43}$$

$$[F_i \ -B_i] = R_{a(i)Y'}R_{Y'Y'}^{-1} \tag{44}$$

and we get

$$\begin{aligned}[F_i \ B_i] = [&\sigma_a^2\mathcal{T}_{U,i}^H(\sigma_a^2\mathcal{T}_{U,i}\mathcal{T}_{U,i}^H + \sigma_v^2 I)^{-1} \\ &\sigma_a^2\mathcal{T}_{U,i}^H(\sigma_a^2\mathcal{T}_{U,i}\mathcal{T}_{U,i}^H + \sigma_v^2 I)^{-1}\bar{\mathcal{T}}_{U,i}],\end{aligned} \tag{45}$$

$$\begin{aligned}[F_i \ B_i] = \Bigg[&\left(\mathcal{T}_{U,i}^H\mathcal{T}_{U,i} + \frac{\sigma_v^2}{\sigma_a^2}I\right)^{-1} \\ &\mathcal{T}_{U,i}^H\left(\mathcal{T}_{U,i}^H\mathcal{T}_{U,i} + \frac{\sigma_v^2}{\sigma_a^2}I\right)^{-1}\mathcal{T}_{U,i}^H\bar{\mathcal{T}}_{U,i}\Bigg].\end{aligned} \tag{46}$$

Then,

$$\begin{aligned}F &= \left(\mathrm{diag}\left(\mathcal{T}_U^H\mathcal{T}_U + \frac{\sigma_v^2}{\sigma_a^2}I\right)\right)^{-1}\mathcal{T}_U^H, \\ B &= \left(\mathrm{diag}\left(\mathcal{T}_U^H\mathcal{T}_U + \frac{\sigma_v^2}{\sigma_a^2}I\right)\right)^{-1} \\ &\quad\times(\mathcal{T}_U^H\mathcal{T}_U - \mathrm{diag}(\mathcal{T}_U^H\mathcal{T}_U)).\end{aligned} \tag{47}$$

The MMSE NCDFE has a very simple structure: the forward filter is proportional to the matched filter and the feedback filter to the cascade of the channel and the forward filter without the central coefficient. Fig. 6 shows the structure of the NCDFE.

All the ISI is removed if there are no errors in the non-causal feedback: the NCDFE attains then the matched filter bound. But, like the decision feedback equalizer, the NCDFE suffers from the error propagation phenomenon.

The burst mode SNR is

$$\mathrm{SNR}_i(\text{MMSE NCDFE}) = \frac{\sigma_a^2}{\sigma_v^2}\left(\mathcal{T}_U^H\mathcal{T}_U + \frac{\sigma_v^2}{\sigma_a^2}I\right)_{i,i}. \tag{48}$$

In the continuous-processing case.

$$\begin{aligned}&\hat{a}_{\text{MMSE NCDFE}}(k) \\ &\quad= (\boldsymbol{H}^\dagger(q)\boldsymbol{H}(q) + \sigma_v^2/\sigma_a^2)^{-1}(\boldsymbol{H}^\dagger(q)\boldsymbol{y}(k) \\ &\qquad- (\boldsymbol{H}^\dagger(q)\boldsymbol{H}(q) - \|\boldsymbol{H}\|^2)\mathrm{dec}(\hat{a}_{\text{MMSE NCDFE}}(k))).\end{aligned} \tag{49}$$

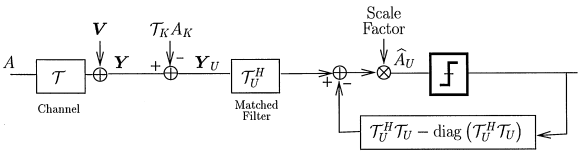

Fig. 6. Structure of the NCDFE.

