Performance Bounds for Multi-Step Prediction Based Blind Equalization

J. Mannerkoski*, V. Koivunen, Senior Member, IEEE
Signal Processing Laboratory
Helsinki Univ. of Technology
P.O. Box 3000
FIN-02015 HUT, FINLAND
{jukka.mannerkoski,visa.koivunen}@hut.fi

Desmond P. Taylor, Fellow, IEEE
Electrical and Electronic Eng.
University of Canterbury
Private Bag 4800
Christchurch, NEW ZEALAND
taylor@elec.canterbury.ac.nz

Index Terms – blind equalization, prediction-error filtering, cyclostationarity.

Abstract

Blind equalization attempts to remove the interference caused by a communication channel without using any known training sequences. Blind equalizers may be implemented with linear prediction-error filters (PEF). For many practical channel types, a suitable delay at the output of the equalizer allows for achieving a small estimation error. The delay cannot be controlled with one-step predictors. Consequently, multi-step PEF based algorithms have been suggested as a solution to the problem. The derivation of the existing algorithms is based on the assumption of a noiseless channel, which results in zero-forcing equalization. We consider the effects of additive noise at the output of the multi-step PEF. Analytical error bounds for two PEF-based blind equalizers in the presence of noise are derived. The obtained results are verified with simulations. The effect of energy concentration in the channel impulse response on the error bound is also addressed.

*This work was supported by the Academy of Finland.

1 Introduction

A communication channel introduces inter-symbol interference (ISI) in the received sequence. Blind equalization attempts to remove the ISI without the knowledge of the transmitted sequence. If multiple samples per symbol are available at the receiver, blind equalizers can be derived from the second-order cyclostationary statistics of the received signal [3, 11]. Zero-forcing (ZF) and minimum mean-square error (MMSE) equalizers can be implemented with linear multichannel prediction-error filters (PEF) [9, 10]. However, one-step forward prediction can only result in zero-delay equalization and one-step backward prediction in maximum-delay equalization. For many practical channels, a smaller equalization error may be achieved by controlling the delay (c.f. [5]). The selected delay should reflect the distribution of energy in the channel impulse response.

Multi-step prediction has been suggested as a solution to the arbitrary-delay equalization problem in [4], [10] and [8]. The approach is based on the assumption of zero noise, so that the multi-step PEF perfectly reduces the degree of the ISI in the received sequence. Consequently, the prediction error can be modeled as an output of a truncated channel with no additive noise. Based on this model, simple blind ZF-equalizers have been developed. The equalizer of [4] employs a combination of two multi-step PEF outputs. We call it the Forward-Forward PEF equalizer (FF-PEF). The equalizer of [10] consists of a cascade of a multi-step forward and a one-step backward PEF, denoted here the Forward-Backward PEF equalizer (FB-PEF). Other PEF-based algorithms have also been developed (c.f, [4, 8]). We have chosen the FF-PEF and FB-PEF for this paper based on their relative simplicity and robustness.

With both considered algorithms, the transmitted symbol can be reconstructed without error in the absence of noise (ZF-equalization). In this paper, we investigate how the additive noise affects the interpretation of the PEF output as an output of the truncated channel. In the presence of additive white noise, the PEF output includes a colored noise component. Moreover, the multi-step PEF is no longer able to reduce the degree of the ISI perfectly, but contains an interference component from the previous symbols at its output. The residual ISI and the colored noise introduce error in the system output. We derive an analytical expression for the error.

This paper is organized as follows. The system model is presented in section 2. Multi-step forward prediction-error filtering is considered in the context of channel equalization in section 3. In section 4, the FF-PEF and the FB-PEF are described and a lower bound for the mean-square error (MSE) — implicitly containing the variance and the estimation bias — of the equalizer output is derived for both algorithms. Analytical and simulated results are presented in section 5. Section 6 concludes the paper.

2 System model

Let $x(k)$ denote the symbol emitted by a random digital source at time kT, T being the symbol period. The channel is considered to be linear and time-invariant (LTI) during a time window sufficient to allow its estimation. The continuous time received signal is

$$y(t) = \sum_{k=-\infty}^{\infty} x(k)h(t-kT) + v(t) \tag{1}$$

where $h(t)$ is the complex baseband channel impulse response, $x(k)$ is the sequence of complex information symbols and $v(t)$ is additive noise.

When $P \geq 2$ samples of $y(t)$ per symbol interval are available at the receiver, the system can be modeled as consisting of P subchannels. The ith subchannel response is defined as $h_i(n) \triangleq h(t_0 + iT/P + nT)$, $n = 0, \ldots, L_h - 1$, where L_h is the subchannel length. Its output sequence $y_i(n) \triangleq y(t_0 + iT/P + nT)$ is

$$y_i(n) = \sum_{k=0}^{L_h-1} x(n-k)h_i(k) + v_i(n) \tag{2}$$

where $i = 0, \ldots, P-1$, $-\infty < n < \infty$, $v_i(n)$ are samples of $v(t)$ corresponding to $y_i(n)$ and the symbol period is normalized to $T = 1$. The received sequence $y(n/P)$ is cyclostationary, which enables blind equalization of the channel from the second-order statistics of the received sequence, if the subchannels $h_i(n)$ have no zero in common [2].

In each symbol interval, a vector $\mathbf{y}(n) = [y_0(n), \ldots, y_{P-1}(n)]^T$ of length P is received, $()^T$ denoting transposition. The channel impulse response can also be represented in vector form as $\mathbf{h}(n) = [h_0(n), \ldots, h_{P-1}(n)]^T$ and the noise as $\mathbf{v}(n) = [v_0(n), \ldots, v_{P-1}(n)]^T$. The scalar cyclostationary process $y(n/P)$ is thus transformed into a wide-sense stationary vector process.

