Equalization Techniques for Fading Channels

Geert Leus[1] and Marc Moonen[2]

[1] Delft University of Technology, Delft, The Netherlands

Email: leus@cas.et.tudelft.nl

[2] Katholieke Universiteit Leuven, Leuven, Belgium

Email: moonen@esat.kuleuven.ac.be

CONTENTS

I. INTRODUCTION

Due to the distortive character of the propagation environment, transmitted data symbols will spread out in time and will interfere with each other, a phenomenon called Inter Symbol Interference (ISI). The degree of ISI depends on the data rate: the higher the data rate, the more ISI is introduced. On the other hand, changes in the propagation environment, e.g., due to mobility in wireless communications, introduce channel time-variation, which could be very harmful. Mitigating these fading channel effects, also referred to as *channel equalization*, constitutes a major challenge in current and future communication systems.

In order to design a good channel equalizer, a practical channel model has to be derived. First of all, we can write the overall system as a symbol rate Single-Input Multiple-Output (SIMO) system, where the multiple outputs are obtained by multiple receive antennas and/or fractional sampling. Then, looking at a fixed time-window, we can distinguish between *Time-InVariant (TIV)* and *Time-Varying (TV)* channels. For TIV channels, we will model the channel by a TIV FIR channel, whereas for TV channels, it will be convenient to model the channel time-variation by means of a Basis Expansion Model (BEM), leading to a BEM FIR channel [40], [14], [33].

For TIV channels, channel equalizers have been extensively studied in literature (see for instance [30, ch. 10], [19, ch. 10], [15, ch. 5], [12] and references therein). For TV channels, on the other hand, they have only been introduced recently. Instead of focusing on complex Maximum Likelihood (ML) or Maximum A Posteriori (MAP) equalizers, we will discuss more practical *finite-length* linear and decision feedback equalizers. We derive Minimum Mean-Square Error (MMSE) solutions, which strike an optimal balance between ISI removal and noise enhancement. By setting the signal power to infinity, these MMSE solutions can easily be transformed into Zero-Forcing (ZF) solutions that completely remove the ISI. We mainly focus on equalizer design based on channel knowledge, and briefly mention channel estimation algorithms and direct equalizer design algorithms, which do not require channel knowledge.

In this chapter, we distinguish between block equalizers and serial equalizers (as already mentioned, only practical finite-length versions will be considered). Block equalizers treat both TIV and TV channels in a similar fashion, and will therefore be described in a unified way. *Block Linear Equalizers (BLEs)* [34], [16] as well as *Block Decision Feedback Equalizers (BDFEs)* [37], [16] will be discussed.

What the serial equalizers is concerned, we will focus on both *Serial Linear Equalizers (SLEs)* and *Serial Decision Feedback Equalizers (SDFEs)*. It will turn out to be convenient to use the same model for the serial equalizer as for the channel. Hence, for a TIV FIR channel, we will use a TIV FIR serial equalizer [42], [1], whereas for a BEM FIR channel, we will use a BEM FIR serial equalizer [20], [3], [2]. For an SDFE, this means that both the feedforward and the feedback filter are modeled this way. Note that in the past a TIV FIR serial equalizer has been employed to equalize a BEM FIR channel, but this requires a symbol rate SIMO channel with many outputs for the linear ZF solution to exist [24]. However, when a BEM FIR serial equalizer is used to equalize a BEM FIR channel, only a symbol rate SIMO channel with two outputs is required for the linear ZF solution to exist [20], [3]. We will discuss serial equalization for TIV and TV channels in parallel, in order to show the similarities between the two approaches.

Finally, for TIV channels, it is also possible to adopt *Frequency Domain (FD) equalization*, which can be viewed as a structured block equalization. We will discuss *FD Linear Equalizers (FDLEs)* [32], [6], [10] as well as *FD Decision Feedback Equalizers (FDDFEs)* [4], [10].

Note that throughout this chapter, we will mainly focus our attention on wireless communications. However, most of the proposed techniques can also be adopted for other types of communications, i.e., wireline communications, optical communications, underwater communications, ...

Notation: We use upper (lower) bold face letters to denote matrices (column vectors). Superscripts *, ^{T}, and ^{H} represent complex conjugate, transpose, and Hermitian, respectively. We denote the Kronecker delta by $\delta[n]$ and the Kronecker product by $\otimes$. The convolution operation is represented by $\star$. We denote the $N \times N$ identity matrix as $\mathbf{I}_N$ and the $M \times N$ all-zero matrix as $\mathbf{0}_{M\times N}$. For a column vector $\mathbf{x}$, diag$\{\mathbf{x}\}$ denotes the diagonal matrix with $\mathbf{x}$ on the diagonal, whereas for a square matrix $\mathbf{X}$, diag$\{\mathbf{X}\}$ denotes the diagonal matrix with the diagonal of $\mathbf{X}$ on the diagonal. Next, $[\mathbf{x}]_{i_1:i_2}$ denotes the subvector of $\mathbf{x}$ containing entries i_1 to i_2 (if $i_1 : i_2$ is replaced by i only the ith entry is considered), and $[\mathbf{X}]_{r_1:r_2,c_1:c_2}$ denotes the submatrix of $\mathbf{X}$ on the intersection of rows r_1 to r_2 and columns c_1 to c_2 (if $r_1 : r_2$ ($c_1 : c_2$) is replaced by r (c), only the rth row (cth column) is considered; if r_1 and r_2 (c_1 and c_2) are omitted all rows (columns) are considered). Finally, Q$\{\cdot\}$ represents a decision device that optimally maps soft symbol estimates into hard symbol estimates.

II. Wireless Channel Model

In this section, we discuss the channel model, which is of course a crucial ingredient when deriving means to mitigate fading channel effects. As already mentioned, we will focus our attention on wireless channels, but note that the proposed channel model also holds for many other applications. More specifically, we consider a baseband description of a wireless system with 1 transmit and M receive antennas. For the mth receive antenna, the symbol sequence $x[n]$ is filtered by the transmit filter $g_{\text{tr}}(t)$, distorted by the physical channel $g_{\text{ch}}^{(m)}(t;\tau)$, corrupted by additive noise $v^{(m)}(t)$, and finally filtered by the receive filter $g_{\text{rec}}(t)$. The received signal at the mth receive antenna $y^{(m)}(t)$ can then be written as

$$y^{(m)}(t) = \sum_{n=-\infty}^{\infty} g^{(m)}(t;t-nT)x[n] + w^{(m)}(t),$$

where T is the symbol period, $w^{(m)}(t) := g_{\text{rec}}(t) \star v^{(m)}(t)$ is the additive noise signal at the mth receive antenna, and [13, ch. 1]

$$g^{(m)}(t;\tau) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{\text{rec}}(s)g_{\text{tr}}(\tau-\theta-s)g_{\text{ch}}^{(m)}(t-s;\theta)dsd\theta \tag{1}$$

is the composite channel for the mth receive antenna. Note that the larger the number of receive antennas M, the smaller the probability that at some time instant all M channels are in a deep fade. As a result, the larger the number of receive antennas M, the better the performance. This phenomenon is known as *receive antenna diversity* [29].

Symbol rate sampling, i.e., sampling the M receive antennas at rate $1/T$ is one option, but when the channel bandwidth is larger than $1/(2T)$, the rate $1/T$ is lower than the Nyquist rate. This causes aliasing, which could deteriorate the performance. *Fractional sampling*, i.e., sampling the M receive antennas at rate P/T with $P > 1$, can solve this problem [43], [39]. However, note that since the channel bandwidth is never significantly larger than $1/T$, the performance will not increase much when increasing P beyond $P = 2$.

Focusing on the general case, where the M receive antennas are sampled at rate P/T with $P \geq 1$, each rate-P/T received sequence can be split into P rate-$1/T$ received sequences. The pth rate-$1/T$ received sequence at the mth receive antenna $y^{(mP+p)}[n] := y^{(m)}((nP+p)T/P)$ can be written as

$$y^{(mP+p)}[n] := \sum_{\nu=-\infty}^{\infty} g^{(mP+p)}[n;\nu]x[n-\nu] + w^{(mP+p)}[n], \tag{2}$$

where $w^{(mP+p)}[n] := w^{(m)}((nP+p)T/P)$ and $g^{(mP+p)}[n;\nu] := g^{(m)}((nP+p)T/P;(\nu P+p)T/P)$. Hence, we obtain a symbol rate Single-Input Multiple-Output (SIMO) system with $A = MP$ outputs, which are obtained by multiple receive antennas and/or fractional sampling.

Most wireless links experience multipath propagation, where clusters of reflected or scattered rays arrive at the receiver. All the rays within the same cluster experience the same delay, but each of them is characterized by its own complex gain and frequency offset. Hence, we can express the physical channel $g_{\text{ch}}^{(m)}(t;\tau)$ as [17, ch. 1], [5, ch. 3], [9], [13, ch. 1]

$$g_{\text{ch}}^{(m)}(t;\tau) = \sum_c \delta(\tau - \tau_c^{(m)}) \sum_r G_{c,r}^{(m)} e^{j2\pi f_{c,r}^{(m)} t}, \tag{3}$$

where $\tau_c^{(m)}$ is the delay of the cth cluster related to the mth receive antenna, and $G_{c,r}^{(m)}$ and $f_{c,r}^{(m)}$ are respectively the complex gain and frequency offset of the rth ray of the cth cluster related to the mth receive antenna. Assuming the time-variation of the physical channel $g_{\text{ch}}^{(m)}(t;\tau)$ over the span of the receive filter $g_{\text{rec}}(t)$ is negligible, we can replace $g_{\text{ch}}^{(m)}(t-s;\theta)$ by $g_{\text{ch}}^{(m)}(t;\theta)$ in (1), leading to

$$\begin{aligned} g^{(m)}(t;\tau) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g_{\text{rec}}(s) g_{\text{tr}}(\tau - \theta - s) ds \right) g_{\text{ch}}^{(m)}(t;\theta) d\theta \\ &= \int_{-\infty}^{\infty} \psi(\tau - \theta) g_{\text{ch}}^{(m)}(t;\theta) d\theta \\ &= \sum_c \psi(\tau - \tau_c^{(m)}) \sum_r G_{c,r}^{(m)} e^{j2\pi f_{c,r}^{(m)} t}, \end{aligned}$$

where $\psi(t) := g_{\text{rec}}(t) \star g_{\text{tr}}(t)$. This means that the channel $g^{(mP+p)}[n;\nu]$ can be expressed as

$$\begin{aligned} g^{(mP+p)}[n;\nu] &= g^{(m)}((nP+p)T/P;(\nu P+p)T/P) \\ &= \sum_c \psi((\nu P+p)T/P - \tau_c^{(m)}) \sum_r G_{c,r}^{(m)} e^{j2\pi f_{c,r}^{(m)}(nP+p)T/P}. \end{aligned} \tag{4}$$

The above channel model has a rather complex structure, which complicates, if not prevents, the development of a low-complexity equalization structure that blends well with the channel structure. Moreover, the above channel model contains a large number of parameters, which causes a major problem when trying to estimate the channel. Hence, we will have to look for other channel models that are well-structured and contain a smaller number of parameters. Therefore, we will look at a limited time window $t \in [0, NT)$, which corresponds to $n \in \{0, 1, \ldots, N-1\}$. Depending on the ratio of NT over $1/f_{\max}$, where $f_{\max}$ is the overall *Doppler spread* of all M

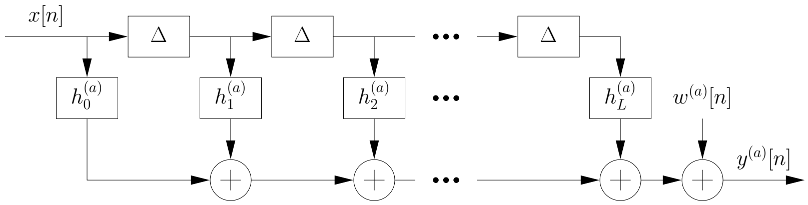

Fig. 1. TIV FIR input-output relation (Δ represents a unit delay).

channels:

$$f_{\max} := \max_{m,c,r}\{|f_{c,r}^{(m)}|\},$$

we call the channel related to the ath output $g^{(a)}[n;\nu]$ Time-Invariant (TIV) or Time-Varying (TV) (note that $a \in \{0, 1, \dots, A-1\}$).

A. TIV Channels

We refer to a channel as a TIV channel, if the channel time-variation over NT is negligible, i.e., if NT is much smaller than $1/f_{\max}$. Assuming that each composite channel satisfies $g^{(m)}(t;\tau) = 0$ for $\tau \notin [0,(L+1)T)$, each TIV channel $g^{(a)}[n;\nu]$ can be modeled for $n \in \{0, 1, \dots, N-1\}$ by a so-called TIV FIR channel:

$$h^{(a)}[n;\nu] = \sum_{l=0}^{L} \delta[\nu - l] h_l^{(a)}. \tag{5}$$

The above TIV FIR channel is well-structured and contains a small number of parameters. Hence, we have obtained a practical channel model.

From (2), the TIV FIR input-output relation for $n \in \{0, 1, \dots, N-1\}$ can be written as (see also Figure 1)

$$y^{(a)}[n] = \sum_{l=0}^{L} h_l^{(a)} x[n-l] + w^{(a)}[n]. \tag{6}$$

B. TV Channels

We refer to a channel as a TV channel, if the channel time-variation over NT is not negligible, i.e., if NT is not much smaller than $1/f_{\max}$. Assuming that each composite channel satisfies

$g^{(m)}(t;\tau) = 0$ for $\tau \notin [0,(L+1)T)$, each TV channel $g^{(a)}[n;\nu]$ can be modeled for $n \in \{0,1,\ldots,N-1\}$ by a so-called TV FIR channel:

$$h^{(a)}[n;\nu] = \sum_{l=0}^{L} \delta[\nu - l] h_l^{(a)}[n]. \tag{7}$$

As the TIV FIR channel, the TV FIR channel is well-structured, but in contrast to the TIV FIR channel, the TV FIR channel contains a large number of parameters, which is objectionable. The key in finding a TV channel model that is well-structured *and* contains a small number of parameters, is to model the channel time-variation using a so-called Basis Expansion Model (BEM) [40], [14], [33].

Assuming each composite channel satisfies $g^{(m)}(t;\tau) = 0$ for $\tau \notin [0,(L+1)T)$, each TV channel $g^{(a)}[n;\nu]$ can be modeled for $n \in \{0,1,\ldots,N-1\}$ by a so-called BEM FIR channel:

$$h^{(a)}[n;\nu] = \sum_{l=0}^{L} \delta[\nu - l] \sum_{q=-Q/2}^{Q/2} h_{q,l}^{(a)} e^{j2\pi qn/K}, \tag{8}$$

where the parameters Q and K should be selected such that $Q/(2KT) \approx f_{\max}$. Note that in general Q can be kept very small as long as NT is smaller than $1/(2f_{\max})$, as illustrated in the next example. Hence, we have again obtained a practical channel model that is well-structured and contains a small number of parameters.