3.3.2. The unbiased/ZF-MMSE NCDFE

As seen in Section 3.1.2, the unbiased MMSE estimate $\hat{A}_{\mathscr{U}\text{MMSE NCDFE}}$ is a scale version of $\hat{A}_{\text{MMSE NCDFE}}$. We find

$$F = (\text{diag}(\mathscr{T}_U^{\text{H}}\mathscr{T}_U))^{-1}\mathscr{T}_U^{\text{H}},$$

$$G = (\text{diag}(\mathscr{T}_U^{\text{H}}\mathscr{T}_U))^{-1}(\mathscr{T}_U^{\text{H}}\mathscr{T}_U - \text{diag}(\mathscr{T}_U^{\text{H}}\mathscr{T}_U)). \tag{50}$$

As all the ISI is removed by the NCDFE, the ZF NDFE and the $\mathscr{U}$MMSE NCDFE are the same. If $n_1 = n_2 = N-1$, the burst mode filters of the NCDFE are time-invariant. When $n_1 < N-1$ and $n_2 < N-1$, the filters vary only at the edges and are otherwise time-invariant. It appears that another interest of the burst mode NCDFE is that it is as easy to implement than its continuous-processing version.

The burst mode SNR is

$$\text{SNR}_i(\mathscr{U}/\text{ZF-MMSE NCDFE}) = \frac{\sigma_a^2}{\sigma_v^2}(\mathscr{T}_U^{\text{H}}\mathscr{T}_U)_{i,i}. \tag{51}$$

No special conditions are required for the ZF and MMSE NCDFEs to be defined as $\text{diag}(\mathscr{T}_U^{\text{H}}\mathscr{T}_U)$ and $\text{diag}\left(\mathscr{T}_U^{\text{H}}\mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2}I\right)$ are invertible.

4. Performance comparisons

In this section, we discuss the performance of the equalizers in terms of SNR and probability of error. In Fig. 7, the SNR curves are drawn for a channel $\boldsymbol{H}_1$ of length 7 with 3 subchannels which coefficients were randomly chosen ($\boldsymbol{H}_1$ is given in Appendix E). The SNR per channel is 10 dB. The input symbols are drawn from a BPSK ($\sigma_a^2 = 1$) and the number of unknown input symbols in the burst is $L = 30$.

4.1. Case of no known symbols

In Fig. 7 (left), the case of no known symbols is shown. We notice that degradations appear at the ends of the burst. The middle symbols appear in N outputs. When no symbols are known, the first and last unknown symbols of the burst appear in strictly less than N outputs, so that there is less information about those symbols in the observations.

The SNR in the middle of the burst converges to the continuous-processing level as the burst length increases.

4.2. Case of $N-1$ known symbols at each end of the burst

We assume now that $n_1 = n_2 = N-1$. The SNR curves are drawn in Fig. 7 (right). This time,

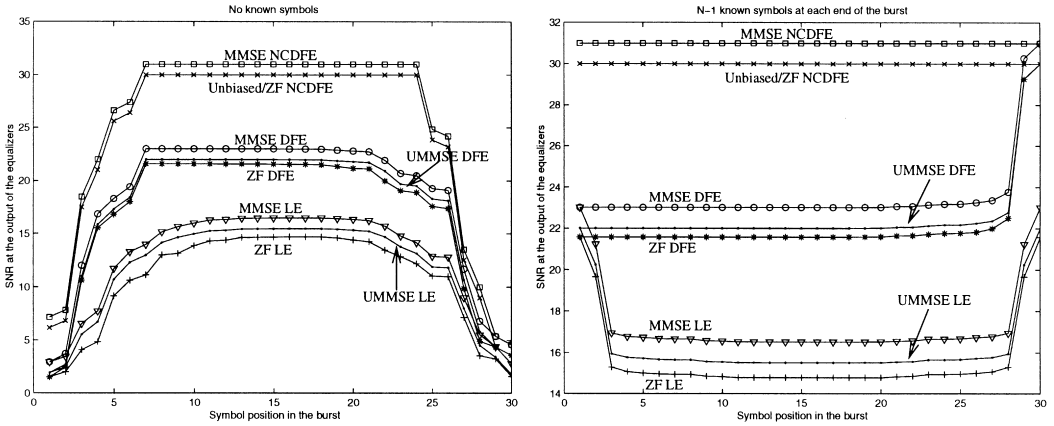

Fig. 7. SNRs at the output of the different equalizers when no symbols are known (left) and when $N-1$ symbols at each end of the burst are known (right).

burst processing performs better than continuous-processing. The middle observations contain N symbols. After eliminating the contributions of the known symbols the outputs at the edges contain strictly less than N symbols, so that there is more information on those symbols. This explains why the symbols are better estimated at both ends for the LEs.