We consider a finite symbol sequence of length M and define a matrix $\mathcal{H}_M$ with dimension $MP \times (L_h + M - 1)$ consisting of the channel coefficient vectors,

$$\mathcal{H}_M = \begin{bmatrix} \mathbf{h}(0) & \dots & \mathbf{h}(L_h-1) & & \mathbf{0}_{P\times(M-1)} \\ & \ddots & & \ddots & \\ \mathbf{0}_{P\times(M-1)} & & \mathbf{h}(0) & \dots & \mathbf{h}(L_h-1) \end{bmatrix} \tag{3}$$

Stacking M received vector samples into an $(MP \times 1)$-vector $\mathbf{y}_M(n) = [\mathbf{y}^T(n), \dots, \mathbf{y}^T(n - M + 1)]^T$, we can write a matrix equation as

$$\mathbf{y}_M(n) = \mathcal{H}_M \mathbf{x}_{L_h+M-1}(n) + \mathbf{v}_M(n) \tag{4}$$

where the input vector is $\mathbf{x}_{L_h+M-1}(n) = [x(n), \dots, x(n - L_h - M + 2)]^T$ [12]. The block-Toeplitz correlation matrix $\mathbf{R}_{M,y}$ of $\mathbf{y}_M(n)$ is defined as

$$\mathbf{R}_{M,y} = \begin{bmatrix} \mathbf{r}(0) & \mathbf{r}(1) & \dots & \mathbf{r}(M-1) \\ \mathbf{r}^H(1) & \mathbf{r}(0) & \dots & \mathbf{r}(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}^H(M-1) & \mathbf{r}^H(M-2) & \dots & \mathbf{r}(0) \end{bmatrix} \tag{5}$$

where $\mathbf{r}(i-j) = \mathrm{E}\{\mathbf{y}(n-j)\mathbf{y}^H(n-i)\}$ and $()^H$ denotes Hermitian transposition.

At this stage, we state the assumptions used throughout the paper:

A1. The input sequence $x(k)$ is zero-mean and white with unit variance.

A2. The additive noise is zero-mean and white with variance σ_v^2.

A3. The sequences $x(k)$ and $v(k)$ are uncorrelated.

A4. The matrix $\mathcal{H}_M$ has full column rank, i.e. the subchannels $h_i(n)$ have no common zeros. Also, $h_i(0) \neq 0$ for some i and $h_i(L_h - 1) \neq 0$ for some i.

A5. The dimensions of $\mathcal{H}_M$ obey $MP \geq L_h + M - 1$.

Using assumptions A1 – A3, we can write the exact correlation matrix of $\mathbf{y}_M(n)$ as

$$\mathbf{R}_{M,y} = \mathcal{H}_M \mathcal{H}_M^H + \sigma_v^2 \mathbf{I}_{MP} \tag{6}$$

where $\mathbf{I}_{MP}$ is an $MP \times MP$ identity matrix.

3 Multi-step forward prediction-error filtering

In the absence of noise, multi-step prediction can be used to decrease the degree of ISI in the received signal without error. Consequently, the PEF output can be modeled as an output of a truncated channel. In this section, we consider the interference component that additive noise introduces at the output of the truncated channel. When the predictor coefficients are computed with the noise component present in (6), the resulting ZF-equalizers are biased [4]. However, we choose not to estimate and subtract the noise term from the block-Toeplitz correlation matrix, because it would consequently become singular. As considered in [1], this would result in robustness concerns when the channel order is not known. Accurate estimates of the noise power and channel order are needed for the unbiased ZF-equalizers, which is not necessarily feasible in practice. Simulation results presented in section 5, showing that unbiased estimation may also lead to an increased MSE, support the choice.

A k-step forward predictor of order M produces an estimate $\hat{\mathbf{y}}(n)$ of the received symbol $\mathbf{y}(n)$ based on the M previous symbols $\mathbf{y}_M(n-k)$. The filter coefficients $\mathbf{w}_{M,k}$ are obtained from the Wiener-Hopf equations

$$\mathbf{w}_{M,k} = \mathbf{r}_k \mathbf{R}_{M,y}^{-1} \tag{7}$$

where $\mathbf{r}_k = [\mathbf{r}(k), \ldots, \mathbf{r}(k+M-1)]$. A corresponding k-step forward prediction error filter of order M produces the error $\mathbf{f}_{M,k}(n) = \mathbf{y}(n) - \hat{\mathbf{y}}(n)$ as its output. The PEF coefficients may be represented in matrix form as $\mathbf{A}_{M,k} = [\mathbf{I}_P \;\mathbf{0}_{P\times P(k-1)} \;\; -\mathbf{w}_{M,k}]$. The k-step prediction error is then

$$\mathbf{f}_{M,k}(n) = \mathbf{A}_{M,k}\mathbf{y}_{M+k}(n) \tag{8}$$

The PEF output $\mathbf{f}_{M,k}(n)$ can be given a meaningful interpretation in the context of channel equalization as the output of a truncated channel $\tilde{\mathbf{h}}(n) \triangleq \mathbf{h}(n), n = 0, \ldots, k-1$.

Theorem 1 *The output* $\mathbf{f}_{M,k}(n)$ *of the* k*-step PEF is given by*

$$\mathbf{f}_{M,k}(n) = \sum_{j=0}^{k-1} \mathbf{h}(j)x(n-j) + \boldsymbol{\xi}_{M,k}(n) \tag{9}$$

where $\boldsymbol{\xi}_{M,k}(n) = \sigma_v^2 \mathbf{w}_{M,k} \mathcal{H}_M^{\dagger H} \mathbf{x}_{L_h+M-1}(n-k) + \mathbf{A}_{M,k}\mathbf{v}_{M+k}(n)$, $\mathcal{H}_M^{\dagger}$ *is the pseudo-inverse of* $\mathcal{H}_M$ *and* $MP \geq L_h + M - 1$.