From (2), the BEM FIR input-output relation for $n \in \{0,1,\ldots,N-1\}$ can be written as (see also Figure 2)

$$y^{(a)}[n] = \sum_{l=0}^{L} \sum_{q=-Q/2}^{Q/2} h_{q,l}^{(a)} e^{j2\pi qn/K} x[n-l] + w^{(a)}[n]. \tag{9}$$

Example: Consider a channel $g_{\text{ch}}^{(0)}(t)$ consisting of 5 clusters of 100 reflected or scattered rays. The delay of the cth cluster is given by $\tau_c = cT/2$ ($c \in \{0,1,2,3,4\}$). Assuming that $g_{\text{tr}}(t)$ and $g_{\text{rec}}(t)$ are rectangular functions over $[0,T)$ with height $1/T$, and thus $\psi(t) = g_{\text{rec}}(t) \star g_{\text{tr}}(t)$ is a triangular function over $[0,2T)$ with height 1, we can thus assume that $L = 3$. The complex gain and frequency offset of the rth ray of the cth cluster are given by $G_{c,r}^{(0)} = e^{j\theta_{c,r}^{(0)}}/\sqrt{100}$ and $f_{c,r}^{(0)} = \cos(\phi_{c,r}^{(0)}) f_{\max}$, where $\theta_{c,r}^{(0)}$ and $\phi_{c,r}^{(0)}$ are uniformly distributed over $[0,2\pi)$. Assuming that $f_{\max} = 1/(400T)$, we now show that the BEM FIR channel is very accurate when Q and K are selected such that $Q/(2KT) \approx f_{\max} = 1/(400T)$. To illustrate that Q can be kept very small as long as NT is smaller than $1/(2f_{\max})$, we consider the extreme case of

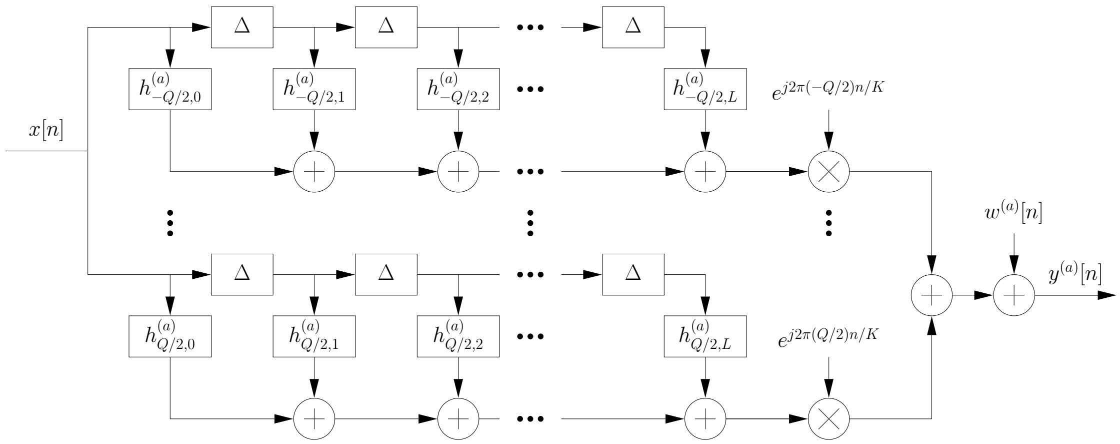

Fig. 2. BEM FIR input-output relation (Δ represents a unit delay).

$Q = 2$ and $NT = 1/(2f_{\max}) = 200T$. To satisfy $Q/(2KT) \approx f_{\max} = 1/(400T)$, we then take $K = 400$. Assuming fractional sampling with a factor $P = 2$, Figure 3 shows the modulus of the 8 TV channel taps $\{\{g^{(a)}[n;l]\}_{a=0}^{1}\}_{l=0}^{3}$ and the modulus of the 8 BEM FIR channel taps $\{\{h^{(a)}[n;l]\}_{a=0}^{1}\}_{l=0}^{3}$ obtained by least squares fitting over $n \in \{0, 1, \ldots, 199\}$. Clearly the approximation for $n \in \{0, 1, \ldots, 199\}$ is very good. □

III. System Model

From now on we will adopt the TIV FIR channel of (5) for TIV channels and the BEM FIR channel of (8) for TV channels. We restrict our attention to outputs $y^{(a)}[n]$ for $n \in \{0, 1, \ldots, N-1\}$.

Defining the $(N+L) \times 1$ data symbol block $\mathbf{x} := [x[-L], \ldots, x[N-1]]^T$, the $N \times 1$ received sample block at the ath output $\mathbf{y}^{(a)} := [y^{(a)}[0], \ldots, y^{(a)}[N-1]]^T$ can be written as

$$\mathbf{y}^{(a)} = \mathbf{H}^{(a)}\mathbf{x} + \mathbf{w}^{(a)}, \tag{10}$$

where $\mathbf{w}^{(a)}$ is similarly defined as $\mathbf{y}^{(a)}$, and $\mathbf{H}^{(a)}$ is an $N \times (N+L)$ channel matrix. The definition of the latter depends on whether we are dealing with TIV or TV channels. In either case, defining $\mathbf{y} := [\mathbf{y}^{(0)T}, \ldots, \mathbf{y}^{(A-1)T}]^T$, we obtain

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w},$$

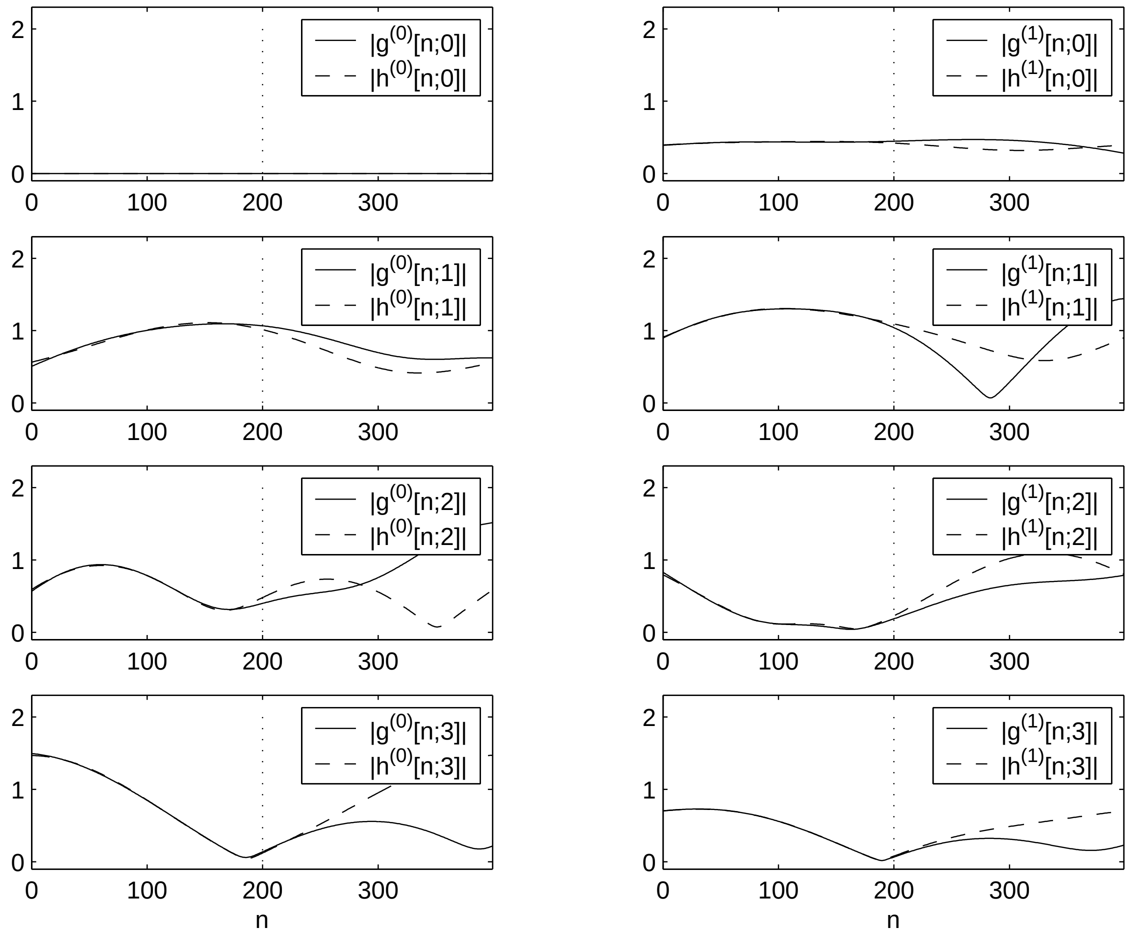

Fig. 3. Illustration of the tight fit of the BEM FIR channel.

where $\mathbf{w}$ is similarly defined as $\mathbf{y}$, and $\mathbf{H} := [\mathbf{H}^{(0)T}, \ldots, \mathbf{H}^{(A-1)T}]^T$. Note that throughout this chapter we will assume perfect knowledge of $\mathbf{H}$. In Section X, we will give a few hints on how to estimate $\mathbf{H}$ in practice.

A. TIV Channels

In case of TIV channels, the $N \times (N + L)$ channel matrix $\mathbf{H}^{(a)}$ is given by

$$\mathbf{H}^{(a)} = \sum_{l=0}^{L} h_l^{(a)} \mathbf{Z}_l, \tag{11}$$

where $\mathbf{Z}_l := [\mathbf{0}_{N\times(L-l)}, \mathbf{I}_N, \mathbf{0}_{N\times l}]$. Substituting (11) in (10), the $N \times 1$ received sample block at the ath output can then be written as

$$\mathbf{y}^{(a)} = \sum_{l=0}^{L} h_l^{(a)} \mathbf{Z}_l \mathbf{x} + \mathbf{w}^{(a)}. \tag{12}$$

B. TV Channels

In case of TV channels, the $N \times (N+L)$ channel matrix $\mathbf{H}^{(a)}$ is given by

$$\mathbf{H}^{(a)} = \sum_{l=0}^{L} \sum_{q=-Q/2}^{Q/2} h_{q,l}^{(a)} \mathbf{D}_q \mathbf{Z}_l, \tag{13}$$

where $\mathbf{D}_q := \text{diag}\{[1, e^{j2\pi q/K}, \ldots, e^{j2\pi q(N-1)/K}]^T\}$ and $\mathbf{Z}_l := [\mathbf{0}_{N\times(L-l)}, \mathbf{I}_N, \mathbf{0}_{N\times l}]$. Substituting (13) in (10), the $N \times 1$ received sample block at the ath output can then be written as

$$\mathbf{y}^{(a)} = \sum_{l=0}^{L} \sum_{q=-Q/2}^{Q/2} h_{q,l}^{(a)} \mathbf{D}_q \mathbf{Z}_l \mathbf{x} + \mathbf{w}^{(a)}. \tag{14}$$

IV. Block Equalization

For block equalization, we will assume that $\mathbf{x} = [\mathbf{0}_{1\times L}, \mathbf{s}^T, \mathbf{0}_{1\times L}]^T$, where $\mathbf{s}$ is an $(N-L)\times 1$ data symbol block. This corresponds to zero padding based block transmission where L zeros are padded after each data symbol block of length $N-L$. The received sample block at the ath output can then be written as

$$\mathbf{y}^{(a)} = \bar{\mathbf{H}}^{(a)} \mathbf{s} + \mathbf{w}^{(a)}, \tag{15}$$

where $\bar{\mathbf{H}}^{(a)} := [\mathbf{H}^{(a)}]_{:,L+1:N}$. We further obtain

$$\mathbf{y} = \bar{\mathbf{H}} \mathbf{s} + \mathbf{w},$$

where $\bar{\mathbf{H}} := [\bar{\mathbf{H}}^{(0)T}, \ldots, \bar{\mathbf{H}}^{(A-1)T}]^T$.

Zero padding can be viewed as a special case of known symbol padding, where the same L known symbols are padded after each data symbol block of length $N-L$. When not all zero, these known symbols can aid synchronization and channel estimation (for TIV channels this has been discussed in [7], [21], [31]). However, for the sake of simplicity, we will stick to zero padding. All results presented for zero padding can easily be modified for the more general known symbol padding case.

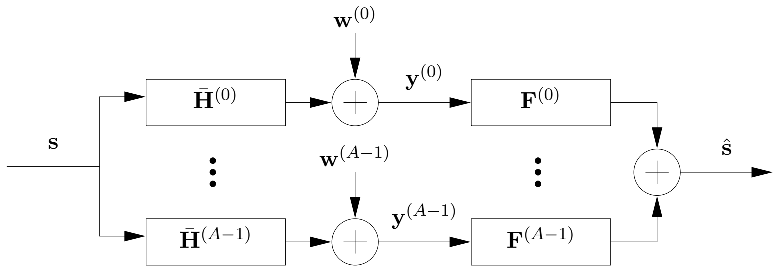

Fig. 4. Block linear equalization.

Zero padding can also be viewed as a special case of linear precoding [34], [35]. Actually, for TIV channels, zero padding turns out to be the best type of linear precoding in terms of performance at high SNR [34], [35]. For TV channels, on the other hand, this is not the case, and other linear precoding strategies with a better performance have been suggested [26] (see also [23] for a multiuser scenario). However, since we do not want to focus on linear precoder design, we will stick to zero padding. All results presented for zero padding can easily be modified for the more general linear precoding case.

We could also adopt cyclic prefix based block transmission [32]. However, as we will show later on, zero padding based block transmission is closely related to cyclic prefix based block transmission. Hence, we will not discuss it in this chapter. A comprehensive overview of different types of block transmission is presented in [46].

A. Block Linear Equalization

In this section, we discuss block linear equalization [34], [16] (see also [25], [18] for a similar approach in the CDMA context). We consider zero padding based block transmission (see (15)). As illustrated in Figure 4, we adopt a Block Linear Equalizer (BLE), consisting of a block filter $\mathbf{F}^{(a)}$ for the ath output, in order to find an estimate of $\mathbf{s}$:

$$\hat{\mathbf{s}} = \sum_{a=0}^{A-1} \mathbf{F}^{(a)}\mathbf{y}^{(a)} = \left(\sum_{a=0}^{A-1} \mathbf{F}^{(a)}\bar{\mathbf{H}}^{(a)}\right)\mathbf{s} + \sum_{a=0}^{A-1} \mathbf{F}^{(a)}\mathbf{w}^{(a)}.$$

Defining $\mathbf{F} := [\mathbf{F}^{(0)}, \ldots, \mathbf{F}^{(A-1)}]$, we then obtain

$$\hat{\mathbf{s}} = \mathbf{F}\mathbf{y} = \mathbf{F}\bar{\mathbf{H}}\mathbf{s} + \mathbf{F}\mathbf{w}.$$

Let us focus on the MMSE BLE, which minimizes the MSE $\mathcal{J} = \mathcal{E}\{\|\mathbf{s} - \hat{\mathbf{s}}\|^2\}$. Defining the data and noise covariance matrices as $\mathbf{R}_s := \mathcal{E}\{\mathbf{s}\mathbf{s}^H\}$ and $\mathbf{R}_w := \mathcal{E}\{\mathbf{w}\mathbf{w}^H\}$, respectively, the MSE can be expressed as

$$\mathcal{J} = \mathrm{tr}\big\{\mathbf{F}(\bar{\mathbf{H}}\mathbf{R}_s\bar{\mathbf{H}}^H + \mathbf{R}_w)\mathbf{F}^H - 2\Re\{\mathbf{R}_s\bar{\mathbf{H}}^H\mathbf{F}^H\} + \mathbf{R}_s\big\}.$$

Solving $\partial\mathcal{J}/\partial\mathbf{F} = \mathbf{0}$, we obtain

$$\begin{aligned}\mathbf{F}_{MMSE} &= \mathbf{R}_s\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\mathbf{R}_s\bar{\mathbf{H}}^H + \mathbf{R}_w)^{-1}\\ &= (\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}} + \mathbf{R}_s^{-1})^{-1}\bar{\mathbf{H}}^H\mathbf{R}_w^{-1},\end{aligned}$$

where the second equality is obtained by using the matrix inversion lemma. Assuming that $\bar{\mathbf{H}}$ has full column rank, the corresponding ZF BLE can be obtained by setting the signal power to infinity ($\mathbf{R}_s^{-1} = \mathbf{0}$):

$$\mathbf{F}_{ZF} = (\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}})^{-1}\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}.$$

Assuming the data sequence and the additive noises are mutually uncorrelated and white with variance σ_s^2 and σ_v^2, respectively, the data and noise covariance matrices can be computed in closed form:

$$\begin{aligned}\mathbf{R}_s &= \sigma_s^2\mathbf{I}_{N-L},\\ \mathbf{R}_w &= \sigma_v^2\mathbf{I}_M \otimes \begin{bmatrix}\mathbf{\Phi}_{N,0} & \cdots & \mathbf{\Phi}_{N,P-1}\\ \vdots & & \vdots\\ \mathbf{\Phi}_{N,-P+1} & \cdots & \mathbf{\Phi}_{N,0}\end{bmatrix},\end{aligned}$$

where $\mathbf{\Phi}_{I,p}$ is the $I \times I$ matrix defined as

$$[\mathbf{\Phi}_{I,p}]_{i,i'} := \int_{-\infty}^{\infty} g_{\mathrm{rec}}(\tau)g_{\mathrm{rec}}(\tau + (i'-i)T + pT/P)d\tau.$$

B. Block Decision Feedback Equalization

In this section, we discuss block decision feedback equalization [37], [16] (see also [8], [18] for a similar approach in the CDMA context). We again consider zero padding based block transmission (see (15)). As illustrated in Figure 5, we adopt a Block Decision Feedback Equalizer

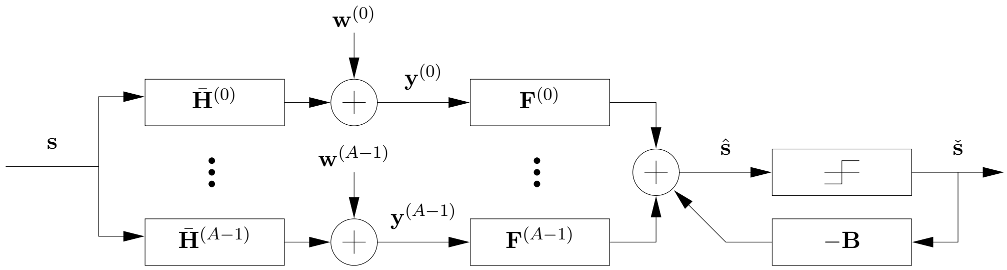

Fig. 5. Block decision feedback equalization.