For the DFEs, things are slightly different at the beginning of the burst: the situation is as if the feedback filter had been correctly initialized and the contribution of the past decisions removed, and as the forward filter is anticausal we tend to the continuous-processing case as the number of data tends to infinity. For the last symbol of the burst, the estimation process is the same as that of the NCDFE. We notice that the NCDFE has a constant SNR over the burst equal to the one of continuous-processing.

4.3. Equalizers comparisons

4.3.1. In terms of SNR

The following comparisons are deduced from the amount of a priori information used for estimating the unknown symbols.

- Within each class of equalizers, LE, DFE, NCDFE:

$$\mathrm{SNR}_i(\mathrm{MMSE}) \geqslant \mathrm{SNR}_i(\mathrm{UMMSE}) \geqslant \mathrm{SNR}_i(\mathrm{ZF\text{-}MMSE}). \quad (52)$$

- For each criterion, MMSE, $\mathscr{U}$MMSE and ZF-MMSE,

$$\mathrm{SNR}_i(\mathrm{NCDFE}) \geqslant \mathrm{SNR}_i(\mathrm{DFE}) \geqslant \mathrm{SNR}_i(\mathrm{LE}). \quad (53)$$

4.3.2. In terms of probabilities of error

For unbiased equalizers, a higher SNR implies a lower probability of error: MMSE ZF equalizers will then have a higher probability of error than the corresponding unbiased MMSE equalizers. However, it is not obvious to rank the MMSE equalizers w.r.t. the ZF equalizers because they are biased. In fact, people would tend to believe that a MMSE equalizer performs better than the corresponding MMSE-ZF equalizer.

In the case of constant modulus modulations, MMSE equalizers have the same performance as the corresponding unbiased MMSE equalizers and so a higher performance than MMSE ZF equalizers. For non-constant-modulus constellations, the bias in MMSE equalizers may have a stronger effect than its higher SNR compared to MMSE-ZF equalizers. This is all the more true as the difference in SNRs between the different equalizers tends to be lower as subchannels are added.

Fig. 8 treats of the DFE case. We plot the probabilities of error for the channel $\boldsymbol{H}_2$ (see Appendix E) for the different DFEs. In the error probability computations of the MMSE and $\mathscr{U}$MMSE, the symbols other than the current symbol of interest are approximated as Gaussian random variables. The input symbols belong to a 4-PAM constellation. No symbols are assumed known; the number of known symbols is equal to $L = 34$. In order to see better the difference between the different curves, we only plot the probability of error for the central coefficients. We notice here that the MMSE equalizer has poorer performance than the MMSE-ZF equalizer, and that the $\mathscr{U}$MMSE performs the best.

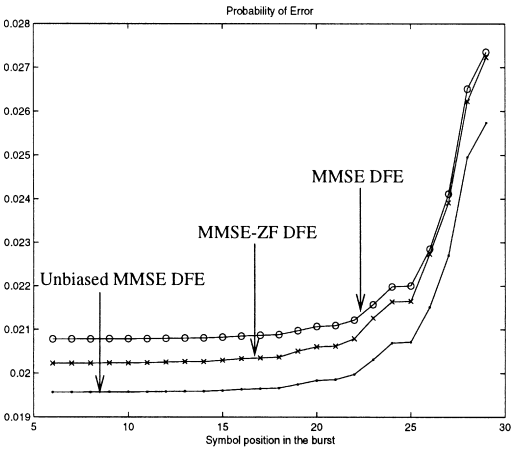

Fig. 8. Probability of error for the MMSE-ZF DFE, the MMSE DFE and the unbiased MMSE DFE.

5. Applying continuous-processing equalizers to the burst case

As already mentioned, burst processing involves time-varying filters. We may wonder if it is worth implementing these time-varying filters, because of complexity reasons, and if simply applying the time-invariant filters corresponding to continuous-processing in burst mode could give acceptable performance.

For that purpose we will consider the case of $N-1$ known symbols at each end of the input burst. We will show that the continuous-processing filters also give better SNR at the ends of the burst than in the middle and always give strictly better SNR than in the continuous-processing case.