The proof is in the Appendix. In the noiseless case, the number of transmitted symbols affecting any received symbol (degree of ISI) is reduced from L_h to k. As shown in the Appendix, the additive noise introduces an unwanted component $\boldsymbol{\xi}_{M,k}(n)$ comprising of residual ISI and colored noise at the output of the partial channel. Fig. 1 summarizes the ideas presented above.

The covariance matrix $\mathbf{J}_{\boldsymbol{\xi}} = \mathrm{E}\{\boldsymbol{\xi}_{M,k}(n)\boldsymbol{\xi}_{M,k}^H(n)\}$ of the error in (9) is expanded in Theorem 2.

Theorem 2 *The covariance matrix of the residual ISI plus noise at the* $\tilde{\mathbf{h}}(n)$ *output is*

$$\mathbf{J}_{\boldsymbol{\xi}} = \boldsymbol{\gamma}_{M,k}\boldsymbol{\gamma}_{M,k}^H + \sigma_v^2 \mathbf{A}_{M,k}\mathbf{A}_{M,k}^H \tag{10}$$

where the $P \times (L_h + M - 1)$*-matrix* $\boldsymbol{\gamma}_{M,k} \triangleq \sigma_v^2 \mathbf{w}_{M,k}\mathcal{H}_M^{\dagger H}$.

The proof is obvious from Theorem 1 using assumptions A1 – A3.

4 Blind equalization based on multi-step prediction

We consider two blind ZF-equalizers that are based on multi-step prediction. Modeling the prediction error as an output of a truncated channel makes the derivation of ZF-equalizers straightforward. The FF-PEF, introduced in [4], employs two forward PEFs in parallel. The FB-PEF was introduced in [10], where it was suggested that a forward and a backward PEF should be used in cascade. An analytical expression for the minimum MSE containing both the squared bias term and the variance is derived for each equalizer.

4.1 The FF-PEF: Two forward prediction-error filters in parallel

The FF-PEF is based on the outputs of k- and $(k+1)$-step PEFs [4]. A k-delay ZF-equalizer can be obtained after acknowledging from (9) that in the absence of noise

$$\mathbf{F}_{M,k}(n) \triangleq \mathbf{f}_{M,k+1}(n) - \mathbf{f}_{M,k}(n) = \mathbf{h}(k)x(n-k) \tag{11}$$

The transmitted symbol can then be extracted as $x(n-k) = \boldsymbol{\nu}(k)\mathbf{F}_{M,k}(n)$, where $\boldsymbol{\nu}(k) = \mathbf{h}^H(k)/[\mathbf{h}^H(k)\mathbf{h}(k)]$, which can be estimated from the correlation matrix $\mathbf{P}_{\mathbf{F}_{M,k}} = \mathrm{E}[\mathbf{F}_{M,k}(n)\ \mathbf{F}^H_{M,k}(n)]$ [1]. Let us denote the unit-norm eigen vector corresponding to the non-zero eigen value of $\mathbf{P}_{\mathbf{F}_{M,k}}$ by $\mathbf{u}$ and the corresponding eigen value by λ. In the absence of noise, $\mathbf{P}_{\mathbf{F}_{M,k}} = \mathbf{h}(k)\mathbf{h}^H(k)$ and $\boldsymbol{\nu}(k) = \mathbf{u}^H/\sqrt{\lambda}$ up to a unitary complex constant. This scalar rotation is an inherent part of any SOS-based blind equalizer and can be solved by differential encoding in applications.

When additive noise is present,

$$\hat{\mathbf{F}}_{M,k}(n) = \mathbf{h}(k)x(n-k) + \boldsymbol{\xi}_{M,k+1}(n) - \boldsymbol{\xi}_{M,k}(n) \tag{12}$$

The error term $\boldsymbol{\xi}_{M,k+1}(n) - \boldsymbol{\xi}_{M,k}(n)$ accumulates into the correlation matrix $\hat{\mathbf{P}}_{\mathbf{F}_{M,k}} = \mathrm{E}[\hat{\mathbf{F}}_{M,k}(n)\ \hat{\mathbf{F}}^H_{M,k}(n)]$. Using assumptions A1 – A3, we can write

$$\hat{\mathbf{P}}_{\mathbf{F}_{M,k}} = \mathbf{h}(k)\mathbf{h}^H(k) + \mathbf{J}_{\boldsymbol{\gamma}} + \mathbf{J}_v + \mathbf{J}_{\mathbf{h}(k)} \tag{13}$$

where

$$\begin{aligned}
\mathbf{J}_{\boldsymbol{\gamma}} \triangleq{}& \boldsymbol{\gamma}_{M,k+1}\boldsymbol{\gamma}^H_{M,k+1} + \boldsymbol{\gamma}_{M,k}\boldsymbol{\gamma}^H_{M,k} \\
& -\boldsymbol{\gamma}_{M,k+1}\mathbf{D}_1\boldsymbol{\gamma}^H_{M,k} - \boldsymbol{\gamma}_{M,k}\mathbf{D}_1^H\boldsymbol{\gamma}^H_{M,k+1} \\
\mathbf{J}_v \triangleq{}& \sigma_v^2[\mathbf{A}_{M,k+1}\mathbf{A}^H_{M,k+1} + \mathbf{A}_{M,k}\mathbf{A}^H_{M,k} \\
& -\mathbf{A}_{M,k+1}\mathbf{I}^H_{\mathrm{left}}\mathbf{A}^H_{M,k} - \mathbf{A}_{M,k}\mathbf{I}_{\mathrm{left}}\mathbf{A}^H_{M,k+1}] \\
\mathbf{J}_{\mathbf{h}(k)} \triangleq{}& -\boldsymbol{\gamma}_{M,k}\mathbf{j}\mathbf{h}^H(k) - \mathbf{h}(k)\mathbf{j}^H\boldsymbol{\gamma}^H_{M,k}
\end{aligned}$$