(BDFE), consisting of a block feedforward filter $\mathbf{F}^{(a)}$ for the ath output and a block feedback filter $\mathbf{B}$, in order to find an estimate of $\mathbf{s}$:

$$\hat{\mathbf{s}} = \sum_{a=0}^{A-1} \mathbf{F}^{(a)}\mathbf{y}^{(a)} - \mathbf{B}\check{\mathbf{s}} = \left(\sum_{a=0}^{A-1} \mathbf{F}^{(a)}\bar{\mathbf{H}}^{(a)}\right)\mathbf{s} - \mathbf{B}\check{\mathbf{s}} + \sum_{a=0}^{A-1} \mathbf{F}^{(a)}\mathbf{w}^{(a)}, \tag{16}$$

where $\check{\mathbf{s}} := \mathrm{Q}\{\hat{\mathbf{s}}\}$. In order to feedback decisions in a causal way, we require $\mathbf{B}$ to be a zero diagonal upper triangular matrix. Defining $\mathbf{F} := [\mathbf{F}^{(0)}, \ldots, \mathbf{F}^{(A-1)}]$, and assuming past decisions are correct (a common assumption in DFE design), i.e., $\check{\mathbf{s}} = \mathbf{s}$, we obtain

$$\hat{\mathbf{s}} = \mathbf{F}\mathbf{y} - \mathbf{B}\mathbf{s} = \mathbf{F}\bar{\mathbf{H}}\mathbf{s} - \mathbf{B}\mathbf{s} + \mathbf{F}\mathbf{w}.$$

Let us focus on the MMSE BDFE, which minimizes the MSE $\mathcal{J} = \mathcal{E}\{\|\mathbf{s} - \hat{\mathbf{s}}\|^2\}$. In a similar fashion as for the BLE, the MSE can be expressed as

$$\mathcal{J} = \mathrm{tr}\big\{\mathbf{F}(\bar{\mathbf{H}}\mathbf{R}_s\bar{\mathbf{H}}^H + \mathbf{R}_w)\mathbf{F}^H - 2\Re\{(\mathbf{B}+\mathbf{I}_{N-L})\mathbf{R}_s\bar{\mathbf{H}}^H\mathbf{F}^H\} + (\mathbf{B}+\mathbf{I}_{N-L})\mathbf{R}_s(\mathbf{B}+\mathbf{I}_{N-L})^H\big\}. \tag{17}$$

Solving $\partial\mathcal{J}/\partial\mathbf{F} = \mathbf{0}$, we obtain

$$\mathbf{F}_{MMSE} = (\mathbf{B} + \mathbf{I}_{N-L})\mathbf{R}_s\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\mathbf{R}_s\bar{\mathbf{H}}^H + \mathbf{R}_w)^{-1} \tag{18}$$

$$= (\mathbf{B} + \mathbf{I}_{N-L})(\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}} + \mathbf{R}_s^{-1})^{-1}\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}, \tag{19}$$

where the second equality is again obtained by using the matrix inversion lemma. Next, substituting (18) in (17) results after some calculation into

$$\mathcal{J} = \mathrm{tr}\big\{(\mathbf{B} + \mathbf{I}_{N-L})\mathbf{R}_{MMSE}(\mathbf{B} + \mathbf{I}_{N-L})^H\big\},$$

where $\mathbf{R}_{MMSE} = (\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}} + \mathbf{R}_s^{-1})^{-1}$. Solving $\partial\mathcal{J}/\partial\mathbf{B} = \mathbf{0}$ under the constraint that $\mathbf{B}$ is a zero diagonal upper triangular matrix, we finally obtain

$$\mathbf{B}_{MMSE} = \mathrm{diag}\{\mathrm{chol}\{\mathbf{R}_{MMSE}^{-1}\}\}^{-1}\mathrm{chol}\{\mathbf{R}_{MMSE}^{-1}\} - \mathbf{I}_{N-L},$$

where chol$\{\mathbf{A}\}$ represents the upper triangular matrix that satisfies the following Cholesky decomposition: $\mathbf{A} = \text{chol}\{\mathbf{A}\}^H\text{chol}\{\mathbf{A}\}$. To summarize, the MMSE BDFE is given by

$$\mathbf{F}_{MMSE} = (\mathbf{B}_{MMSE} + \mathbf{I}_{N-L})(\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}} + \mathbf{R}_s^{-1})^{-1}\bar{\mathbf{H}}^H\mathbf{R}_w^{-1},$$
$$\mathbf{R}_{MMSE} = (\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}} + \mathbf{R}_s^{-1})^{-1},$$
$$\mathbf{B}_{MMSE} = \text{diag}\{\text{chol}\{\mathbf{R}_{MMSE}^{-1}\}\}^{-1}\text{chol}\{\mathbf{R}_{MMSE}^{-1}\} - \mathbf{I}_{N-L}.$$

Assuming $\bar{\mathbf{H}}$ has full column rank, the corresponding ZF BDFE can again be obtained by setting the signal power to infinity ($\mathbf{R}_s^{-1} = \mathbf{0}$):

$$\mathbf{F}_{ZF} = (\mathbf{B}_{ZF} + \mathbf{I}_{N-L})(\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}})^{-1}\bar{\mathbf{H}}^H\mathbf{R}_w^{-1},$$
$$\mathbf{R}_{ZF} = (\bar{\mathbf{H}}^H\mathbf{R}_w^{-1}\bar{\mathbf{H}})^{-1},$$
$$\mathbf{B}_{ZF} = \text{diag}\{\text{chol}\{\mathbf{R}_{ZF}^{-1}\}\}^{-1}\text{chol}\{\mathbf{R}_{ZF}^{-1}\} - \mathbf{I}_{N-L}.$$

V. Serial Linear Equalization

In this section, we discuss serial linear equalization. We do not focus on zero padding based block transmission, but on the serial transmission model (see (12) and (14)). Hence, we assume that all entries of $\mathbf{x}$ contain data symbols. Note, however, that we will not estimate the edges of $\mathbf{x}$ and only estimate the middle part of $\mathbf{x}$ (denoted as $\mathbf{x}_\star$). The edges are either estimated in a previous step (top entries of $\mathbf{x}$) or will be estimated in a next step (bottom entries of $\mathbf{x}$).

We adopt a Serial Linear Equalizer (SLE), consisting of a serial filter $f^{(a)}[n;\nu]$ for the ath output, in order to find an estimate of $x[n-d]$ (see Figure 6):

$$\hat{x}[n-d] = \sum_{a=0}^{A-1}\sum_{\nu=-\infty}^{\infty} f^{(a)}[n;\nu]y^{(a)}[n-\nu], \tag{20}$$

where d represents the synchronization delay. To discuss the structure of this SLE in more detail, we distinguish between TIV and TV channels. Both cases will give rise to a related data model, which allows us to treat the equalizer design in a joint fashion.

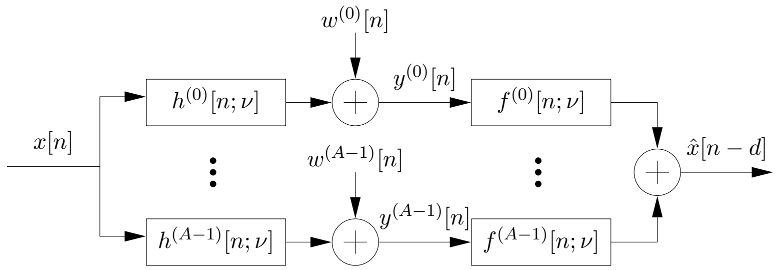

Fig. 6. Serial linear equalization.

A. TIV Channels

Since for a TIV channel, the TIV FIR channel of (5) was applied, it is also convenient to use a TIV FIR serial filter $f^{(a)}[n;\nu]$ [42]. In other words, we design each serial equalizer $f^{(a)}[n;\nu]$ to have $L'+1$ TIV taps:

$$f^{(a)}[n;\nu] = \sum_{l'=0}^{L'} \delta[\nu - l'] f_{l'}^{(a)}.$$

An estimate of $x[n-d]$ is then computed as

$$\hat{x}[n-d] = \sum_{a=0}^{A-1} \sum_{l'=0}^{L'} f_{l'}^{(a)} y^{(a)}[n-l']. \tag{21}$$

Defining the l'th time-shifted received sequence related to the ath output as

$$\mathbf{y}_{l'}^{(a)} := \bar{\mathbf{Z}}_{l'} \mathbf{y}^{(a)},$$

where $\bar{\mathbf{Z}}_{l'} := [\mathbf{0}_{(N-L')\times(L'-l')}, \mathbf{I}_{N-L'}, \mathbf{0}_{(N-L')\times l'}]$, and introducing

$$\mathbf{x}_\star := [x[L'-d], \ldots, x[N-d-1]]^T,$$

this means that an estimate of $\mathbf{x}_\star$ is obtained as

$$\hat{\mathbf{x}}_\star^T = \sum_{a=0}^{A-1} \mathbf{f}^{(a)T} \mathbf{Y}^{(a)},$$

where $\mathbf{f}^{(a)}$ is the $(L'+1)\times 1$ vector given by $\mathbf{f}^{(a)} := [f_{L'}^{(a)}, \ldots, f_0^{(a)}]^T$, and $\mathbf{Y}^{(a)}$ is the $(L'+1)\times(N-L')$ matrix given by $\mathbf{Y}^{(a)} := [\mathbf{y}_{L'}^{(a)}, \ldots, \mathbf{y}_0^{(a)}]^T$.

Let us now rewrite $\mathbf{Y}^{(a)}$ as a function of the TIV FIR channel parameters and the data symbols. The l'th time-shifted received sequence related to the ath output can be written as

$$\begin{aligned}\mathbf{y}_{l'}^{(a)} &:= \bar{\mathbf{Z}}_{l'}\mathbf{y}^{(a)} \\ &= \sum_{l=0}^{L} h_l^{(a)}\bar{\mathbf{Z}}_{l'}\mathbf{Z}_l\mathbf{x} + \mathbf{w}_{l'}^{(a)} \\ &= \sum_{l=0}^{L} h_l^{(a)}\tilde{\mathbf{Z}}_{l+l'}\mathbf{x} + \mathbf{w}_{l'}^{(a)},\end{aligned}$$

where $\mathbf{w}_{l'}^{(a)}$ is similarly defined as $\mathbf{y}_{l'}^{(a)}$ and $\tilde{\mathbf{Z}}_k := [\mathbf{0}_{(N-L')\times(L+L'-k)}, \mathbf{I}_{N-L'}, \mathbf{0}_{(N-L')\times k}]$. Introducing $k := l + l'$, and defining $\mathbf{x}_k := \tilde{\mathbf{Z}}_k\mathbf{x}$ (note that $\mathbf{x}_\star = \mathbf{x}_d$), we can also write this as

$$\mathbf{y}_{l'}^{(a)} = \sum_{k=0}^{L+L'} h_{k-l'}^{(a)}\mathbf{x}_k + \mathbf{w}_{l'}^{(a)}.$$

Defining $\mathbf{X} := [\mathbf{x}_{L+L'}, \ldots, \mathbf{x}_0]^T$, $\mathbf{Y}^{(a)}$ can then be expressed as

$$\mathbf{Y}^{(a)} = \boldsymbol{\mathcal{H}}^{(a)}\mathbf{X} + \mathbf{W}^{(a)},$$

where $\mathbf{W}^{(a)}$ is similarly defined as $\mathbf{Y}^{(a)}$ and $\boldsymbol{\mathcal{H}}^{(a)}$ is the $(L'+1)\times(L+L'+1)$ Toeplitz matrix given by

$$\boldsymbol{\mathcal{H}}^{(a)} := \begin{bmatrix} h_L^{(a)} & \ldots & h_0^{(a)} & & 0 \\ & \ddots & & \ddots & \\ 0 & & h_L^{(a)} & \ldots & h_0^{(a)} \end{bmatrix}.$$

Defining $\mathbf{Y} := [\mathbf{Y}^{(0)T}, \ldots, \mathbf{Y}^{(A-1)T}]^T$, we then obtain

$$\mathbf{Y} = \boldsymbol{\mathcal{H}}\mathbf{X} + \mathbf{W}, \tag{22}$$

where $\mathbf{W}$ is similarly defined as $\mathbf{Y}$ and $\boldsymbol{\mathcal{H}} := [\boldsymbol{\mathcal{H}}^{(0)T}, \ldots, \boldsymbol{\mathcal{H}}^{(A-1)T}]^T$. Hence, we obtain

$$\hat{\mathbf{x}}_\star^T = \sum_{a=0}^{A-1} \mathbf{f}^{(a)T}\mathbf{Y}^{(a)} = \mathbf{f}^T\mathbf{Y} = \mathbf{f}^T\boldsymbol{\mathcal{H}}\mathbf{X} + \mathbf{f}^T\mathbf{W}, \tag{23}$$

where $\mathbf{f} := [\mathbf{f}^{(0)T}, \ldots, \mathbf{f}^{(A-1)T}]^T$.

B. TV Channels

Since for a TV channel, the BEM FIR channel of (8) was applied, it is also convenient to use a BEM FIR serial filter $f^{(a)}[n;\nu]$ [20], [3]. In other words, we design each serial filter $f^{(a)}[n;\nu]$ to have $L'+1$ TV taps, where the time-variation of each tap is modeled by $Q'+1$ complex exponentials:

$$f^{(a)}[n;\nu] = \sum_{l'=0}^{L'} \delta[\nu - l'] \sum_{q'=-Q'/2}^{Q'/2} e^{j2\pi q' n/K} f_{q',l'}^{(a)}.$$

An estimate of $x[n-d]$ is then computed as

$$\hat{x}[n-d] = \sum_{a=0}^{A-1} \sum_{l'=0}^{L'} \sum_{q'=-Q'/2}^{Q'/2} e^{j2\pi q' n/K} f_{q',l'}^{(a)} y^{(a)}[n-l'].$$

Defining the q'th frequency-shifted and l'th time-shifted received sequence related to the ath output as

$$\mathbf{y}_{q',l'}^{(a)} := \bar{\mathbf{D}}_{q'} \bar{\mathbf{Z}}_{l'} \mathbf{y}^{(a)},$$

where $\bar{\mathbf{D}}_{q'} := \text{diag}\{[1, e^{j2\pi q'/K}, \dots, e^{j2\pi q'(N-L'-1)/K}]^T\}$ and $\bar{\mathbf{Z}}_{l'} := [\mathbf{0}_{(N-L')\times(L'-l')}, \mathbf{I}_{N-L'}, \mathbf{0}_{(N-L')\times l'}]$, and introducing

$$\mathbf{x}_\star := [x[L'-d], \dots, x[N-d-1]]^T,$$

this means that an estimate of $\mathbf{x}_\star$ is obtained as

$$\hat{\mathbf{x}}_\star^T = \sum_{a=0}^{A-1} \mathbf{f}^{(a)T} \mathbf{Y}^{(a)},$$

where $\mathbf{f}^{(a)}$ is the $(L'+1)(Q'+1)\times 1$ vector given by $\mathbf{f}^{(a)} := [f_{Q'/2,L'}^{(a)}, \dots, f_{Q'/2,0}^{(a)}, \dots, f_{-Q'/2,0}^{(a)}]^T$, and $\mathbf{Y}^{(a)}$ is the $(L'+1)(Q'+1)\times(N-L')$ matrix given by $\mathbf{Y}^{(a)} := [\mathbf{y}_{Q'/2,L'}^{(a)}, \dots, \mathbf{y}_{Q'/2,0}^{(a)}, \dots, \mathbf{y}_{-Q'/2,0}^{(a)}]^T$.