For the LEs, the contribution of the known symbols is removed at the end of the observation data. For the DFEs, the initialization is done by putting the $N-1$ leading known symbols in the memory of the feedback filter. Only the trailing known symbols are removed from the processing data.

In both cases, we put the channel outputs before and after the data to be processed equal to zero. The only difference with the continuous-processing case is that we have a finite input symbol sequence, but also a finite noise sequence. As will be seen in the simulations of the next section, for the DFE, the way we proceed is equivalent to the continuous-processing case at the beginning of the burst.

For the LEs, the different reasonings will be held for zero delay non-causal continuous-processing filters. For the DFEs, the forward filter is assumed to be anticausal (zero delay) the feedback filter is causal and FIR (of the same length as the channel). As the channel output is zero outside the time interval of the processing data, these filters will involve only a finite number of data.

In the MMSE ZF case, the MSE contains only the noise contributions. Since the noise is only finite length, the MSE is smaller at the edges. The MSE of MMSE (unbiased or not) equalizers outputs contains residual ISI also. This variance gets also reduced as the input sequence becomes finite length.

For the NCDFE, the leading and trailing symbols are both put in the memory of the feedback filter. In this case, the optimal burst mode feedforward and feedback filters are time invariant and are the same as the continuous mode filters. This fact reinforces the interest of the NCDFE.

5.1. MSE calculations

The outputs of the different linear equalizers based on the continuous-processing filters may be written as:

$$\hat{A}_U = F\boldsymbol{Y}_U \tag{54}$$

where F is a structured matrix containing the coefficients of the continuous processing filter.

In general,

$$\mathrm{MSE}_i = (\sigma_a^2(F\mathscr{T}_U - I)(F\mathscr{T}_U - I)^{\mathrm{H}} + \sigma_v^2 FF^{\mathrm{H}})_{ii}, \tag{55}$$

where $F\mathscr{T}_U = I$ in the ZF case.

The outputs of the different DFEs be may written as

$$\hat{A}_U = F\boldsymbol{Y}_U - (B - I')A'. \tag{56}$$

A' contains A_U and the leading symbols, $I' = [I\ 0]$, F contains the coefficients of the continuous-processing forward filter. B the coefficients of the continuous-processing feedback filter.

In general,

$$\mathrm{MSE}_i = (\sigma_a^2(F\mathscr{T}_U - B)(F\mathscr{T}_U - B)^{\mathrm{H}} + \sigma_v^2 FF^{\mathrm{H}})_{ii}, \tag{57}$$

where $F\mathscr{T}_U = B$ in the ZF case.

In Fig. 9, we present the case of the MMSE LE and MMSE DFE: we compare performances for channel $\boldsymbol{H}_1$. The input symbols are drawn from a BPSK. The length of the filters for the LEs and for the feedforward filter for the DFE is equal to $3N$. The number of unknown input symbols is $L = 30$.

6. Conclusion

We have derived the optimal structure of the burst mode equalizers for three classes of equalizers: linear, decision feedback and non-causal

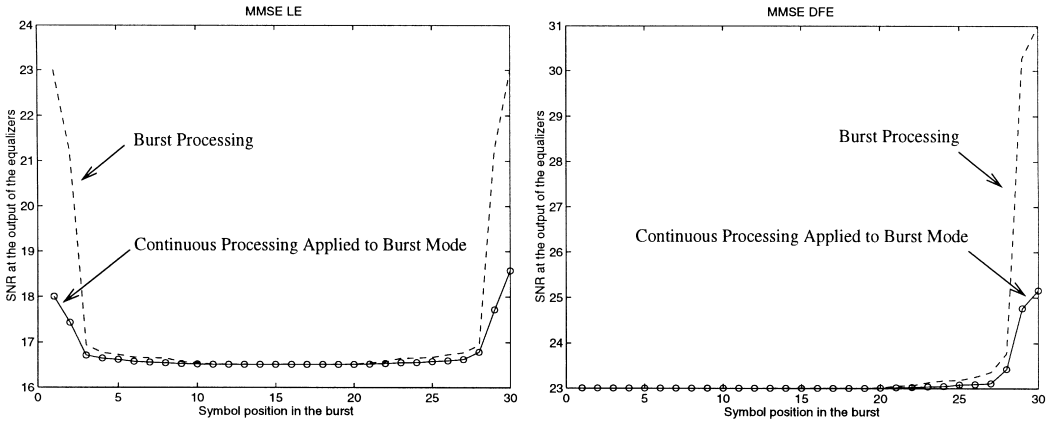

Fig. 9. SNR curves: optimal burst processing compared to continuous processing applied to burst mode for the MMSE LE (left) and MMSE DFE (right).