and $\mathbf{D}_1$ is an $(L_h + M - 1)$-square matrix with ones on its first diagonal above the main diagonal and zeros elsewhere, $\mathbf{I}_{\mathrm{left}} = [\mathbf{I}_{P(M+k)}\ \mathbf{0}_{P\times P(M+k)}]$ and the $L_h + M - 1$-vector $\mathbf{j} = [1\ 0 \ldots 0]^H$.

The additional terms in $\hat{\mathbf{P}}_{\mathbf{F}_{M,k}}$ result in an error in the estimate $\hat{\boldsymbol{\nu}}(k) = \hat{\mathbf{u}}^H/\sqrt{\hat{\lambda}}$, where $\hat{\lambda}$ is the largest eigen value of $\hat{\mathbf{P}}_{\mathbf{F}_{M,k}}$ and $\hat{\mathbf{u}}$ the corresponding eigen vector. Let us define $\boldsymbol{\nu}_e(k) = \hat{\boldsymbol{\nu}}(k) - \boldsymbol{\nu}(k)$. The FF-PEF output is now

$$\begin{aligned}\hat{x}(n-k) =& \hat{\boldsymbol{\nu}}(k)[\mathbf{h}(k)x(n-k) + \boldsymbol{\xi}_{M,k+1}(n) - \boldsymbol{\xi}_{M,k}(n)] \\ =& x(n-k) + \hat{\boldsymbol{\nu}}(k)[\boldsymbol{\xi}_{M,k+1}(n) - \boldsymbol{\xi}_{M,k}(n)] + \boldsymbol{\nu}_e(k)\mathbf{h}(k)x(n-k)\end{aligned} \tag{14}$$

The minimum achievable MSE $J_{\min} = \mathrm{E}\{[\hat{x}(n) - x(n)]^2\}$ of the symbol estimate is given in Theorem 3.

Theorem 3 *A lower bound for the MSE of the FF-PEF blind equalizer output is*

$$J_{min} = \hat{\boldsymbol{\nu}}(k)[\mathbf{J}_{\gamma} + \mathbf{J}_v]\hat{\boldsymbol{\nu}}^H(k) + J_{\boldsymbol{\nu}_e} \tag{15}$$

where

$$\begin{aligned}J_{\boldsymbol{\nu}_e} \triangleq& \boldsymbol{\nu}_e(k)\mathbf{h}(k)\mathbf{h}^H(k)\boldsymbol{\nu}_e^H(k) \\ -& \hat{\boldsymbol{\nu}}(k)\boldsymbol{\gamma}_{M,k}\mathbf{j}\mathbf{h}^H(k)\boldsymbol{\nu}_e^H(k) - \boldsymbol{\nu}_e(k)\mathbf{h}(k)\mathbf{j}^H\boldsymbol{\gamma}_{M,k}^H\hat{\boldsymbol{\nu}}^H(k)\end{aligned} \tag{16}$$

The proof is straightforward from (14) using assumptions A1 – A3 and the result in (13). Component $\mathbf{J}_{\gamma}$ is due to the residual ISI and $\mathbf{J}_v$ to the colored noise. $J_{\boldsymbol{\nu}_e}$ results from the error in estimating $\boldsymbol{\nu}(k)$, which also affects the whole expression of $J_{\min}$ in $\hat{\boldsymbol{\nu}}(k)$.

4.2 The FB-PEF: A forward and a backward prediction-error filter in cascade

A blind d-delay equalizer can be realized as a combination of a multi-step forward and a one-step backward PEF, as suggested in [10]. The FB-PEF directly exploits the interpretation of the forward PEF output as an output of the truncated channel. In this section, we first present the relation between a one-step backward prediction error of the received sequence and the transmitted sequence. Next, we introduce an extension to the system model of section 2. In order to illustrate what the error of the FB-PEF is comprised of, we give three ways to present the multi-step PEF output in vector form. Finally, we derive an analytic expression for the MSE of the symbol estimate.

4.2.1 Backward prediction-error filtering

In one-step backward linear prediction the task is to use a linear combination of the M_b symbols $\mathbf{y}_{M_b}(n)$ to estimate the previously received symbol $\mathbf{y}(n-M_b)$. The Wiener-Hopf equations for backward linear prediction can be written as [6]

$$\mathbf{w}_{b,M_b} = \mathbf{r}_1^b \mathbf{R}_{M_b}^{-1} \tag{17}$$

where the $\mathbf{w}_{b,M_b}$ are the filter coefficients, M_b is the backward prediction order and $\mathbf{r}_1^b = [\mathbf{r}^H(M_b), \mathbf{r}^H(M_b-1), \ldots, \mathbf{r}^H(1)]$ (c.f. (5)).