Let us now rewrite $\mathbf{Y}^{(a)}$ as a function of the BEM FIR channel parameters and the data symbols. Using the property $\bar{\mathbf{Z}}_{l'} \mathbf{D}_q = e^{j2\pi q(L'-l')/K} \bar{\mathbf{D}}_q \bar{\mathbf{Z}}_{l'}$, the q'th frequency-shifted and l'th time-shifted received sequence related to the ath output can be written as

$$\begin{aligned}
\mathbf{y}_{q',l'}^{(a)} &:= \bar{\mathbf{D}}_{q'} \bar{\mathbf{Z}}_{l'} \mathbf{y}^{(a)} \\
&= \sum_{l=0}^{L} \sum_{q=-Q/2}^{Q/2} h_{q,l}^{(a)} e^{j2\pi q(L'-l')/K} \bar{\mathbf{D}}_{q'} \bar{\mathbf{D}}_q \bar{\mathbf{Z}}_{l'} \mathbf{Z}_l \mathbf{x} + \mathbf{w}_{q',l'}^{(a)} \\
&= \sum_{l=0}^{L} \sum_{q=-Q/2}^{Q/2} e^{j2\pi q(L'-l')/K} h_{q,l}^{(a)} \bar{\mathbf{D}}_{q+q'} \tilde{\mathbf{Z}}_{l+l'} \mathbf{x} + \mathbf{w}_{q',l'}^{(a)},
\end{aligned}$$

where $\mathbf{w}_{q',l'}^{(a)}$ is similarly defined as $\mathbf{y}_{q',l'}^{(a)}$ and $\tilde{\mathbf{Z}}_k := [\mathbf{0}_{(N-L')\times(L+L'-k)}, \mathbf{I}_{N-L'}, \mathbf{0}_{(N-L')\times k}]$. Introducing $k := l + l'$ and $p := q + q'$, and defining $\mathbf{x}_{p,k} := \bar{\mathbf{D}}_p \tilde{\mathbf{Z}}_k \mathbf{x}$ (note that $\mathbf{x}_\star = \mathbf{x}_{0,d}$), we can also write this as

$$\mathbf{y}_{q',l'}^{(a)} = \sum_{k=0}^{L+L'} \sum_{p=-(Q+Q')/2}^{(Q+Q')/2} e^{j2\pi(p-q')(L'-l')/K} h_{p-q',k-l'}^{(a)} \mathbf{x}_{p,k} + \mathbf{w}_{q',l'}^{(a)}.$$

Defining $\mathbf{X} := [\mathbf{x}_{Q/2+Q'/2,L+L'}, \ldots, \mathbf{x}_{Q/2+Q'/2,0}, \ldots, \mathbf{x}_{-Q/2-Q'/2,0}]^T$, $\mathbf{Y}^{(a)}$ can then be expressed as

$$\mathbf{Y}^{(a)} = \boldsymbol{\mathcal{H}}^{(a)} \mathbf{X} + \mathbf{W}^{(a)},$$

where $\mathbf{W}^{(a)}$ is similarly defined as $\mathbf{Y}^{(a)}$ and $\boldsymbol{\mathcal{H}}^{(a)}$ is the $(Q'+1)(L'+1)\times(Q+Q'+1)(L+L'+1)$ matrix given by

$$\boldsymbol{\mathcal{H}}^{(a)} := \begin{bmatrix} \boldsymbol{\Omega}^{Q/2}\boldsymbol{\mathcal{H}}_{Q/2}^{(a)} & \ldots & \boldsymbol{\Omega}^{-Q/2}\boldsymbol{\mathcal{H}}_{-Q/2}^{(a)} & & \mathbf{0} \\ & \ddots & & \ddots & \\ \mathbf{0} & & \boldsymbol{\Omega}^{Q/2}\boldsymbol{\mathcal{H}}_{Q/2}^{(a)} & \ldots & \boldsymbol{\Omega}^{-Q/2}\boldsymbol{\mathcal{H}}_{-Q/2}^{(a)} \end{bmatrix},$$

with $\boldsymbol{\mathcal{H}}_q^{(a)}$ the $(L'+1)\times(L+L'+1)$ Toeplitz matrix given by

$$\boldsymbol{\mathcal{H}}_q^{(a)} := \begin{bmatrix} h_{q,L}^{(a)} & \ldots & h_{q,0}^{(a)} & & 0 \\ & \ddots & & \ddots & \\ 0 & & h_{q,L}^{(a)} & \ldots & h_{q,0}^{(a)} \end{bmatrix},$$

and $\boldsymbol{\Omega} := \text{diag}\{[1, e^{j2\pi/K}, \ldots, e^{j2\pi L'/K}]^T\}$. Defining $\mathbf{Y} := [\mathbf{Y}^{(0)T}, \ldots, \mathbf{Y}^{(A-1)T}]^T$, we then obtain

$$\mathbf{Y} = \boldsymbol{\mathcal{H}}\mathbf{X} + \mathbf{W}, \tag{24}$$

where $\mathbf{W}$ is similarly defined as $\mathbf{Y}$ and $\boldsymbol{\mathcal{H}} := [\boldsymbol{\mathcal{H}}^{(0)T}, \ldots, \boldsymbol{\mathcal{H}}^{(A-1)T}]^T$. Hence, we obtain

$$\hat{\mathbf{x}}_\star^T = \sum_{a=0}^{A-1} \mathbf{f}^{(a)T}\mathbf{Y}^{(a)} = \mathbf{f}^T\mathbf{Y} = \mathbf{f}^T\boldsymbol{\mathcal{H}}\mathbf{X} + \mathbf{f}^T\mathbf{W}, \tag{25}$$

where $\mathbf{f} := [\mathbf{f}^{(0)T}, \ldots, \mathbf{f}^{(A-1)T}]^T$.

C. Equalizer Design

Noticing the equivalence between (23) and (25) (although with different matrix/vector definitions), we can now proceed with the SLE design for TIV and TV channels in a joint fashion.

Let us focus on the MMSE SLE, which minimizes the MSE $\mathcal{J} = \mathrm{E}\{\|\mathbf{x}_\star - \hat{\mathbf{x}}_\star\|^2\}$. Defining the data and noise covariance matrices as $\mathbf{R}_X := \mathrm{E}\{\mathbf{X}\mathbf{X}^H\}$ and $\mathbf{R}_W = \mathrm{E}\{\mathbf{W}\mathbf{W}^H\}$, respectively, the MSE can be expressed as

$$\mathcal{J} = \mathbf{f}^T(\boldsymbol{\mathcal{H}}\mathbf{R}_X\boldsymbol{\mathcal{H}}^H + \mathbf{R}_W)\mathbf{f}^* - 2\Re\{\mathbf{e}^T\mathbf{R}_X\boldsymbol{\mathcal{H}}^H\mathbf{f}^*\} + \mathbf{e}^T\mathbf{R}_X\mathbf{e}^*.$$

For TIV channels, $\mathbf{e}$ is the $(L+L'+1)\times 1$ unit vector with a 1 in position $L+L'+1-d$. For TV channels, $\mathbf{e}$ is the $(Q+Q'+1)(L+L'+1)\times 1$ unit vector with a 1 in position $(Q+Q')(L+L'+1)/2+L+L'+1-d$. Solving $\partial\mathcal{J}/\partial\mathbf{f} = \mathbf{0}$, we obtain

$$\begin{aligned}\mathbf{f}_{MMSE}^T &= \mathbf{e}^T\mathbf{R}_X\boldsymbol{\mathcal{H}}^H(\boldsymbol{\mathcal{H}}\mathbf{R}_X\boldsymbol{\mathcal{H}}^H + \mathbf{R}_W)^{-1} \\ &= \mathbf{e}^T(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}} + \mathbf{R}_X^{-1})^{-1}\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1},\end{aligned} \tag{26}$$

where the second equality is obtained by using the matrix inversion lemma. Assuming that $\boldsymbol{\mathcal{H}}$ has full column rank, the corresponding ZF SLE can be obtained by setting the signal power to infinity ($\mathbf{R}_X^{-1} = \mathbf{0}$):

$$\mathbf{f}_{ZF}^T = \mathbf{e}^T(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}})^{-1}\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}. \tag{27}$$

Assuming the data sequence and the additive noises are mutually uncorrelated and white with variance σ_x^2 and σ_v^2, respectively, the data and noise covariance matrices can be computed in closed form. For TIV channels, the data and noise covariance matrices are given by

$$\begin{aligned}\mathbf{R}_X &= \sigma_x^2\mathbf{I}_{L+L'+1}, \\ \mathbf{R}_W &= \sigma_v^2\mathbf{I}_M \otimes \begin{bmatrix} \boldsymbol{\Phi}_{L'+1,0} & \cdots & \boldsymbol{\Phi}_{L'+1,P-1} \\ \vdots & & \vdots \\ \boldsymbol{\Phi}_{L'+1,-P+1} & \cdots & \boldsymbol{\Phi}_{L'+1,0} \end{bmatrix}.\end{aligned}$$

For TV channels, the data and noise covariance matrices are given by

$$\begin{aligned}\mathbf{R}_X &= \sigma_x^2\mathbf{J}_{Q+Q'+1} \otimes \mathbf{I}_{L+L'+1}, \\ \mathbf{R}_W &= \sigma_v^2\mathbf{I}_M \otimes \begin{bmatrix} \mathbf{J}_{Q'+1}\otimes\boldsymbol{\Phi}_{L'+1,0} & \cdots & \mathbf{J}_{Q'+1}\otimes\boldsymbol{\Phi}_{L'+1,P-1} \\ \vdots & & \vdots \\ \mathbf{J}_{Q'+1}\otimes\boldsymbol{\Phi}_{L'+1,-P+1} & \cdots & \mathbf{J}_{Q'+1}\otimes\boldsymbol{\Phi}_{L'+1,0} \end{bmatrix},\end{aligned}$$

where $\mathbf{J}_I$ is the $I\times I$ matrix defined as

$$[\mathbf{J}_I]_{i,i'} = \sum_{n=0}^{N-L'-1} e^{j2\pi(i'-i)n/K}.$$

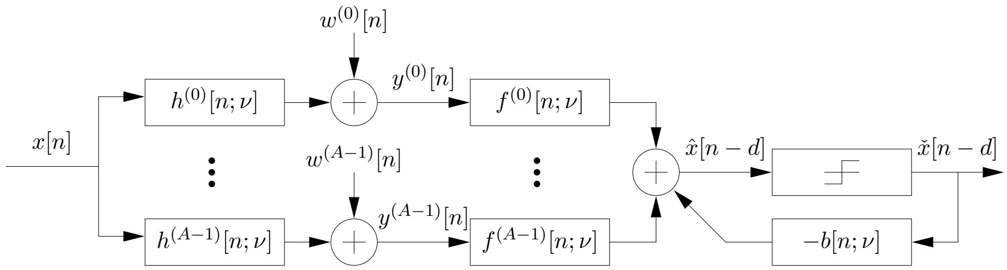

Fig. 7. Serial decision feedback equalization.

VI. Serial Decision Feedback Equalization

In this section, we discuss serial decision feedback equalization. As before, we do not focus on zero padding based block transmission, but on the serial transmission model (see (12) and (14)).

We adopt a Serial Decision Feedback Equalizer (SDFE), consisting of a serial feedforward filter $f^{(a)}[n;\nu]$ for the ath output and a serial feedback filter $b[n;\nu]$, in order to find an estimate of $x[n-d]$ (see Figure 7):

$$\hat{x}[n-d] = \sum_{a=0}^{A-1} \sum_{\nu=-\infty}^{\infty} f^{(a)}[n;\nu] y^{(a)}[n-\nu] - \sum_{\nu=-\infty}^{\infty} b[n;\nu]\check{x}[n-d-\nu],$$

where d again represents the synchronization delay and $\check{x}[n] := \mathrm{Q}\{\hat{x}[n]\}$. To discuss the structure of this SDFE in more detail, we again distinguish between TIV and TV channels. Both cases will again give rise to a related data model, which allows us to treat the equalizer design in a joint fashion.

A. TIV Channels

Since for a TIV channel, the TIV FIR channel of (5) was applied, it is also convenient to use a TIV FIR serial feedforward filter $f^{(a)}[n;\nu]$ and a TIV FIR serial feedback filter $b[n;\nu]$ [1]. In other words, we design each serial feedforward filter $f^{(a)}[n;\nu]$ to have $L'+1$ TIV taps:

$$f^{(a)}[n;\nu] = \sum_{l'=0}^{L'} \delta[\nu - l'] f_{l'}^{(a)}, \tag{28}$$

and the serial feedback filter $b[n;\nu]$ to have $L''+1$ TIV taps:

$$b[n;\nu] = \sum_{l''=0}^{L''} \delta[\nu - l''] b_{l''}, \tag{29}$$

where in order to feedback decisions in a causal way, we require $b_0 = 0$. An estimate of $x[n-d]$ is then computed as

$$\hat{x}[n-d] = \sum_{a=0}^{A-1} \sum_{l'=0}^{L'} f_{l'}^{(a)} y^{(a)}[n-l'] - \sum_{l''=1}^{L''} b_{l''} \check{x}[n-d-l'']. \tag{30}$$

Using the notation introduced in Section V-A, and assuming that past decisions are correct, we can write this as

$$\begin{aligned} \hat{\mathbf{x}}_\star^T = \sum_{a=0}^{A-1} \mathbf{f}^{(a)T} \mathbf{Y}^{(a)} - \mathbf{b}^T \mathbf{P} \mathbf{X} = \mathbf{f}^T \mathbf{Y} - \mathbf{b}^T \mathbf{P} \mathbf{X} \\ = \mathbf{f}^T \boldsymbol{\mathcal{H}} \mathbf{X} - \mathbf{b}^T \mathbf{P} \mathbf{X} + \mathbf{f}^T \mathbf{W}, \end{aligned} \tag{31}$$

where $\mathbf{b}$ is the $(L''+1)\times 1$ vector given by $\mathbf{b} = [b_{L''}, \ldots, b_1, 0]^T$, and $\mathbf{P}$ is the $(L''+1)\times(L+L'+1)$ selection matrix given by

$$\mathbf{P} := [\mathbf{0}_{(L''+1)\times(L+L'-L''-d)}, \mathbf{I}_{L''+1}, \mathbf{0}_{(L''+1)\times d}].$$

B. *TV Channels*

Since for a TV channel, the BEM FIR channel of (8) was applied, it is also convenient to use a BEM FIR serial feedforward filter $f^{(a)}[n;\nu]$ and a BEM FIR serial feedback filter $b[n;\nu]$ [2]. In other words, we design each serial feedforward filter $f^{(a)}[n;\nu]$ to have $L'+1$ TV taps, where the time-variation of each tap is modeled by $Q'+1$ complex exponentials:

$$f^{(a)}[n;\nu] = \sum_{l'=0}^{L'} \delta[\nu - l'] \sum_{q'=-Q'/2}^{Q'/2} e^{j2\pi q' n/K} f_{q',l'}^{(a)}, \tag{32}$$

and the serial feedback filter $b[n;\nu]$ to have $L''+1$ TV taps, where the time-variation of each tap is modeled by $Q''+1$ complex exponentials:

$$b[n;\nu] = \sum_{l''=0}^{L''} \delta[\nu - l''] \sum_{q''=-Q''/2}^{Q''/2} e^{j2\pi q'' n/K} b_{q'',l''}, \tag{33}$$

where in order to feedback decisions in a causal way, we require $b_{-Q''/2,0} = \cdots = b_{Q''/2,0} = 0$. An estimate of $x[n-d]$ is then computed as

$$\hat{x}[n-d] = \sum_{a=0}^{A-1} \sum_{l'=0}^{L'} \sum_{q'=-Q'/2}^{Q'/2} e^{j2\pi q' n/K} f_{q',l'}^{(a)} y^{(a)}[n-l'] - \sum_{l''=1}^{L''} \sum_{q''=-Q''/2}^{Q''/2} e^{j2\pi q'' n/K} b_{q'',l''} \check{x}[n-d-l'']. \tag{34}$$

Using the notation introduced in Section V-B, and assuming that past decisions are correct, we can write this as

$$\hat{\mathbf{x}}_\star^T = \sum_{a=0}^{A-1} \mathbf{f}^{(a)T}\mathbf{Y}^{(a)} - \mathbf{b}^T\mathbf{P}\mathbf{X} = \mathbf{f}^T\mathbf{Y} - \mathbf{b}^T\mathbf{P}\mathbf{X}$$

$$= \mathbf{f}^T\boldsymbol{\mathcal{H}}\mathbf{X} - \mathbf{b}^T\mathbf{P}\mathbf{X} + \mathbf{f}^T\mathbf{W}, \tag{35}$$

where $\mathbf{b}$ is the $((Q''+1)L''+1)\times 1$ vector given by $\mathbf{b} = [b_{Q''/2,L''}, \ldots, b_{Q''/2,1}, \ldots, b_{1,1}, b_{0,L''}, \ldots, b_{0,1}, 0, b_{-1,L''}, \ldots, b_{-1,1}, \ldots, b_{-Q''/2,1}]^T$ and $\mathbf{P}$ is the $((Q''+1)L''+1)\times(Q+Q'+1)(L+L'+1)$ selection matrix given by

$$\mathbf{P} := \begin{bmatrix} & \mathbf{I}_{Q''/2}\otimes\mathbf{P}_1 & & \\ \mathbf{0}_{\alpha\times\beta} & & \mathbf{P}_2 & & \mathbf{0}_{\alpha\times\beta} \\ & & & \mathbf{I}_{Q''/2}\otimes\mathbf{P}_1 & \end{bmatrix},$$

with $\alpha := (Q''+1)L''+1$, $\beta := (Q+Q'-Q'')(L+L'+1)/2$, $\mathbf{P}_1 := [\mathbf{0}_{L''\times(L+L'-L''-d)}, \mathbf{I}_{L''}, \mathbf{0}_{L''\times(d+1)}]$, and $\mathbf{P}_2 := [\mathbf{0}_{(L''+1)\times(L+L'-L''-d)}, \mathbf{I}_{L''+1}, \mathbf{0}_{(L''+1)\times d}]$.