decision feedback equalizers. Three different criteria have been considered: the MMSE, the unbiased MMSE, and the zero forcing criteria. The problem of finding the equalizer filters have been formulated in terms of linear estimation based on the data and certain a priori information, which allows an easy classification of the equalizers in terms of performance. The SNR degradations have been studied as a function of the position of the unknown symbols in the bursts and as a function of the presence of known symbols. The more favorable situation for burst mode is when pre- and postamble sequence of known symbols are attached at each end of the burst: in this case burst mode equalization performs better than continuous processing. At last we have shown how time-varying burst mode filters can be approximated by time-invariant filters in the situation where the pre- and postamble sequences have the same length as the channel memory: time-invariant filters still have better performance than the continuous-processing level and allows a lower complexity for implementation than the time-varying optimal burst mode filters. The case of the NCDFE appears also of particular interest: it is potentially the most powerful equalizer as it can eliminate all the ISI, and has a particularly simple structure. In particular, when the pre- and postamble sequences have the same length as the channel memory, the NCDFE filters are time-invariant.

Appendix A

$$R_{A_u Y'} R_{Y'Y'}^{-1} \boldsymbol{Y}' = R_{A_u Y'} R_{Y'Y'}^{-1} \left(\begin{bmatrix} \boldsymbol{Y}_u \\ 0 \end{bmatrix} + \begin{bmatrix} \mathscr{T}_K A_K \\ A_K \end{bmatrix} \right), \tag{A.1}$$

$$R_{A_U Y'} = [\sigma_a^2 \ \mathscr{T}_U^{\mathrm{H}} \quad 0] = [R_{A_U Y_U} \quad 0], \tag{A.2}$$

$$R_{Y'Y'} = \begin{bmatrix} \sigma_a^2 \mathscr{T}\mathscr{T}^{\mathrm{H}} + \sigma_v^2 I & \sigma_a^2 \mathscr{T}_K \\ \sigma_a^2 \mathscr{T}_K^{\mathrm{H}} & \sigma_a^2 I \end{bmatrix}. \tag{A.3}$$

The element (1,1) of $R_{Y'Y'}^{-1}$ is

$$R_{Y'Y'}^{-1}(1,1) = (\sigma_a^2 \mathscr{T}_U \mathscr{T}_U^{\mathrm{H}} + \sigma_a^2 \mathscr{T}_K \mathscr{T}_K^{\mathrm{H}} + \sigma_v^2 I - \sigma_a^2 \mathscr{T}_K \mathscr{T}_K^{\mathrm{H}})^{-1} = R_{Y_U Y_U}^{-1}. \tag{A.4}$$

The element (1,2) of $R_{Y'Y'}^{-1}$ is

$$R_{Y'Y'}^{-1}(1,2) = -R_{Y'Y'}^{-1}(1,1)\mathscr{T}_K = -R_{Y_U Y_U}^{-1}\mathscr{T}_K. \tag{A.5}$$

Then

$$R_{A_u Y'} R_{Y'Y'}^{-1} \begin{bmatrix} \boldsymbol{Y}_u \\ 0 \end{bmatrix} = R_{A_U Y_U} R_{Y'Y'}^{-1}(1,1) \boldsymbol{Y}_U = R_{A_U Y_U} R_{Y_U Y_U}^{-1} \boldsymbol{Y}_U \tag{A.6}$$

and

$$R_{A_u Y'} R_{Y'Y'}^{-1} \begin{bmatrix} \mathscr{T}_K A_K \\ A_K \end{bmatrix} = R_{A_u Y'} R_{Y_U Y_U}^{-1} [I \quad -\mathscr{T}_K] \begin{bmatrix} \mathscr{T}_K A_K \\ A_K \end{bmatrix} = 0. \tag{A.7}$$