A backward prediction error of the received sequence is defined as

$$\mathbf{b}_{M_b}(n) = \mathbf{B}_{M_b}\mathbf{y}_{M_b+1}(n) \tag{18}$$

where the filter coefficients are represented in matrix form as $\mathbf{B}_{M_b} = [-\mathbf{w}_{b,M_b}\ \mathbf{I}_P]$. The backward prediction error is related to the transmitted symbol as [7]

$$\mathbf{b}_{M_b}(n) = \mathbf{h}(L_h-1)x(n-L_h-M_b+1) \tag{19}$$

resulting in a delay of (L_h+M_b-1) in the equalizer output. This is, however, maximum-delay equalization, which suffers from similar limitations to zero-delay equalization.

As shown in (9), in the noiseless case the output of the k-step forward PEF equals the output of the truncated channel $\tilde{\mathbf{h}}(n)$ of length k. Therefore, using a cascade of a k-step forward and 1-step backward PEF results in a $(k+M_b-1)$-delay equalizer, where the delay can be controlled. The additive interference $\boldsymbol{\xi}_{M,k}(n)$ in (9) induces an error in the equalizer output, because the derivation is based on the assumption of zero noise. The interference also propagates into the computation of the backward PEF coefficients $\mathbf{B}_{M_b}$.

4.2.2 The multi-step PEF output in vector form

In order to use the multi-step PEF output $\mathbf{f}_{M,k}(n)$ as an input to the backward PEF in (18), we need to present the PEF outputs in vector form. We consider a vector of M_b outputs as $\tilde{\mathbf{y}}_{M_b}(n) = [\mathbf{f}_{M,k}^T(n), \ldots, \mathbf{f}_{M,k}^T(n-M_b-1)]^T$, where the symbol $\tilde{\mathbf{y}}_{M_b}(n)$ is chosen to emphasize the interpretation of $\mathbf{f}_{M,k}(n)$ as an output of the partial channel $\tilde{\mathbf{h}}(n)$. Three different expressions for $\tilde{\mathbf{y}}_{M_b}(n)$ will be used later in the analysis of the FB-PEF performance. First, we can write (c.f. (8))

$$\tilde{\mathbf{y}}_{M_b}(n) = \mathcal{A}_{M_b,k}\mathbf{y}_{P(M+k+M_b-1)}(n) \tag{20}$$

where the $M_bP \times P(M+k+M_b-1)$-matrix may be written as

$$\mathcal{A}_{M_b,k} = \begin{bmatrix} \mathbf{A}_{M,k} & & \mathbf{0}_{P\times P(M_b-1)} \\ \mathbf{0}_{P\times P} & \mathbf{A}_{M,k} & \mathbf{0}_{P\times P(M_b-2)} \\ & \ddots & \\ \mathbf{0}_{P\times P(M_b-1)} & & \mathbf{A}_{M,k} \end{bmatrix} \tag{21}$$

Alternatively, we get a second expression for $\tilde{\mathbf{y}}_{M_b}(n)$ by subdividing the rhs of (20) into two parts consisting of an original sequence component and a colored noise component. From Theorem 1, we see that

$$\tilde{\mathbf{y}}_{M_b}(n) = \Pi_{M_b}\mathbf{x}_{L_h+M+k+M_b-2}(n) + \mathcal{A}_{M_b,k}\mathbf{v}_{M+k+M_b-1}(n) \tag{22}$$

where Π_{M_b} is the $M_bP \times (L_h+M+k+M_b-2)$-matrix

$$\Pi_{M_b} = \begin{bmatrix} \mathbf{h}(0) & \dots & \mathbf{h}(k-1) & \boldsymbol{\gamma}_{M,k} & \mathbf{0}_{P\times(M_b-1)} \\ & \ddots & & \ddots & \\ \mathbf{0}_{P\times(M_b-1)} & \mathbf{h}(0) & \dots & \mathbf{h}(k-1) & \boldsymbol{\gamma}_{M,k} \end{bmatrix} \tag{23}$$

with $\boldsymbol{\gamma}_{M,k} = \sigma_v^2\mathbf{w}_{M,k}\mathcal{H}_M^{\dagger H}$ as defined earlier.

We will further subdivide the first component on the rhs of (22) into two parts to obtain the third representation for $\tilde{\mathbf{y}}_{M_b}(n)$. From (22) and (23), we see that $\tilde{\mathbf{y}}_{M_b}(n)$ is comprised of the partial channel output, the residual ISI and the colored noise as

$$\begin{aligned}\tilde{\mathbf{y}}_{M_b}(n) =& \tilde{\mathcal{H}}_{M_b}\mathbf{x}_{k+M_b-1}(n) + \Gamma_{M_b,k}\mathbf{x}_{L_h+M+M_b-2}(n-k) \\ &+\mathcal{A}_{M_b,k}\mathbf{v}_{M+k+M_b-1}(n)\end{aligned} \tag{24}$$

using the $M_bP \times (k+M_b-1)$-matrix

$$\tilde{\mathcal{H}}_{M_b} = \begin{bmatrix} \mathbf{h}(0) & \dots & \mathbf{h}(k-1) & & \mathbf{0}_{P\times(M_b-1)} \\ & \ddots & & \ddots & \\ \mathbf{0}_{P\times(M_b-1)} & & \mathbf{h}(0) & \dots & \mathbf{h}(k-1) \end{bmatrix} \tag{25}$$

and the $M_b P \times (L_h + M + M_b - 2)$-matrix

$$\Gamma_{M_b,k} = \begin{bmatrix} \boldsymbol{\gamma}_{M,k} & & \mathbf{0}_{P\times(M_b-1)} \\ \mathbf{0}_{P\times 1} & \boldsymbol{\gamma}_{M,k} & \mathbf{0}_{P\times(M_b-2)} \\ & \ddots & \\ \mathbf{0}_{P\times(M_b-1)} & & \boldsymbol{\gamma}_{M,k} \end{bmatrix} \tag{26}$$

The matrices $\mathcal{A}_{M_b,k}$, Π_{M_b}, $\tilde{\mathcal{H}}_{M_b}$ and $\Gamma_{M_b,k}$ will be used later to express the MSE of the FB-PEF output.