C. *Equalizer Design*

As in Section V-C, noticing the equivalence between (31) and (35), we can proceed with the SDFE design for TIV and TV channels in a joint fashion.

Let us focus on the MMSE SDFE, which minimizes the MSE $\mathcal{J} = \mathrm{E}\{\|\mathbf{x}_\star - \hat{\mathbf{x}}_\star\|^2\}$. In a similar fashion as for the SLE, the MSE can be expressed as

$$\mathcal{J} = \mathbf{f}^T(\boldsymbol{\mathcal{H}}\mathbf{R}_X\boldsymbol{\mathcal{H}}^H + \mathbf{R}_W)\mathbf{f}^* - 2\Re\{(\mathbf{b}+\mathbf{e})^T\mathbf{P}\mathbf{R}_X\boldsymbol{\mathcal{H}}^H\mathbf{f}^*\} + (\mathbf{b}+\mathbf{e})^T\mathbf{P}\mathbf{R}_X\mathbf{P}^H(\mathbf{b}+\mathbf{e})^*. \tag{36}$$

For TIV channels, $\mathbf{e}$ is the $(L''+1)\times 1$ unit vector with a 1 in position $L''+1$. For TV channels, $\mathbf{e}$ is the $((Q''+1)L''+1)\times 1$ unit vector with a 1 in position $Q''L''/2+L''+1$. Solving $\partial\mathcal{J}/\partial\mathbf{f} = \mathbf{0}$, we obtain

$$\mathbf{f}_{MMSE}^T = (\mathbf{b}+\mathbf{e})^T\mathbf{P}\mathbf{R}_X\boldsymbol{\mathcal{H}}^H(\boldsymbol{\mathcal{H}}\mathbf{R}_X\boldsymbol{\mathcal{H}}^H + \mathbf{R}_W)^{-1} \tag{37}$$

$$= (\mathbf{b}+\mathbf{e})^T\mathbf{P}(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}} + \mathbf{R}_X^{-1})^{-1}\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}, \tag{38}$$

where the second equality is again obtained by using the matrix inversion lemma. Next, substituting (37) in (36) results after some calculation into

$$\mathcal{J} = (\mathbf{b}+\mathbf{e})^T\mathbf{R}_{MMSE}(\mathbf{b}+\mathbf{e})^*,$$

where $\mathbf{R}_{MMSE} = \mathbf{P}(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}} + \mathbf{R}_X^{-1})^{-1}\mathbf{P}^H$. Solving $\partial\mathcal{J}/\partial\mathbf{b} = \mathbf{0}$ under the constraint that $\mathbf{e}^T\mathbf{b} = 0$, we finally obtain

$$\mathbf{b}_{MMSE}^T = \frac{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}}{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}\mathbf{e}} - \mathbf{e}^T.$$

To summarize, the MMSE SDFE is given by

$$\begin{aligned}
\mathbf{f}_{MMSE}^T &= (\mathbf{b}_{MMSE} + \mathbf{e})^T\mathbf{P}(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}} + \mathbf{R}_X^{-1})^{-1}\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}, \\
\mathbf{R}_{MMSE} &= \mathbf{P}(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}} + \mathbf{R}_X^{-1})^{-1}\mathbf{P}^H, \\
\mathbf{b}_{MMSE}^T &= \frac{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}}{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}\mathbf{e}} - \mathbf{e}^T.
\end{aligned}$$

Assuming $\boldsymbol{\mathcal{H}}$ has full column rank, the corresponding ZF SDFE can again be obtained by setting the signal power to infinity ($\mathbf{R}_X^{-1} = \mathbf{0}$):

$$\begin{aligned}
\mathbf{f}_{ZF}^T &= (\mathbf{b}_{ZF} + \mathbf{e})^T\mathbf{P}(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}})^{-1}\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}, \\
\mathbf{R}_{ZF} &= \mathbf{P}(\boldsymbol{\mathcal{H}}^H\mathbf{R}_W^{-1}\boldsymbol{\mathcal{H}})^{-1}\mathbf{P}^H, \\
\mathbf{b}_{ZF}^T &= \frac{\mathbf{e}^T\mathbf{R}_{ZF}^{-1}}{\mathbf{e}^T\mathbf{R}_{ZF}^{-1}\mathbf{e}} - \mathbf{e}^T.
\end{aligned}$$

VII. Frequency Domain Equalization for TIV Channels

For TIV channels, a popular method to reduce the implementation complexity of block equalization is based on Frequency Domain (FD) processing. To explain this FD equalization, we resort again to the zero padding based block transmission applied for block equalization.

Rewriting the data model for zero padding (15) as

$$\mathbf{y}^{(a)} = \mathbf{H}_c^{(a)}\mathbf{u} + \mathbf{w}^{(a)}, \tag{39}$$

where $\mathbf{u} := [\mathbf{s}^T, \mathbf{0}_{1\times L}]^T$ and $\mathbf{H}_c^{(a)} := [[\mathbf{H}^{(a)}]_{:,L+1:N}, [\mathbf{H}^{(a)}]_{:,1:L} + [\mathbf{H}^{(a)}]_{:,N+1:N+L}]$, we observe a similarity with the data model for cyclic prefix based block transmission [32], with the exception that the symbols in the cyclic prefix are now zero. Hence, for TIV channels, where $\mathbf{H}_c^{(a)}$ is circulant, we can simplify (39) using Fast Fourier Transform (FFT) operations, as for cyclic prefix based block transmission [46].

Defining the N-point normalized FFT of $\mathbf{u}$ as $\tilde{\mathbf{u}} := \mathbf{G}\mathbf{u}$ and the N-point normalized FFT of $\mathbf{y}^{(a)}$ as $\tilde{\mathbf{y}}^{(a)} := \mathbf{G}\mathbf{y}^{(a)}$, we obtain

$$\begin{aligned}
\tilde{\mathbf{y}}^{(a)} &= \mathbf{G}\mathbf{H}_c^{(a)}\mathbf{G}^H\tilde{\mathbf{u}} + \tilde{\mathbf{w}}^{(a)} \\
&= \tilde{\mathbf{H}}^{(a)}\tilde{\mathbf{u}} + \tilde{\mathbf{w}}^{(a)},
\end{aligned} \tag{40}$$

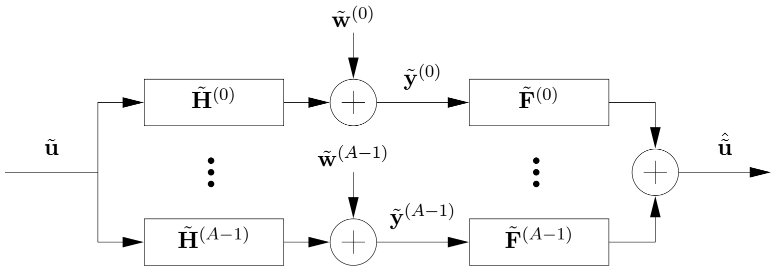

Fig. 8. Frequency domain linear equalization.

where $\tilde{\mathbf{w}}^{(a)}$ is similarly defined as $\tilde{\mathbf{y}}^{(a)}$ and $\tilde{\mathbf{H}}^{(a)} := \text{diag}\{\sqrt{N}\mathbf{G}[h_0^{(a)}, \dots, h_L^{(a)}, \dots, 0]^T\}$. Defining $\tilde{\mathbf{y}} := [\tilde{\mathbf{y}}^{(0)T}, \dots, \tilde{\mathbf{y}}^{(A-1)T}]^T$, we then obtain

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\tilde{\mathbf{u}} + \tilde{\mathbf{w}}, \tag{41}$$

where $\tilde{\mathbf{w}}$ is similarly defined as $\tilde{\mathbf{y}}$, and $\tilde{\mathbf{H}} := [\tilde{\mathbf{H}}^{(0)T}, \dots, \tilde{\mathbf{H}}^{(A-1)T}]^T$. This data model will allow us to use simplified Frequency Domain (FD) processing, as illustrated next. Once an estimate $\hat{\tilde{\mathbf{u}}}$ of $\tilde{\mathbf{u}}$ is obtained, an estimate of $\mathbf{s}$ can be computed as

$$\hat{\mathbf{s}} = [\mathbf{I}_{N-L}, \mathbf{0}_{(N-L)\times L}]\mathbf{G}^H\hat{\tilde{\mathbf{u}}}.$$

Note that $\hat{\mathbf{s}}$ also implies an estimate $\hat{x}[n]$ of $x[n]$.

A. FD Linear Equalization

As illustrated in Figure 8, a FD Linear Equalizer (FDLE) computes an estimate of $\tilde{\mathbf{u}}$ using a FD filter $\tilde{\mathbf{F}}^{(a)}$ for the ath output [32], [6], [10]:

$$\hat{\tilde{\mathbf{u}}} = \sum_{a=0}^{A-1}\tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{y}}^{(a)} = \left(\sum_{a=0}^{A-1}\tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{H}}^{(a)}\right)\tilde{\mathbf{u}} + \sum_{a=0}^{A-1}\tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{w}}^{(a)}.$$

Defining $\tilde{\mathbf{F}} := [\tilde{\mathbf{F}}^{(0)}, \dots, \tilde{\mathbf{F}}^{(A-1)}]$, we then obtain

$$\hat{\tilde{\mathbf{u}}} = \tilde{\mathbf{F}}\tilde{\mathbf{y}} = \tilde{\mathbf{F}}\tilde{\mathbf{H}}\tilde{\mathbf{u}} + \tilde{\mathbf{F}}\tilde{\mathbf{w}}.$$

Let us focus on the MMSE FDLE, which minimizes the MSE $\mathcal{J} = \mathcal{E}\{\|\tilde{\mathbf{u}} - \hat{\tilde{\mathbf{u}}}\|^2\}$. Defining the FD data and noise covariance matrices as $\mathbf{R}_{\tilde{u}} := \mathcal{E}\{\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H\}$ and $\mathbf{R}_{\tilde{w}} := \mathcal{E}\{\tilde{\mathbf{w}}\tilde{\mathbf{w}}^H\}$, respectively, the MSE can be expressed as

$$\mathcal{J} = \text{tr}\big\{\tilde{\mathbf{F}}(\tilde{\mathbf{H}}\text{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H + \text{diag}\{\mathbf{R}_{\tilde{w}}\})\tilde{\mathbf{F}}^H - 2\Re\{\text{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H\tilde{\mathbf{F}}^H\} + \text{diag}\{\mathbf{R}_{\tilde{u}}\}\big\}.$$

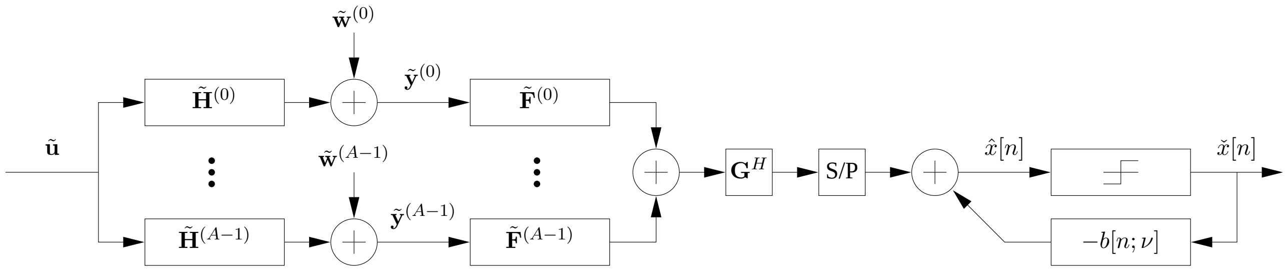

Fig. 9. Frequency domain decision feedback equalization.

Solving $\partial\mathcal{J}/\partial\tilde{\mathbf{F}} = \mathbf{0}$, we obtain

$$\begin{aligned}\tilde{\mathbf{F}}_{MMSE} &= \text{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H(\tilde{\mathbf{H}}\text{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H + \text{diag}\{\mathbf{R}_{\tilde{w}}\})^{-1} \\ &= (\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}} + \text{diag}\{\mathbf{R}_{\tilde{u}}\}^{-1})^{-1}\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1},\end{aligned}$$

where the second equality is obtained by using the matrix inversion lemma. Assuming $\tilde{\mathbf{H}}$ has full column rank, the corresponding ZF FDLE can be obtained by setting the signal power to infinity ($\mathbf{R}_{\tilde{u}}^{-1} = \mathbf{0}$):

$$\tilde{\mathbf{F}}_{ZF} = (\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}.$$

Assuming the data sequence and the additive noises are mutually uncorrelated and white with variance σ_s^2 and σ_v^2, respectively, the FD data and noise covariance matrices can be computed in closed form:

$$\mathbf{R}_{\tilde{u}} = \sigma_s^2\mathbf{G}\begin{bmatrix}\mathbf{I}_{N-L} & \mathbf{0}_{(N-L)\times L} \\ \mathbf{0}_{L\times(N-L)} & \mathbf{0}_{L\times L}\end{bmatrix}\mathbf{G}^H,$$

$$\mathbf{R}_{\tilde{w}} = \sigma_v^2\mathbf{I}_M \otimes \begin{bmatrix}\mathbf{G}\boldsymbol{\Phi}_{N,0}\mathbf{G}^H & \cdots & \mathbf{G}\boldsymbol{\Phi}_{N,P-1}\mathbf{G}^H \\ \vdots & & \vdots \\ \mathbf{G}\boldsymbol{\Phi}_{N,-P+1}\mathbf{G}^H & \cdots & \mathbf{G}\boldsymbol{\Phi}_{N,0}\mathbf{G}^H\end{bmatrix}.$$

B. FD Decision Feedback Equalization

Due to the inherent delay of FD processing, the feedback part of any FD decision feedback equalization approach has to be implemented in the time domain, e.g., by means of a serial filter. Therefore, as illustrated in Figure 9, a FD Decision Feedback Equalizer (FDDFE) computes an estimate of $\tilde{\mathbf{u}}$ using a FD feedforward filter $\tilde{\mathbf{F}}^{(a)}$ for the ath output and a serial feedback filter $b[n;\nu]$ [4], [10]:

$$\begin{aligned}\hat{\tilde{\mathbf{u}}} &= \sum_{a=0}^{A-1} \tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{y}}^{(a)} - \sum_{n=0}^{N-1}[\mathbf{G}]_{:,n+1} \sum_{\nu=-\infty}^{\infty} b[n;\nu]\check{x}[n-\nu] \\ &= \left(\sum_{a=0}^{A-1} \tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{H}}^{(a)}\right)\tilde{\mathbf{u}} - \sum_{n=0}^{N-1}[\mathbf{G}]_{:,n+1} \sum_{\nu=-\infty}^{\infty} b[n;\nu]\check{x}[n-\nu] + \sum_{a=0}^{A-1}\tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{w}}^{(a)},\end{aligned}$$

where $\check{x}[n] = \mathrm{Q}\{\hat{x}[n]\}$. In order to avoid overlap with the previous symbol block and allow for a simple computation, we design the serial feedback filter $b[n;\nu]$ as a TIV FIR filter with $L+1$ taps:

$$b[n;\nu] = \sum_{l=0}^{L} \delta[\nu - l] b_l,$$

where in order to feedback decisions in a causal way, we require $b_0 = 0$. Hence, we can write

$$\hat{\tilde{\mathbf{u}}} = \left(\sum_{a=0}^{A-1} \tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{H}}^{(a)}\right)\tilde{\mathbf{u}} - \sum_{n=0}^{N-1}[\mathbf{G}]_{:,n+1} \sum_{l=1}^{L} b_l\check{x}[n-l] + \sum_{a=0}^{A-1}\tilde{\mathbf{F}}^{(a)}\tilde{\mathbf{w}}^{(a)},$$