Then, have the result

$$R_{A_u Y'} R_{Y'}^{-1} Y' = R_{A_u Y_U} R_{Y_U Y_U}^{-1} Y_U. \quad \text{(A.8)}$$

Appendix B

We prove that

$$\mathrm{SNR}_i(\mathscr{U}\mathrm{MMSE}) = \mathrm{SNR}_i(\mathrm{MMSE}) - 1. \quad \text{(B.1)}$$

From expression (20), we deduce

$$(R_{\hat{A}_{U,\mathscr{U}\mathrm{MMSE}} \hat{A}_{U,\mathscr{U}\mathrm{MMSE}}})_{i,i} = \sigma_a^2 \mathscr{D}_{i,i} \quad \text{and}$$

$$(R_{\hat{A}_{U,\mathscr{U}\mathrm{MMSE}} A_U})_{i,i} = \sigma_a^2, \quad \text{(B.2)}$$

$$\begin{aligned} &R_{\tilde{A}_{U,\mathscr{U}\mathrm{MMSE}} \tilde{A}_{U,\mathscr{U}\mathrm{MMSE}}} \\ &\quad = E(A_U - \hat{A}_{U,\mathscr{U}\mathrm{MMSE}})(A_U - \hat{A}_{U,\mathscr{U}\mathrm{MMSE}})^{\mathrm{H}} \\ &\quad = R_{A_U A_U} - R_{A_U \hat{A}_{U,\mathscr{U}\mathrm{MMSE}}} - R_{\hat{A}_{U,\mathscr{U}\mathrm{MMSE}} A_U} \\ &\qquad + R_{\hat{A}_{U,\mathscr{U}\mathrm{MMSE}} \hat{A}_{U,\mathscr{U}\mathrm{MMSE}}}. \end{aligned} \quad \text{(B.3)}$$

and we find

$$\mathrm{MSE}_i(\mathscr{U}\mathrm{MMSE}) = \sigma_a^2(\mathscr{D}_{i,i} - 1). \quad \text{(B.4)}$$

Noting that

$$\sigma_a^2 \mathscr{D}_{i,i}^{-1} = \sigma_a^2 - \mathrm{MSE}_i(\mathrm{MMSE}) \quad \text{(B.5)}$$

$$\mathrm{MSE}_i(\mathscr{U}\mathrm{MMSE}) = \sigma_a^2 \frac{\mathrm{MSE}_i(\mathrm{MMSE})}{\sigma_a^2 - \mathrm{MSE}_i(\mathrm{MMSE})} \quad \text{(B.6)}$$

from which we get (B.1).

Appendix C

We derive the expressions of the feedforward filter F and feedback filter B of the MMSE-DFE. We do not use directly expression (34), but rather a slightly more elegant way.

The expression of the estimate of the unknown symbols by the MMSE-DFE is

$$\hat{A}_{U,\mathrm{MMSE\ DFE}} = F Y_U - B A_U \quad \text{(C.1)}$$

with B strictly triangular inferior. The MMSE criterion is written as

$$\min_{F,B} \|A_U - (F Y_U - B A_U)\|^2 \quad \text{(C.2)}$$

and

$$A_U - \hat{A}_{U,\mathrm{MMSE\ DFE}} = (I + B)A_U - F Y_U. \quad \text{(C.3)}$$

Using the orthogonality principle, which states that the error on A_U should be orthogonal to Y_U, we find

$$(I + B)\sigma_a^2 \mathscr{T}_U^{\mathrm{H}} - F R_{Y_U Y_U} = 0 \quad \text{(C.4)}$$

from which we get

$$\begin{aligned} F &= \sigma_a^2 (I + B) \mathscr{T}_U^{\mathrm{H}} R_{Y_U Y_U}^{-1} \Rightarrow F \\ &= (I + B)\left(\mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U + \frac{\sigma_v^2}{\sigma_a^2} I \right)^{-1} \mathscr{T}_U^{\mathrm{H}} \end{aligned} \quad \text{(C.5)}$$

and

$$A_U - \hat{A}_{U,\mathrm{MMSE\ DFE}} = (I + B)(A_U - \hat{A}_{U,\mathrm{MMSE\ LE}}). \quad \text{(C.6)}$$