4.2.3 Analysis of the FB-PEF performance

The analysis of this system is somewhat more tedious than the previous one. From (24) we can see that $\tilde{\mathbf{y}}_{M_b}(n)$ includes an interference term in addition to the partial channel output, when noise is present. This results in an error in computing the backward PEF coefficients $\mathbf{B}_{M_b}$. In the ideal noiseless case, the backward PEF $\mathbf{B}^0_{M_b}$ would be computed from (17) using the correlation matrix of the interference-free output of the truncated channel $\tilde{\mathbf{h}}(n)$,

$$\mathbf{R}^0_{M_b,\tilde{y}} = \tilde{\mathcal{H}}_{M_b}\tilde{\mathcal{H}}^H_{M_b} \tag{27}$$

In reality, the coefficients $\mathbf{B}_{M_b}$ are computed from a correlation matrix that includes the residual ISI and the colored noise (c.f. (4) and (20))

$$\mathbf{R}_{M_b,\tilde{y}} = \mathcal{A}_{M_b,k}\mathcal{H}_{M_b}\mathcal{H}^H_{M_b}\mathcal{A}^H_{M_b,k} + \sigma_v^2\mathcal{A}_{M_b,k}\mathcal{A}^H_{M_b,k} \tag{28}$$

which results in an error in the backward PEF coefficients, namely

$$\mathbf{B}^{\mathrm{err}}_{M_b} = \mathbf{B}_{M_b} - \mathbf{B}^0_{M_b} \tag{29}$$

Consequently, the output of the backward PEF $\mathbf{B}_{M_b}$ can be defined as

$$\mathbf{b}_{M_b}(n) = \mathbf{B}^0_{M_b}\tilde{\mathbf{y}}_{M_b+1}(n) + \mathbf{B}^{\mathrm{err}}_{M_b}\tilde{\mathbf{y}}_{M_b+1}(n) \tag{30}$$

which can be written as

$$\mathbf{b}_{M_b}(n) = \mathbf{h}(k-1)x(n-k-M_b+1) + \boldsymbol{\varepsilon}(n) \tag{31}$$

where we have defined the error according to (24) as

$$\begin{aligned}\boldsymbol{\varepsilon}(n) \triangleq& \mathbf{B}^{0}_{M_b}\Gamma_{M_b+1,k}\mathbf{x}_{L_h+M+M_b-1}(n-k) \\ &+\mathbf{B}^{0}_{M_b}\mathcal{A}_{M_b+1,k}\mathbf{v}_{M+k+M_b}(n)+\mathbf{B}^{\mathrm{err}}_{M_b}\tilde{\mathbf{y}}_{M_b+1}(n)\end{aligned} \tag{32}$$

The two first components on the rhs of (32) are the residual ISI and colored noise in $\tilde{\mathbf{y}}_{M_b+1}(n)$ at the output of the ideal filter. The third component is due to the error in computing the backward PEF coefficients.

In the absence of noise, the equalizer output is $x(n-k-M_b+1) = \boldsymbol{\nu}(k-1)\mathbf{b}_{M_b}(n)$, where $\boldsymbol{\nu}(k-1) = \mathbf{h}^H(k-1)/[\mathbf{h}^H(k-1)\mathbf{h}(k-1)]$, which is computed as $\mathbf{u}/\sqrt{\lambda}$, where λ is the largest eigen value of $\mathbf{P_b} = \mathrm{E}[\mathbf{b}_{M_b}(n)\mathbf{b}^H_{M_b}(n)]$ and $\mathbf{u}$ is the corresponding unit-norm eigen vector. In a noiseless system, $\mathbf{P_b} = \mathbf{h}(k-1)\mathbf{h}^H(k-1)$ and $\boldsymbol{\nu}(k-1)$ as well as the transmitted symbol are retrieved perfectly.

Additive noise results in an error in $\hat{\mathbf{P}}_{\mathbf{b}} = \mathrm{E}\{[\mathbf{h}(k-1)x(n-k-M_b+1)+\boldsymbol{\varepsilon}(n)][\mathbf{h}(k-1)x(n-k-M_b+1)+\boldsymbol{\varepsilon}(n)]^H\}$ and in the consequent estimate of $\boldsymbol{\nu}(k-1)$ obtained from the eigen analysis of $\hat{\mathbf{P}}_{\mathbf{b}}$ as $\hat{\boldsymbol{\nu}}(k-1) = \hat{\mathbf{u}}/\sqrt{\hat{\lambda}}$. Using assumptions A1 – A3

$$\hat{\mathbf{P}}_{\mathbf{b}} = \mathbf{h}(k-1)\mathbf{h}^H(k-1) + \mathbf{J}_{\boldsymbol{\gamma}} + \mathbf{J}_v + \mathbf{J}_{\mathrm{err}} + \mathbf{J}_{\mathrm{cross}} + \mathbf{J}_{\mathbf{h}(k-1)} \tag{33}$$