Defining $\tilde{\mathbf{F}} := [\tilde{\mathbf{F}}^{(0)}, \ldots, \tilde{\mathbf{F}}^{(A-1)}]$, and assuming past decisions are correct, i.e., $\check{x}[n] = x[n]$, we then obtain

$$\hat{\tilde{\mathbf{u}}} = \tilde{\mathbf{F}}\tilde{\mathbf{y}} - \mathbf{G}\mathbf{B}_c\mathbf{u} = \tilde{\mathbf{F}}\tilde{\mathbf{H}}\tilde{\mathbf{u}} - \mathbf{G}\mathbf{B}_c\mathbf{G}^H\tilde{\mathbf{u}} + \tilde{\mathbf{F}}\tilde{\mathbf{w}},$$

where $\mathbf{B}_c$ is a circulant matrix with first column $[\mathbf{b}^T, \mathbf{0}_{1\times(N-L-1)}]^T$, where $\mathbf{b} := [0, b_1, \ldots, b_L]^T$. Let us focus on the MMSE FDDFE, which minimizes the MSE $\mathcal{J} = \mathcal{E}\{\|\tilde{\mathbf{u}} - \hat{\tilde{\mathbf{u}}}\|^2\}$. In a similar fashion as for the FDLE, the MSE can be expressed as

$$\begin{aligned}\mathcal{J} = \mathrm{tr}\big\{&\tilde{\mathbf{F}}(\tilde{\mathbf{H}}\mathrm{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H + \mathrm{diag}\{\mathbf{R}_{\tilde{w}}\})\tilde{\mathbf{F}}^H - 2\Re\{\mathbf{G}(\mathbf{B}_c + \mathbf{I}_N)\mathbf{G}^H\mathrm{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H\tilde{\mathbf{F}}^H\} \\ &+ \mathbf{G}(\mathbf{B}_c + \mathbf{I}_N)\mathbf{G}^H\mathrm{diag}\{\mathbf{R}_{\tilde{u}}\}\mathbf{G}(\mathbf{B}_c + \mathbf{I}_N)^H\mathbf{G}^H\big\}, \qquad (42)\end{aligned}$$

where $\mathbf{B}_c + \mathbf{I}_N$ is a circulant matrix with first column $[(\mathbf{b} + \mathbf{e})^T, \mathbf{0}_{1\times(N-L-1)}]^T$, where $\mathbf{e}$ is the $(L+1)\times 1$ unit vector with a 1 in the first position. Solving $\partial\mathcal{J}/\partial\tilde{\mathbf{F}} = \mathbf{0}$, we obtain

$$\begin{aligned}\tilde{\mathbf{F}}_{MMSE} &= \mathbf{G}(\mathbf{B}_c + \mathbf{I}_N)\mathbf{G}^H\mathrm{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H(\tilde{\mathbf{H}}\mathrm{diag}\{\mathbf{R}_{\tilde{u}}\}\tilde{\mathbf{H}}^H + \mathrm{diag}\{\mathbf{R}_{\tilde{w}}\})^{-1} \qquad (43) \\ &= \mathbf{G}(\mathbf{B}_c + \mathbf{I}_N)\mathbf{G}^H(\tilde{\mathbf{H}}^H\mathrm{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}} + \mathrm{diag}\{\mathbf{R}_{\tilde{u}}\}^{-1})^{-1}\tilde{\mathbf{H}}^H\mathrm{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}, \qquad (44)\end{aligned}$$

where the second equality is again obtained by using the matrix inversion lemma. Next, substituting (43) in (42) results after some calculation into

$$\begin{aligned}\mathcal{J} &= \text{tr}\big\{\mathbf{G}(\mathbf{B}_c+\mathbf{I}_N)\mathbf{G}^H(\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}}+\text{diag}\{\mathbf{R}_{\tilde{u}}\}^{-1})^{-1}\mathbf{G}(\mathbf{B}_c+\mathbf{I}_N)^H\mathbf{G}^H\big\}\\ &= (\mathbf{b}+\mathbf{e})^T\mathbf{R}_{MMSE}(\mathbf{b}+\mathbf{e})^*,\end{aligned}$$

where $\mathbf{R}_{MMSE} = N[\mathbf{G}^T(\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}}+\text{diag}\{\mathbf{R}_{\tilde{u}}\}^{-1})^{-1}\mathbf{G}^*]_{1:L+1,1:L+1}$. Solving $\partial\mathcal{J}/\partial\mathbf{b} = \mathbf{0}$ under the constraint $\mathbf{e}^T\mathbf{b} = 0$, we finally obtain

$$\mathbf{b}_{MMSE}^T = \frac{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}}{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}\mathbf{e}} - \mathbf{e}^T.$$

To summarize, the MMSE FDDFE is given by

$$\begin{aligned}\tilde{\mathbf{F}}_{MMSE} &= \mathbf{G}(\mathbf{B}_{c,MMSE}+\mathbf{I}_N)\mathbf{G}^H(\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}}+\text{diag}\{\mathbf{R}_{\tilde{u}}\}^{-1})^{-1}\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1},\\ \mathbf{R}_{MMSE} &= N[\mathbf{G}^T(\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}}+\text{diag}\{\mathbf{R}_{\tilde{u}}\}^{-1})^{-1}\mathbf{G}^*]_{1:L+1,1:L+1},\\ \mathbf{b}_{MMSE}^T &= \frac{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}}{\mathbf{e}^T\mathbf{R}_{MMSE}^{-1}\mathbf{e}} - \mathbf{e}^T.\end{aligned}$$

Assuming $\tilde{\mathbf{H}}$ has full column rank, the corresponding ZF FDDFE can again be obtained by setting the signal power to infinity ($\mathbf{R}_{\tilde{u}}^{-1} = \mathbf{0}$):

$$\begin{aligned}\tilde{\mathbf{F}}_{ZF} &= \mathbf{G}(\mathbf{B}_{c,ZF}+\mathbf{I}_N)\mathbf{G}^H(\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1},\\ \mathbf{R}_{ZF} &= N[\mathbf{G}^T(\tilde{\mathbf{H}}^H\text{diag}\{\mathbf{R}_{\tilde{w}}\}^{-1}\tilde{\mathbf{H}})^{-1}\mathbf{G}^*]_{1:L+1,1:L+1},\\ \mathbf{b}_{ZF}^T &= \frac{\mathbf{e}^T\mathbf{R}_{ZF}^{-1}}{\mathbf{e}^T\mathbf{R}_{ZF}^{-1}\mathbf{e}} - \mathbf{e}^T.\end{aligned}$$

VIII. Existence of Zero-Forcing Solution

Comparing the MMSE with the ZF solution, the MMSE solution always leads to a better performance than the ZF solution. However, the *existence* of the ZF solution generally gives a good indication of the performance at high SNR. For instance, when the ZF solution does not exist with probability one, e.g., when it never exists because certain dimensionality conditions are not satisfied, the performance will saturate at high SNR. When the ZF solution exists with probability one, the performance will always increase with increasing SNR. The smaller the region for which one comes close to a channel realization for which the ZF solution does not exist, the steeper the slope of the performance curve (or the higher the collected *diversity*). We will now briefly discuss the existence of the ZF solution for the different equalizers introduced previously.

A. Linear Equalizers

Let us first focus on the LEs. The ZF BLE exists if and only if the channel matrix $\bar{\mathbf{H}}$ has full column rank, which requires that $\bar{\mathbf{H}}$ has at least as many rows as columns. Since $\bar{\mathbf{H}}$ has AN rows and $N-L$ columns, this is always the case. However, this does not mean that $\bar{\mathbf{H}}$ always has full column rank. The latter is only true for TIV channels. On the other hand, judging from (27), one would think that the ZF SLE exists if and only if the channel matrix $\boldsymbol{\mathcal{H}}$ has full column rank. However, this is only true if $\boldsymbol{\mathcal{H}}$ is column reduced (see [36] for TIV channels and [20], [2] for TV channels). Assuming this is the case, the existence of the ZF SLE is equivalent with $\boldsymbol{\mathcal{H}}$ having full column rank, which requires that $\boldsymbol{\mathcal{H}}$ has at least as many rows as columns. For TIV channels, this happens when $A(L'+1) \geq L+L'+1$, whereas for TV channels, this happens when $A(Q'+1)(L'+1) \geq (Q+Q'+1)(L+L'+1)$. Clearly, these inequalities can only be satisfied if $A \geq 2$ with a sufficiently large L' for TIV channels and a sufficiently large Q' and L' for TV channels. Hence, we need at least two outputs, which can for instance be achieved by sampling $M=2$ receive antennas at rate $1/T$ ($P=1$) or sampling $M=1$ receive antenna at rate $2/T$ ($P=2$). More detailed sufficient conditions for the ZF SLE to exist can be found in [36] for TIV channels and [20], [2] for TV channels. As already discussed, for TIV channels, we can also adopt FD processing. The ZF FDLE exists if and only if $\tilde{\mathbf{H}}$ has full column rank, which requires that $\tilde{\mathbf{H}}$ has at least as many rows as columns. Since $\tilde{\mathbf{H}}$ has AN rows and N columns, this is again always the case. However, in contrast to $\bar{\mathbf{H}}$, $\tilde{\mathbf{H}}$ does not always have full column rank for TIV channels. It becomes singular when all A channels have a common zero on the N-point FFT grid, i.e., when $\tilde{\mathbf{H}}$ has a zero column.

B. Decision Feedback Equalizers

As far as DFEs are concerned, note that the MMSE and ZF DFEs that we have proposed earlier assume that past decisions are correct, which basically makes the DFEs look linear. Only in this context, the statements we made at the beginning of this section hold. Judging from the equations we presented for the different ZF DFEs, we would tend to think that a ZF DFE exists if and only if the corresponding ZF LE exists. However, other (more complicated) equations could be derived from which we could see that a ZF DFE can also exist when the corresponding ZF LE does not exist. Suffice it to illustrate this for the SDFE. Defining $\mathbf{P}^{\perp}$ as the orthogonal complement of $\mathbf{P}$, i.e., $\mathbf{P}^{\perp T}\mathbf{P} = \mathbf{0}$, it is clear from (31) and (35) that ISI is completely removed and thus a ZF SDFE is obtained if $\mathbf{f}^T\boldsymbol{\mathcal{H}}\mathbf{P}^{\perp T} = \mathbf{0}$, $\mathbf{f}^T\boldsymbol{\mathcal{H}}\mathbf{e} = 1$, and $(\mathbf{b}+\mathbf{e})^T = \mathbf{f}^T\boldsymbol{\mathcal{H}}\mathbf{P}$. A sufficient condition for this to be satisfied is that $[\boldsymbol{\mathcal{H}}\mathbf{P}^{\perp T}, \boldsymbol{\mathcal{H}}\mathbf{e}]$ has full column rank, which requires that $[\boldsymbol{\mathcal{H}}\mathbf{P}^{\perp T}, \boldsymbol{\mathcal{H}}\mathbf{e}]$ has at least as many rows as columns. For TIV channels, this happens when $A(L'+1) \geq (L+L'+1) - L''$, whereas for TV channels, this happens when $A(Q'+1)(L'+1) \geq (Q+Q'+1)(L+L'+1) - (Q''+1)L''$. Again, these inequalities can be satisfied if $A \geq 2$, but this time they can also be satisfied if $A = 1$, i.e., when the ZF SLE does not exist.

IX. COMPLEXITY

In this section, we discuss some complexity issues of the above equalization structures. We can distinguish between design complexity and implementation complexity. The design complexity is the computational cost to design the equalizer, whereas the implementation complexity is the computational cost to equalize the channel once the equalizer has been designed. The block size N will play an important role in these complexities. In the following, we always assume that N is chosen large enough such that blind channel estimation becomes feasible or the overhead of the training symbols for training based channel estimation does not decrease the data transmission rate too much (this basically boils down to choosing $N \gg L+1$ for TIV channels and $N \gg (Q+1)(L+1)$ for TV channels).

A. Design Complexity

Although many equalizer design procedures are possible (see Section X), we will consider equalizer design based on channel knowledge. For the sake of simplicity, we will not exploit the band structure of $\bar{\mathbf{H}}^{(a)}$ or the special structure of $\boldsymbol{\mathcal{H}}^{(a)}$ in the design complexity calculations. Let us first take a look at the linear equalization approaches. To design a BLE, we have to compute the inverse of an $(N-L) \times (N-L)$ matrix, which requires $\mathcal{O}((N-L)^3)$ flops. On the other hand, to compute an SLE, we need the inverse of a $D \times D$ matrix, where $D = L+L'+1$ for TIV channels and $D = (Q+Q'+1)(L+L'+1)$ for TV channels, which requires $\mathcal{O}(D^3)$ flops. As will be illustrated in Section XI, with a D that is much smaller than $N-L$, the performance of the SLE can approach the performance of the BLE for $A > 1$. As a result, the SLE can have a much smaller design complexity than the BLE, without a significant loss in performance for $A > 1$. For $A = 1$, there is a loss in performance at high SNR, since in contrast to the performance of the BLE, the performance of the SLE saturates at high SNR.

Let us now focus on decision feedback equalization approaches. To design the feedback part of a BDFE, we have to compute the Cholesky decomposition of an $(N-L) \times (N-L)$ matrix. Hence, next to the $\mathcal{O}((N-L)^3)$ flops to design the feedforward part (similar to the complexity to design a BLE), we have an extra cost of $\mathcal{O}((N-L)^3)$ flops to design the feedback part. On the other hand, to compute the feedback part of an SDFE, we need the inverse of a $D'' \times D''$ matrix, where $D'' = L''+1$ for TIV channels and $D'' = (Q''+1)L''+1$ for TV channels. Hence, next to the $\mathcal{O}(D^3)$ flops to design the feedforward part (similar to the complexity to design a SLE), we have an extra cost of $\mathcal{O}(D''^3)$ flops to design the feedback part. As will be illustrated in Section XI, with a D that is much smaller than $N-L$ and a D'' that is about half the size of D (and thus also much smaller than $N-L$), the performance of the SDFE can approach the performance of the BDFE. As a result, the SDFE can have a much smaller design complexity than the BDFE, without a significant loss in performance. This now even holds for $A = 1$, since as the performance of the BDFE, the performance of the SDFE does not saturate at high SNR.

Adopting an FDLE (FDDFE) for TIV channels, the FFT processing is computationally the most expensive, and results in a complexity of $\mathcal{O}(N \log_2 N)$ flops. Hence, the FDLE (FDDFE) for TIV channels has a much smaller design complexity than the BLE (BDFE), while their performances are comparable, as will be illustrated in Section XI. The comparison with the design complexity of the SLE (SDFE) for TIV channels depends on the specific scenario.

B. Implementation complexity

The implementation complexity will be defined here as the number of multiply-add (MA) operations required to estimate the transmitted data symbols. For the BLE, estimating the transmitted data symbols requires $N(N-L)$ MA operations per output, with an extra $(N-L)(N-L-1)/2$ MA operations for the BDFE. On the other hand, for the SLE, estimating the transmitted data symbols requires $(N-L')D'$ MA operations per output, with an extra $(N-L')(D''-1)$ MA operations for the SDFE, where $D' = L'+1$ for TIV channels and $D' = (Q'+1)(L'+1)$ for TV channels, and D'' is defined as before. Previously, we mentioned that with a D (and thus also a D') that is much smaller than $N-L$, the performance of the SLE can approach the performance of the BLE for $A > 1$. Hence, the SLE can also have a much smaller implementation complexity than the BLE, without a significant loss in performance for $A > 1$. Above, we also mentioned that with a D (and thus also a D') that is much smaller than $N - L$ and a D'' that is about half the size of D (and thus much smaller than $(N - L - 1)/2$), the performance of the SDFE can approach the performance of the BDFE. Hence, the SDFE can also have a much smaller implementation complexity than the BDFE, without a significant loss in performance.