Then

$$\begin{aligned} &E\|A_U - A_{U,\mathrm{MMSE\ DFE}}\|^2 \\ &\quad = \sigma_v^2 \mathrm{tr}\{(I + B)R^{-1}(I + B)^{\mathrm{H}}\} \end{aligned} \quad \text{(C.7)}$$

(we recall that $R = \mathscr{T}_U^{\mathrm{H}} \mathscr{T}_U + (\sigma_v^2/\sigma_a^2)I$). Consider the UDL decomposition of R:

$$R = L^{\mathrm{H}} D L, \quad \text{(C.8)}$$

$$\begin{aligned} &E\|A_U - A_{U,\mathrm{MMSE\ DFE}}\|^2 \\ &\quad = \sigma_v^2 \mathrm{tr}\{(I + B)L^{-1}D^{-1}L^{-\mathrm{H}}(I + B)^{\mathrm{H}}\}. \end{aligned} \quad \text{(C.9)}$$

The minimization problem

$$\min_{\{C:\ \mathrm{diag}(C) = I\}} \mathrm{tr}\{C^{\mathrm{H}} D^{-1} C\} \quad \text{(C.10)}$$

as for solution $C = I$.

Then (C.9) is minimized when $I + B = L$. Then

$$\begin{aligned} F &= D^{-1} L^{-\mathrm{H}} \mathscr{T}_U^{\mathrm{H}}, \\ B &= L - I. \end{aligned} \quad \text{(C.11)}$$

Appendix D

We derive the expressions of the feedforward filter F and feedback filter B of the MMSE-ZF

DFE. We want to solve

$$\min_{F\mathscr{T}_U^{F,B}-B=I} E\|A_U-(F\boldsymbol{Y}_U-BA_U)\|^2, \tag{D.1}$$

and with constraint that B be strictly triangular inferior. Let the following decomposition of F onto the rows of $\mathscr{T}_U^{\mathrm{H}}$ and its orthogonal complement $\mathscr{T}_U^{\mathrm{H}\perp}$ (the rows of $\mathscr{T}_U^{\mathrm{H}\perp}$ span the orthogonal complement of the rows of $\mathscr{T}_U^{\mathrm{H}}$):

$$\begin{aligned} &F=F_1\mathscr{T}_U^{\mathrm{H}}+F_2\mathscr{T}_U^{\mathrm{H}\perp},\\ &F_2\mathscr{T}_U^{\mathrm{H}\perp}\mathscr{T}_U^{\mathrm{H}}=0\ \Rightarrow\ F_1\mathscr{T}_U^{\mathrm{H}}\mathscr{T}_U\\ &\quad=I+B\ \Rightarrow\ F_1=(I+B)(\mathscr{T}_U^{\mathrm{H}}\mathscr{T}_U)^{-1} \end{aligned} \tag{D.2}$$

$$(\mathrm{D.1})\Leftrightarrow \min_{F\mathscr{T}_U^{F,B}-B=I} E\|FV\|^2\Leftrightarrow \min_{F\mathscr{T}_U^{F,B}-B=I}\sigma_v^2\|FF^{\mathrm{H}}\|^2 \tag{D.3}$$

$$\|FF^{\mathrm{H}}\|^2=\mathrm{tr}\{F_1\mathscr{T}_U^{\mathrm{H}}\mathscr{T}_UF_1^{\mathrm{H}}\}+\mathrm{tr}\{F_2\mathscr{T}_U^{\mathrm{H}\perp}\mathscr{T}_U^{\mathrm{H}\perp^{\mathrm{H}}}F_2^{\mathrm{H}}\} \tag{D.4}$$

(D.4) gives $F_2=0$:

$$(\mathrm{D.1})\Leftrightarrow \min_{B}(I+B)(\mathscr{T}_U^{\mathrm{H}}\mathscr{T}_U)^{-1}(I+B)^{\mathrm{H}}. \tag{D.5}$$

Considering now the UDL factorization of $\mathscr{T}_U^{\mathrm{H}}\mathscr{T}_U$, we obtain as in Appendix C:

$$\begin{aligned} &F=D^{-1}L^{-\mathrm{H}}\mathscr{T}_U^{\mathrm{H}},\\ &B=L-I. \end{aligned} \tag{D.6}$$