where

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\gamma}} \triangleq& \mathbf{B}^{0}_{M_b}\Gamma_{M_b+1,k}\Gamma^H_{M_b+1,k}\mathbf{B}^{0H}_{M_b} \\ \mathbf{J}_v \triangleq& \sigma_v^2\mathbf{B}^{0}_{M_b}\mathcal{A}_{M_b+1,k}\mathcal{A}^H_{M_b+1,k}\mathbf{B}^{0H}_{M_b} \\ \mathbf{J}_{\mathrm{err}} \triangleq& \mathbf{B}^{\mathrm{err}}_{M_b}(\Pi_{M_b+1}\Pi^H_{M_b+1}+\sigma_v^2\mathcal{A}_{M_b+1,k}\mathcal{A}^H_{M_b+1,k})\mathbf{B}^{\mathrm{err}H}_{M_b} \\ \mathbf{J}_{\mathrm{cross}} \triangleq& \mathbf{B}^{0}_{M_b}\Gamma_{M_b+1,k}\mathbf{I}_{\mathrm{right}}\Pi^H_{M_b+1}\mathbf{B}^{\mathrm{err}H}_{M_b} + \mathbf{B}^{\mathrm{err}}_{M_b}\Pi_{M_b+1}\mathbf{I}^H_{\mathrm{right}}\Gamma^H_{M_b+1,k}\mathbf{B}^{0H}_{M_b} \\ &+\sigma_v^2\mathbf{B}^{0}_{M_b}\mathcal{A}_{M_b+1,k}\mathcal{A}^H_{M_b+1,k}\mathbf{B}^{\mathrm{err}H}_{M_b} + \sigma_v^2\mathbf{B}^{\mathrm{err}}_{M_b}\mathcal{A}_{M_b+1,k}\mathcal{A}^H_{M_b+1,k}\mathbf{B}^{0H}_{M_b} \\ \mathbf{J}_{\mathbf{h}(k-1)} \triangleq& \mathbf{h}(k-1)\mathbf{j}^H_{M_b}\Gamma^H_{M_b+1,k}\mathbf{B}^{0H}_{M_b} + \mathbf{B}^{0}_{M_b}\Gamma_{M_b+1,k}\mathbf{j}_{M_b}\mathbf{h}^H(k-1) \\ &+\mathbf{h}(k-1)\mathbf{j}^H_{M_b+k}\Pi^H_{M_b+1}\mathbf{B}^{\mathrm{err}H}_{M_b} + \mathbf{B}^{\mathrm{err}}_{M_b}\Pi_{M_b+1}\mathbf{j}_{M_b+k}\mathbf{h}^H(k-1)\end{aligned}$$

where $\mathbf{I}_{\mathrm{right}} = [\mathbf{0}_{(L_h+M+M_b-1)\times k}\ \mathbf{I}_{L_h+M+M_b-1}]$, the (L_h+M+M_b-1)-vector $\mathbf{j}_{M_b} = [0\ldots 0\,1\,0\ldots 0]^T$, where the 1 is in the M_bth position and the $(L_h+M+k+M_b-1)$-vector $\mathbf{j}_{M_b+k} = [0\ldots 0\ 1\ 0\ldots 0]^T$, with the 1 in the (M_b+k)th position. Straightforward proof of (33) is provided in the Appendix.

Defining $\boldsymbol{\nu}_e(k-1) = \hat{\boldsymbol{\nu}}(k-1) - \boldsymbol{\nu}(k-1)$ and using (31), the FB-PEF output is

$$\begin{aligned}\hat{x}(n-k-M_b+1) =& \hat{\boldsymbol{\nu}}(k-1)\mathbf{b}_{M_b}(n) \\ =& x(n-k-M_b+1) + \hat{\boldsymbol{\nu}}(k-1)\boldsymbol{\varepsilon}(n) \\ &+ \boldsymbol{\nu}_e(k-1)\mathbf{h}(k-1)x(n-k-M_b+1)\end{aligned} \tag{34}$$

The minimum achievable MSE $J_{\min} = \mathrm{E}\{[\hat{x}(n) - x(n)]^2\}$ of the symbol estimate is given in Theorem 4.

Theorem 4 *A lower bound for the MSE of the blind equalizer FB-PEF output is*

$$J_{min} = \hat{\boldsymbol{\nu}}(k-1)[\mathbf{J}_{\gamma} + \mathbf{J}_v + \mathbf{J}_{err} + \mathbf{J}_{cross}]\hat{\boldsymbol{\nu}}^H(k-1) + J_{\boldsymbol{\nu}_e} \tag{35}$$

where

$$\begin{aligned}J_{\boldsymbol{\nu}_e} \triangleq& \boldsymbol{\nu}_e(k-1)\mathbf{h}(k-1)\mathbf{h}^H(k-1)\boldsymbol{\nu}_e^H(k-1) \\ &+ \boldsymbol{\nu}_e(k-1)\mathbf{h}(k-1)\mathbf{j}_{M_b}^H\Gamma_{M_b+1,k}^H\mathbf{B}_{M_b}^{0H}\hat{\boldsymbol{\nu}}^H(k-1) \\ &+ \hat{\boldsymbol{\nu}}(k-1)\mathbf{B}_{M_b}^0\Gamma_{M_b+1,k}\mathbf{j}_{M_b}\mathbf{h}^H(k-1)\boldsymbol{\nu}_e^H(k-1) \\ &+ \boldsymbol{\nu}_e(k-1)\mathbf{h}(k-1)\mathbf{j}_{M_b+k}^H\Pi_{M_b+1}^H\mathbf{B}_{M_b}^{errH}\hat{\boldsymbol{\nu}}^H(k-1) \\ &+ \hat{\boldsymbol{\nu}}(k-1)\mathbf{B}_{M_b}^{err}\Pi_{M_b+1}\mathbf{j}_{M_b+k}\mathbf{h}^H(k-1)\boldsymbol{\nu}_e^H(k-1)\end{aligned} \tag{36}$$

The result can be obtained with a slight modification of the proof of (33), by direct substitution from (32) to (34). Component $\mathbf{J}_{\gamma}$ is due to the residual ISI and $\mathbf{J}_v$ to the colored noise, akin to the $\mathbf{J}_{\gamma}$ and $\mathbf{J}_v$ in Theorem 3. Component $\mathbf{J}_{\mathrm{err}}$ results from the error in the backward PEF coefficients and $\mathbf{J}_{\mathrm{cross}}$ involves the correlated components. Similarly to the FF-PEF, $J_{\boldsymbol{\nu}_e}$ is due to the error in estimating $\boldsymbol{\nu}(k-1)$, affecting also the whole expression of $J_{\min}$ in $\hat{\boldsymbol{\nu}}(k-1)$.