Like the design complexity, the implementation complexity of the FDLE (FDDFE) for TIV channels is completely determined by the complexity of the FFT processing, which is again much smaller than the implementation complexity of the BLE (BDFE). As before, the comparison with the implementation complexity of the SLE (SDFE) for TIV channels depends on the specific scenario. However, we should keep in mind that FD processing is only useful for TIV channels.

X. Channel Estimation and Direct Equalizer Design

Up till now we have focused on equalizer design based on channel knowledge. In practice, however, the channel has to be estimated. There are basically two ways to estimate the channel: training based or blind (intermediate so-called semi-blind approaches are also possible). When training based channel estimation is adopted, training symbols are inserted in $\mathbf{x}$ for serial transmission or in $\mathbf{s}$ for zero padding based block transmission. We refer the interested reader to [45] for TIV channels and [27] for TV channels. When blind channel estimation is adopted, no training symbols are inserted. For serial transmission over TIV channels, many blind channel estimation algorithms have been proposed based on the data model (22) [28], [42], [12]. For serial transmission over TV channels, most of these algorithms can be extended by observing the similarity between the data models (22) and (24) [22]. Instead of only working with time-shifted versions of the received sequences, one then has to make use of time- and frequency-shifted versions of the received sequences. For zero padding based block transmission over TIV channels, an interesting blind channel estimation algorithm has been developed in [35]. For zero padding based block transmission over TV channels, this algorithm can be extended by employing a special type of linear precoding as described in [38].

Next to equalizer design based on channel knowledge, there also exist direct equalizer design algorithms, which do not require channel knowledge. They have mainly been developed for SLEs and can easily be extended to SDFEs. Again, one can distinguish between training based and blind approaches. Training based approaches are fairly easy to develop. Looking at (23) and (25), training based direct equalizer design algorithms basically try to estimate $\mathbf{f}$ based on knowledge of $\mathbf{Y}$ and partial knowledge of $\mathbf{x}_\star$ via least squares fitting for instance. Blind approaches are more difficult to derive. For TIV channels, many blind direct equalizer design algorithms have been proposed based on the data model (22) [42], [12], [11], [44], [41]. For TV channels, most of these algorithms can again be extended by observing the similarity between the data models (22) and (24). As before, instead of only working with time-shifted versions of the received sequences, one then has to make use of time- and frequency-shifted versions of the received sequences. Such direct equalizer design algorithms for TV channels are currently under investigation.

XI. Performance Results

In this section, we compare the performances of the different equalizers discussed in this chapter. We only focus on the MMSE equalizers, which have a better performance than the ZF equalizers. We generate M ($M = 1, 2$) channels $g_{\text{ch}}^{(m)}(t)$ consisting of 5 clusters of 100 reflected or scattered rays. The delay of the cth cluster is given by $\tau_c = cT/2$ ($c \in \{0, 1, 2, 3, 4\}$). Assuming that $g_{\text{tr}}(t)$ and $g_{\text{rec}}(t)$ are rectangular functions over $[0, T)$ with height $1/T$, and thus $\psi(t) = g_{\text{rec}}(t) \star g_{\text{tr}}(t)$ is a triangular function over $[0, 2T)$ with height 1, we can thus assume that $L = 3$. The complex gain and frequency offset of the rth ray of the cth cluster are given by $G_{c,r}^{(m)} = e^{j\theta_{c,r}^{(m)}}/\sqrt{100}$ and $f_{c,r}^{(m)} = \cos(\phi_{c,r}^{(m)}) f_{\max}$, where $\theta_{c,r}^{(m)}$ and $\phi_{c,r}^{(m)}$ are uniformly distributed over $[0, 2\pi)$. We further consider fractional sampling with a factor of $P = 1, 2$. The modulation we use is QPSK with unit modulus. We assume the data sequence and the additive noises are mutually uncorrelated and white. The SNR is defined as $SNR = 5/\sigma_v^2$, where σ_v^2 is the variance of the additive noise. The factor 5 is due to the fact that we consider 5 clusters. For TIV channels, we fit the TIV FIR channel of (5) to the true channel for $n \in \{0, 1, \ldots, N-1\}$, whereas for TV channels, we fit the BEM FIR channel of (8) to the true channel for $n \in \{0, 1, \ldots, N-1\}$. In both cases, we use the obtained channel model parameters to design our equalizer. In practice, we have to estimate the channel model parameters of (5) or (8) using some channel estimation method. This can be a training based method or a blind method (see Section X). Although we make abstraction of this channel estimation procedure in this chapter, it will determine the block size N that we adopt in the simulations.

A. TIV Channels

For the TIV channels case, we consider $f_{\max} = 0$, and use the TIV FIR channel of (5) with $L = 3$ to design our equalizers. Since $f_{\max} = 0$, the TIV FIR assumption will hold for any block size N. We consider a block size $N = 64$. This block size is large enough such that blind channel estimation becomes feasible or the overhead of the training symbols for training based channel estimation does not decrease the data transmission rate too much. Moreover, it is also the block size that has been adopted for the IEEE 802.11a and HIPERLAN/2 WLAN standards (in the context of OFDM). Let us first compare the block equalizers with the frequency domain equalizers for TIV channels. Figure 10 shows the performance of the BLE, BDFE, FDLE, and FDDFE in TIV channels for $M = 1, 2$ and $P = 1, 2$. We observe that the performance of the FDLE (FDDFE) approaches the performance of the BLE (BDFE) for all cases. However, we see that when fractional sampling is employed ($P = 2$), the FDLE (FDDFE) is not capable to improve the performance as much as the BLE (BDFE) does, but the difference between the two approaches is still rather small. Let us next compare the block equalizers with the serial equalizers for TIV channels. For the serial equalizers, we take $L' = 7$, $d = (L + L')/2 = 5$, and $L'' = L + L' - d = 5$. Figure 11 shows the performance of the BLE, BDFE, SLE, and SDFE in TIV channels for $M = 1, 2$ and $P = 1, 2$. We observe that the performance of the SLE approaches the performance of the BLE, except for the case $M = P = 1$ at high SNR, and the performance of the SDFE approaches the performance of the BDFE for all cases.

As mentioned before, the design complexity of the BLE is $\mathcal{O}\{(N - L)^3\}$ flops, with an extra $\mathcal{O}\{(N - L)^3\}$ flops for the BDFE, where $(N - L)^3 \approx 227,000$. On the other hand, the design complexity of the SLE is $\mathcal{O}\{D^3\}$ flops, with an extra $\mathcal{O}\{D''^3\}$ flops for the SDFE, where $D^3 = (L + L' + 1)^3 \approx 1,300$ and $D''^3 = (L'' + 1)^3 \approx 200$. Hence, the design complexity of the SLE (SDFE) is clearly much smaller than the design complexity of the BLE (BDFE). A similar observation holds for the implementation complexity. The BLE requires $N(N-L) = 3,904$ MA operations per output, with an extra $(N-L)(N-L-1)/2 = 1,830$ MA operations for the BDFE, whereas the SLE requires $(N - L')D' = (N - L')(L' + 1) = 456$ MA operations per output, with an extra $(N - L')(D'' - 1) = (N - L')L'' = 285$ MA operations for the SDFE. The major computational cost of designing or implementing the FDLE (FDDFE) is the FFT processing, which results into $\mathcal{O}\{N \log_2 N\}$ flops, where $N \log_2 N = 384$. Hence, compared to the BLE

(BDFE), the design and implementation complexity of the FDLE (FDDFE) are much smaller. Compared to the SLE (SDFE), they are comparable for the chosen block size N. However, they would be larger for a larger block size N.

B. TV Channels

For the TV channels case, we consider $f_{\max} = 1/(400T)$, and use the BEM FIR channel of (8) with $L = 3$ to design our equalizers. We know that in the best case scenario we can take $Q = 2$, and then the number of channel model parameters that would have to be estimated in practice is three times as large as in the TIV channels case. Hence, it would then make sense to take N about three times as large as in the TIV channels case. Let us for instance take $N = 200$. In that case $NT = 200T \leq 1/(2f_{\max}) = 200T$, which means that Q can be kept small to obtain a good BEM FIR approximation, as mentioned in Section II-B. Therefore, we can indeed take $Q = 2$ as assumed above. To satisfy $Q/(2KT) \approx f_{\max} = 1/(400T)$, we then take $K = 400$. As illustrated in example 1, these parameters lead to a tight fit of the BEM FIR channel to the true channel. Since frequency domain equalizers are not useful for TV channels, we only compare the block equalizers with the serial equalizers for TV channels. For the serial equalizers, we take $L' = 7$, $d = (L+L')/2 = 5$, $L'' = L+L'-d = 5$, $Q' = 6$, and $Q'' = (Q+Q')/2 = 4$. Figure 12 shows the performance of the BLE, BDFE, SLE, and SDFE in TV channels for $M = 1, 2$ and $P = 1, 2$. As for the TIV channels, we observe that the performance of the SLE approaches the performance of the BLE, except for the case $M = P = 1$ at high SNR, and the performance of the SDFE approaches the performance of the BDFE for all cases.

Let us again take a look at the design and implementation complexity of the different methods. For the BLE and the BDFE, N has changed compared to the TIV channels case. In other words, the design complexity now depends on $(N - L)^3 \approx 7,645,400$. For the SLE and SDFE, D and D'' have changed compared to the TIV channels case. More specifically, we now have $D^3 = ((L + L' + 1)(Q + Q' + 1))^3 \approx 970,300$ and $D''^3 = ((Q'' + 1)L'' + 1)^3 \approx 17,600$. As for the TIV channels case, we observe that the design complexity of the SLE (SDFE) is much smaller than the design complexity of the BLE (BDFE). A similar observation holds for the implementation complexity. The BLE requires $N(N - L) = 39,400$ MA operations per output, with an extra $(N - L)(N - L - 1)/2 = 19,306$ MA operations for the BDFE, whereas the SLE requires $(N - L')D' = (N - L')(Q' + 1)(L' + 1) = 10,808$ MA operations per output, with an

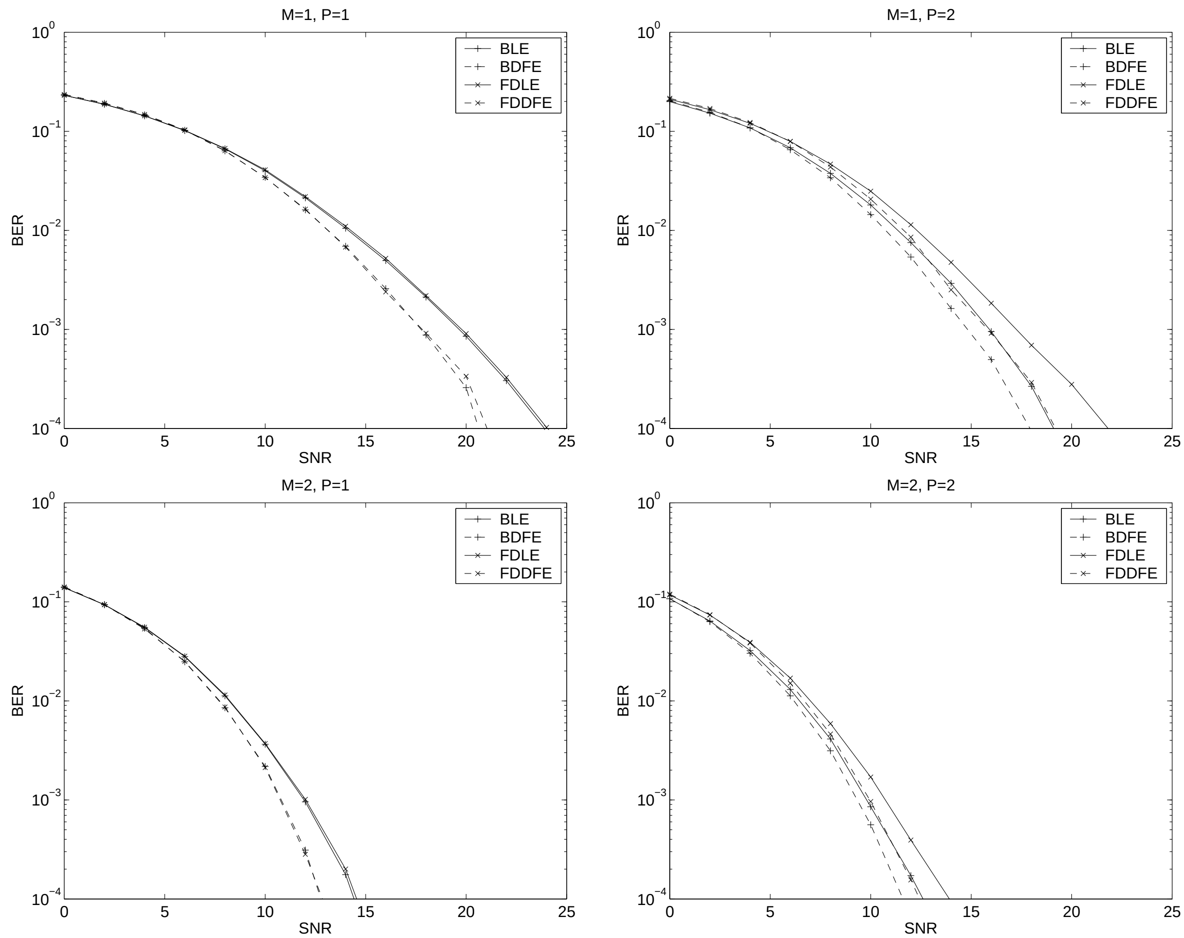

Fig. 10. Comparison of different block and frequency domain equalizers in TIV channels for $M = 1, 2$ and $P = 1, 2$.

extra $(N - L')(D'' - 1) = (N - L')(Q'' + 1)L'' = 4,825$ MA operations for the SDFE. Note, however, that the relative difference between the design or implementation complexity of the SLE (SDFE) and the BLE (BDFE) is smaller than for the TIV channels case. One could argue that for a smaller block size N, the difference would even disappear, but then the block size would not be large enough such that blind channel estimation becomes feasible or the overhead of the training symbols for training based channel estimation does not decrease the data transmission rate too much.

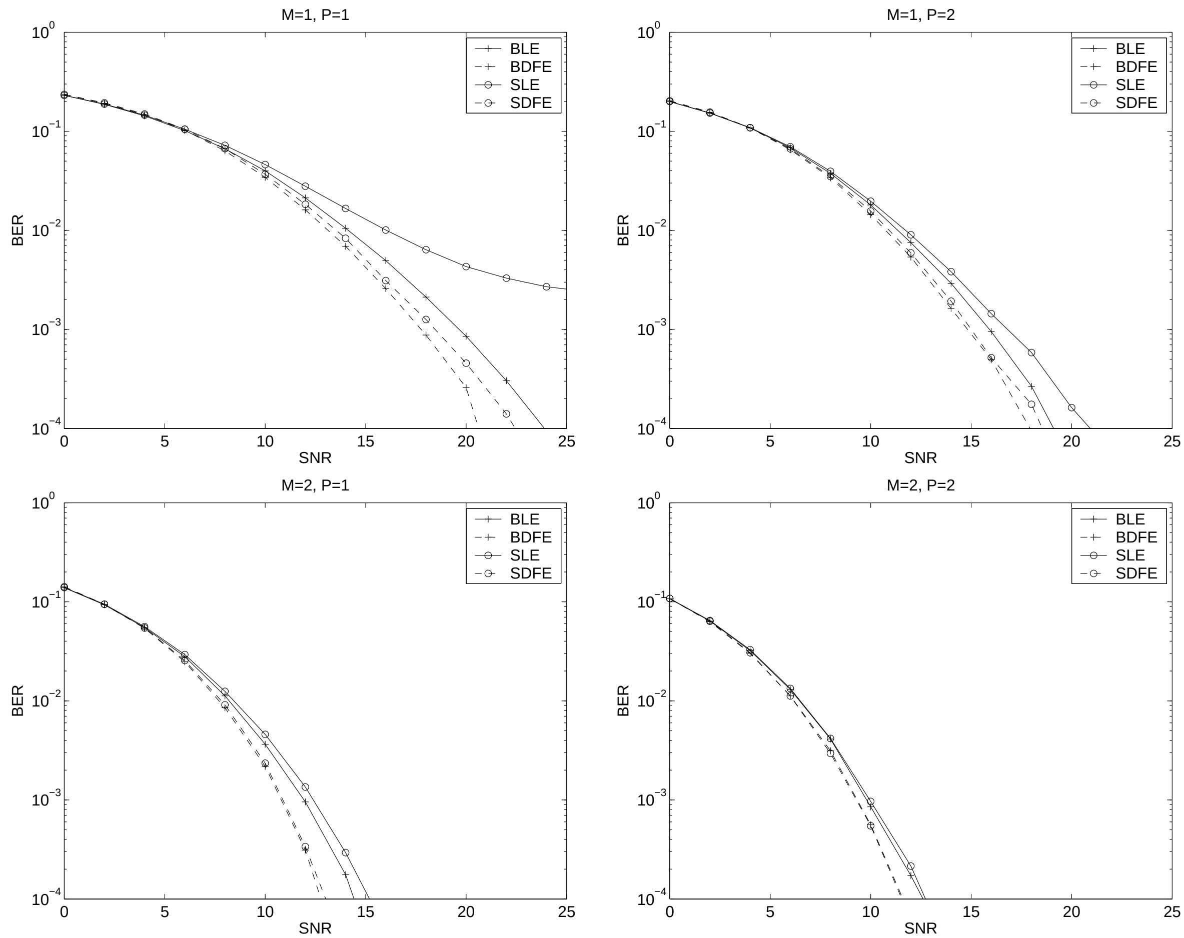

Fig. 11. Comparison of different block and serial equalizers in TIV channels for $M = 1, 2$ and $P = 1, 2$.