Appendix E

The channels used in the simulations are the following:

$$\boldsymbol{H}_1=\begin{bmatrix} -0.7989 & -0.0562 & 0.7562 & 0.3750 & -2.3775 & 0.3180 & 1.6065\\ -0.7652 & 0.5135 & 0.4005 & 1.1252 & -0.2738 & -0.5112 & 0.8476\\ 0.8617 & 0.3967 & -1.3414 & 0.7286 & -0.3229 & -0.0020 & 0.2681 \end{bmatrix},$$

$$\boldsymbol{H}_2=\begin{bmatrix} -0.6776 & -0.4710 & 0.4992 & 0.1558 & -0.7209\\ 0.4617 & -0.5649 & 0.3827 & -0.5692 & 0.3998\\ -1.1939 & -0.4239 & -0.0136 & -0.7488 & -1.3747 \end{bmatrix}. \tag{E.1}$$

References

[1] N. Al-Dhahir, Time-varying versus time-invariant finite-length MMSE-DFE on stationary dispersive channels, IEEE Trans. Commun. 46 (1) (January 1998) 11–15.

[2] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboglu, G.D. Forney Jr., MMSE decision-feedback equalizers and coding – Part I: equalization results, IEEE Trans. Comm. 43 (10) (October 1995) 2582–2594.

[3] J.M. Cioffi, G.D. Forney Jr., Canonical packet transmission on the ISI channel with Gaussian noise, Proceedings of the GLOBECOM 96 Conference, London, Great Britain, November 1996.

[4] J.M. Cioffi, G.D. Forney Jr., Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise, in: Commun. Comput. Control, Signal Process. A Tribute to Thomas Kailath, Kluwer Academic Publishers, Norwell, MA, 1997, pp. 79–127.

[5] E. de Carvalho, Blind and semi-blind multichannel estimation and equalization for Wireless communications, PhD Thesis, ENST, Paris, 1999.

[6] E. de Carvalho, D.T.M. Slock, Burst mode non-causal decision-feedback equalizer based on soft decisions, Proceedings of the 48th Annual Vehicular Technology Conference, Ottawa, Canada, May 1998.

[7] A. Gersho, T.L. Lim, Adaptive cancellation of intersymbol interference for data transmission, Bell Syst. Tech. J. 60 (11) (November 1981) 1997–2021.

[8] G. Kawas Kaleh, Channel equalization for block transmission systems, IEEE J. Selected Areas Commun. 13 (1) (January 1995) 110–121.

[9] M. Kristensson, B.E. Ottersten, D.T.M. Slock, Blind subspace identification of a BPSK communication channel, Proceedings of the 30th Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, November 1996.

[10] J.G. Proakis, Adaptive non linear filtering techniques for data transmission, IEEE Symp. Adapt. Process. (1970) XV.2.1–5.
[11] D.T.M. Slock, E. de Carvalho, Burst mode non-causal decision-feedback equalization and blind MLSE, Proceedings of the GLOBECOM 96 Conference, London, Great Britain, November 1996.
[12] D.T.M. Slock, E. de Carvalho, Unbiased MMSE decision-feedback equalization for packet transmission, Proceedings of the EUSIPCO 96 Conference, Trieste, Italy, September 1996.
[13] D.T.M. Slock, Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction, Proceedings of the ICASSP 94 Conference, Adelaide, Australia, April 1994.
[14] D.T.M. Slock, C.B. Papadias, Blind fractionally-spaced equalization based on cyclostationarity, Proceedings of the Vehicular Technology Conference, Stockholm, Sweden, June 1994.
[15] L. Tong, G. Xu, T. Kailath, A new approach to blind identification and equalization of multipath channels, Proceedings of the 25th Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, November 1991, pp. 856–860.
[16] L. Tong, G. Xu, T. Kailath, Blind identification and equalization based on second-order statistics: a time domain approach, IEEE Trans. Inf. Theory 40 (2) (March 1994) 340–349.
[17] A.J. van der Veen, Analytical method for blind binary signal separation, IEEE Trans. Signal Process. 45 (4) (April 1997) 1078–1082.