XII. SUMMARY

In this chapter, we have presented different practical finite-length equalization structures for TIV and TV channels. We have investigated linear and decision feedback block equalizers, linear and decision feedback serial equalizers, and linear and decision feedback frequency domain equalizers (the latter only apply to TIV channels).

All these channel equalizers are based on a practical channel model. Writing the overall system as a symbol rate SIMO system, where the multiple outputs are obtained by multiple receive antennas and/or fractional sampling, we can distinguish between TIV and TV channels, by looking at the channel time-variation over a fixed time-window. For TIV channels, we have modeled the channel by a TIV FIR channel, whereas for TV channels, it has been convenient

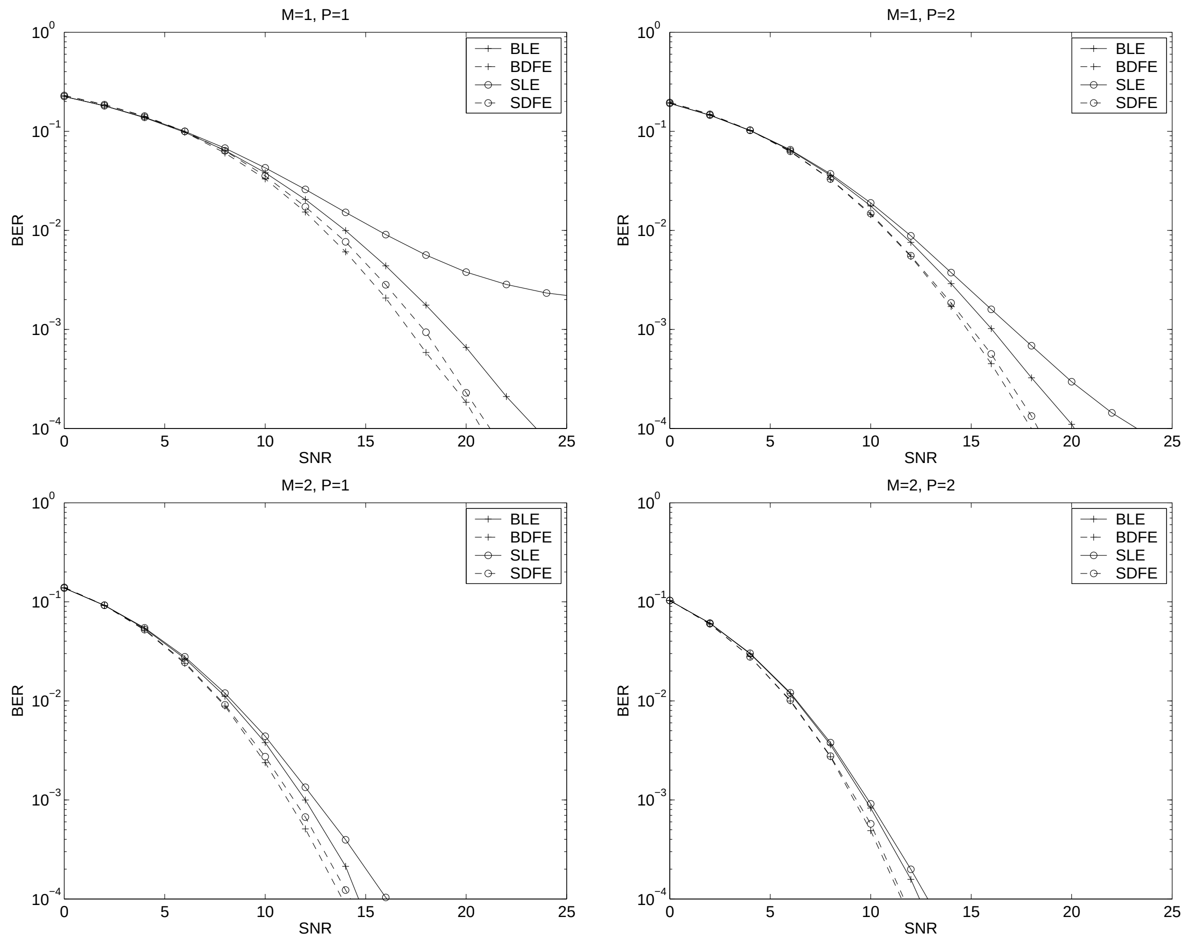

Fig. 12. Comparison of different block and serial equalizers in TV channels for $M = 1, 2$ and $P = 1, 2$.

to model the channel time-variation by means of a Basis Expansion Model (BEM), leading to a BEM FIR channel. Note that for the serial equalizers, we have used the same model for the equalizer as for the channel. Hence, for a TIV FIR channel, we have adopted a TIV FIR serial equalizer, whereas for a BEM FIR channel, we have adopted a BEM FIR serial equalizer. Note that in contrast with equalizing a BEM FIR channel with a TIV FIR serial equalizer, which requires a symbol rate SIMO channel with many outputs for the linear ZF solution to exist, equalizing a BEM FIR channel with a BEM FIR serial equalizer only requires a symbol rate SIMO channel with two outputs for the linear ZF solution to exist.

A complexity analysis and some illustrative simulation results have indicated that the SLE can have a much smaller design and implementation complexity than the BLE, without a significant loss in performance for a system with multiple outputs. For a system with a single output, there is a loss in performance at high SNR, since in contrast to the performance of the BLE, the performance of the SLE saturates at high SNR. Similarly, the SDFE can have a much smaller design and implementation complexity than the BDFE, without a significant loss in performance. This now even holds for a system with a single output, since as the performance of the BDFE, the performance of the SDFE does not saturate at high SNR. Finally, the FDLE (FDDFE) for TIV channels has a much smaller design and implementation complexity than the BLE (BDFE), while their performances are comparable. The comparison with the design and implementation complexity of the SLE (SDFE) for TIV channels depends on the specific scenario. However, note that FD processing can only be applied for TIV channels.

REFERENCES

[1] N. Al-Dhahir and J. M. Cioffi . MMSE Decision-Feedback Equalizers: Finite-Length Results. *IEEE Trans. on Information Theory*, 41(4):961–975, July 1995.

[2] I. Barhumi, G. Leus, and M. Moonen. Time-Varying FIR Decision Feedback Equalization of Doubly-Selective Channels. In *Proc. of IEEE Global Communications Conf. (GLOBECOM'03)*, San Francisco, California, December 2003.

[3] I. Barhumi, G. Leus, and M. Moonen. Time-Varying FIR Equalization of Doubly-Selective Channels. In *Proc. of IEEE Intl. Conf. on Communications (ICC'03)*, Anchorage, Alaska, May 2003.

[4] N. Benvenuto and S. Tomasin. On the Comparison Between OFDM and Single Carrier Modulation With a DFE Using a Frequency-Domain Feedforward Filter. *IEEE Trans. on Communcations*, 50(6):947–955, June 2002.

[5] J. K. Cavers. *Mobile Channel Characteristics*. Kluwer, 2000.

[6] A. Czylwik. Comparison Between Adaptive OFDM and Single Carrier Modulation With Frequency Domain Equalization. In *Proc. of IEEE Vehicular Technology Conference (VTC)*, pages 865–869, Phoenix, Arizona, May 1997.

[7] L. Deneire, B. Gyselinckx, and M. Engels. Training Sequence vs. Cyclic Prefi x: A New Look on Single Carrier Communication. In *Proc. of GLOBECOM*, San Fransisco, California, November/December 2000.

[8] A. Duel-Hallen. A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels. *IEEE Transactions on Communications*, 43(2/3/4):421–434, February/March/April 1995.

[9] A. Duel-Hallen, S. Hu, and H. Hallen. Long-Range Prediction of Fading Channels. *IEEE Signal Processing Mag.*, 17(3):62–75, May 2000.

[10] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson. Frequency Domain Equalization for Single-Carrier Broadband Wireless Systems. *IEEE Communications Magazine*, 40(4):58–66, April 2002.

[11] D. Gesbert, P. Duhamel, and S. Mayrargue. On-Line Blind Multichannel Equalization Based on Mutually Referenced Filters. *IEEE Transactions on Signal Processing*, 45(9):2307–2317, September 1997.

[12] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, editors. *Signal Processing Advances in Wireless & Mobile Communications: Trends in Channel Estimation and Equalization*, volume 1. Prentice Hall, 2000.

[13] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, editors. *Signal Processing Advances in Wireless & Mobile Communications: Trends in Single- and Multi-User Systems*, volume 2. Prentice Hall, 2000.

[14] G. B. Giannakis and C. Tepedelenlioğlu. Basis Expansion Models and Diversity Techniques for Blind Equalization of Time-Varying Channels. *Proc. of the IEEE*, 86(10):1969–1986, 1998.

[15] J. D. Gibson, editor. *The Mobile Communications Handbook*. CRC Press / IEEE Press, second edition, 1999.

[16] A. Ginesi, G. M. Vitetta, and D. D. Falconer. Block Channel Equalization in the Presence of a Cochannel Interferent Signal. *IEEE Journal on Selected Areas in Communications*, 17(11):1853–1862, 1999.

[17] W. C. Jakes, editor. *Microwave Mobile Communications*. Wiley, 1974.

[18] A. Klein, G. Kawas Kaleh, and P. W. Baier. Zero Forcing and Minimum Mean-Square-Error Equalization for Multiple-Access Channels. *IEEE Transactions on Vehicular Technology*, 45(2):276–287, May 1996.

[19] E. A. Lee and D. G. Messerschmitt. *Digital Communications*. Kluwer, second edition, 1994.

[20] G. Leus, I. Barhumi, and M. Moonen. MMSE Time-Varying FIR Equalization of Doubly-Selective Channels. In *Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'03)*, Hong Kong, April 2003.

[21] G. Leus and M. Moonen. Semi-Blind Channel Estimation for Block Transmission with Non-Zero Padding. In *Proc. of Asilomar Conf. on Signals, Systems and Computers*, pages 762–766, Pacific Grove, California, November 2001.

[22] G. Leus and M. Moonen. Deterministic Subspace Based Blind Channel Estimation for Doubly-Selective Channels. In *Proc. of the IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'03)*, Rome, Italy, June 2003.

[23] G. Leus, S. Zhou, and G. B. Giannakis. Orthogonal Multiple Access over Time- and Frequency-Selectvie Channels. *IEEE Trans. on Information Theory*, 49(8):1942–1950, August 2003.

[24] H. Liu and G. B. Giannakis. Deterministic Approaches for Blind Equalization of Time-Varying Channels with Antenna Arrays. *IEEE Trans. on Signal Processing*, 46(11):3003–3013, November 1998.

[25] R. Lupas and S. Verdú. Linear Multiuser Detectors for Synchronous Code-Division Multiple-Access Channels. *IEEE Transactions on Information Theory*, 35(1):123–136, January 1989.

[26] X. Ma and G. B. Giannakis. Maximum-Diversity Transmissions over Doubly-Selective Wireless Channels. *IEEE Trans. on Information Theory*, 49(7):1832–1840, July 2003.

[27] X. Ma, G. B. Giannakis, and S. Ohno. Optimal Training for Block Transmissions over Doubly-Selective Fading Channels. *IEEE Trans. on Signal Processing*, 51(5):1351–1366, May 2003.

[28] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. Subspace Methods for the Blind Identification of Multichannel FIR Filters. *IEEE Transactions on Signal Processing*, 43(2):516–525, February 1995.

[29] A. J. Paulraj and C. B. Papadias. Space-Time Processing for Wireless Communications. *IEEE Signal Processing Mag.*, 14(6):49–83, November 1997.

[30] J. G. Proakis. *Digital Communications*. McGraw-Hill, fourth edition, 2001.

[31] O. Rousseaux, G. Leus, and M. Moonen. Training Based Maximum Likelihood Channel Identification. In *Proc. of IEEE Workshop on Signal Processing Advances for Wireless communications (SPAWC)*, Rome, Italy, June 2003.

[32] H. Sari, G. Karam, and I. Jeanclaude. Transmission Techniques for Digital Terrestrial TV Broadcasting. *IEEE Communications Magazine*, 33(2):100–109, February 1995.

[33] A. M. Sayeed and B. Aazhang. Joint Multipath-Doppler Diversity in Mobile Wireless Communications. *IEEE Trans. on Communications*, 47(1):123–132, January 1999.

[34] A. Scaglione and G. B. Giannakis. Redundant Filterbank Precoders and Equalizers, Part I: Unification and Optimal Designs. *IEEE Trans. on Signal Processing*, 47(7):1988–2006, July 1999.

[35] A. Scaglione and G. B. Giannakis. Redundant Filterbank Precoders and Equalizers, Part II: Blind Channel Estimation, Synchronization, and Direct Equalization. *IEEE Trans. on Signal Processing*, 47(7):2007–2022, July 1999.

[36] D. T. M. Slock. Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction. In *Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'94)*, pages IV/585–IV/588, Adelaide, Australia, April 1994.

[37] A. Stamoulis, G. B. Giannakis, and A. Scaglione. Block FIR Decision-Feedback Equalizers for Filterbank Precoded Transmissions with Blind Channel Estimation Capabilities. *IEEE Trans. on Communications*, 49(1):69–83, January 2001.

[38] C. Tepedelenlioğlu and G. B. Giannakis. Transmitter Redundancy for Blind Estimation and Equalization of Time- and Frequency-Selective Channels. *IEEE Trans. on Signal Processing*, 48(7):2029–2043, July 2000.

[39] J. R. Treichler, I. Fijalkow, and C. R. Johnson, Jr. Fractionally Spaced Equalizers: How long should they really be? *IEEE Signal Processing Magazine*, 13(3):65–81, May 1996.

[40] M. K. Tsatsanis and G. B. Giannakis. Modeling and Equalization of Rapidly Fading Channels. *International Journal of Adaptive Control and Signal Processing*, 10(2/3):159–176, March 1996.

[41] M. K. Tsatsanis and Z. Xu. Performance Analysis of Minimum Variance Receivers. *IEEE Transactions on Signal Processing*, 46(11):3014–3022, November 1998.

[42] J. K. Tugnait, L. Tong, and Z. Ding. Single-User Channel Estimation and Equalization. *IEEE Signal Processing Magazine*, 17(3):17–28, May 2000.

[43] G. Ungerboeck. Fractional Tap-Spacing Equalizer and Consequences for Clock Recovery in Data Modems. *IEEE Trans. on communications*, 24(8):856–864, August 1976.

[44] A.-J. van der Veen, S. Talwar, and A. Paulraj. A Subspace Approach to Blind Space-Time Signal Processing for Wireless Communication Systems. *IEEE Transactions on Signal Processing*, 45(1):173–190, January 1997.

[45] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath. Optimal Training for Frequency-Selective Channels. In *Proc. of ICASSP*, Salt Lake City, Utah, May 2001.

[46] Z. Wang and G. B. Giannakis. Wireless Multicarrier Communications: Where Fourrier Meets Shannon. *IEEE Signal Processing Magazine*, 17(3):29–48, May 2000.