5 Analytical and simulated results

In this section, we evaluate the analytical expression for the residual ISI plus noise at the k-step PEF output. We also compute a lower bound for the MSE of the FF-PEF and FB-PEF blind equalizers and use simulations to verify the results. The magnitude of the employed communication channels 1 and 2 are shown in Fig. 2. They are both empirically measured microwave channels obtained from the web-site *http://spib.rice.edu/spib/microwave.html.* The source sequence is white 4-QAM and the additive noise is white and Gaussian. Unless otherwise stated, the forward prediction order $M = L_h - 1$ and backward $M_b = k - 1$. The effect of prediction order mismatch has been considered in more detail e.g. in [1, 7]. For the Monte-Carlo simulated results, a sample size of N transmitted symbols is used to estimate the correlation matrices. An independent test set of 5000 symbols is used at each evaluated point and the results are averaged over 100 Monte-Carlo runs. The SNR is defined as SNR $:= \mathrm{E}\{\Sigma_{j=0}^{P-1}|y_j'(n)|^2/\Sigma_{k=0}^{P-1}|v_k(n)|^2\}$, where $y_j'(n)$ is the signal component of $y_j(n)$

5.1 Error at the output of the truncated channel

The output of a k-step forward PEF was shown in Theorem 1 to be equal to an output of a truncated channel. The trace of the covariance matrix of the output error (10) is given as a function of the SNR in the top plot of Fig. 3 for channel 1, $k = 1$. The overall error consists of the residual ISI component $\sigma_v^4 \mathbf{w}_{M,k} \mathcal{H}_M^{\dagger H} \mathcal{H}_M^{\dagger} \mathbf{w}_{M,k}^H$ and the colored noise component $\sigma_v^2 \mathbf{A}_{M,k} \mathbf{A}_{M,k}^H$, which are also plotted separately. It can be seen that the majority of the error is due to the colored noise.

A similar graph for channel 2 is shown in the bottom plot of Fig. 3. It is interesting to note that although the colored noise is stronger than for channel 1, the dominant error source is the residual ISI component. It would be possible to make the residual ISI go to zero by choosing the k to equal the channel length. In the context of blind equalization, however, we are interested in a k that is in the proximity of the energy maximum of the channel. A reasonable choice for channel 2 is $k = 4$, as will be seen in section 5.2.

The residual ISI for channel 2 is large compared to the error for channel 1. The same phenomenon can be observed for other channels as well that have the first channel coefficient vector $\mathbf{h}(0)$ small in magnitude. We demonstrate this by evaluating the error (10) for random channels of length $8T$. The uncorrelated channel coefficients are randomly drawn from a complex Gaussian distribution with zero mean and unit variance, resulting in a Rayleigh distribution for the magnitude of each tap. The total error and its components are plotted in Fig. 4 as a function of the relative magnitude $M[\mathbf{h}(0)]$ of $\mathbf{h}(0)$ to the magnitude of the

Table 1: Results for randomly weighted two-ray raised-cosine channels.

$M[\mathbf{h}(0)]$	5%	10%	15%	20%	25%
PEF error	0.0035	0.0031	0.0026	0.0017	0.0011
FF-PEF MSE	0.4749	0.4398	0.3726	0.1878	0.0312
FB-PEF MSE	0.1667	0.1916	0.2363	0.1740	0.0327

entire channel, defined as

$$M[\mathbf{h}(k)] \triangleq \frac{|h_0(k)| + |h_1(k)|}{\sum_{i=0}^{L_h-1}\{|h_0(i)| + |h_1(i)|\}} \times 100\% \tag{37}$$

At each evaluated point, $M[\mathbf{h}(0)] = 1\%, \ldots, 25\%$, the results are averaged over 1000 realizations of the random channel. The SNR is 30 dB and $k = 4$. It is clear from Fig. 4 that the residual ISI — and to a lesser extent also the colored noise — is dependent on $M[\mathbf{h}(0)]$.

The random channels employed in the previous experiment may have an infinite bandwidth. For a more realistic situation, we now choose a two-ray raised cosine channel as in [5]. The channel spans eight symbol intervals and has the form $h(t) = g_1 r_c(t - 0.75T) + g_2 r_c(t - 2.25T)$, where $r_c(t)$ is a raised cosine pulse with a roll-off factor 0.35. The gains g_1 and g_2 are randomly selected from a complex Gaussian distribution and $h(t)$ is sampled at the rate of $T/2$ to obtain the discrete time channel. The first row of Table 1 (PEF error) presents the total error obtained for five different values of $M[\mathbf{h}(0)]$, averaged over 1000 realizations of the weighted raised cosine channel. The same dependence of the error on $M[\mathbf{h}(0)]$ as with the random tap channels is visible.

The average condition number (the ratio of the largest eigen value to the smallest) of $\mathcal{H}_M$ is also plotted in Fig. 4 to show that the large error at low values of $M[\mathbf{h}(0)]$ may be due to the violation of assumption A4, i.e. $\mathcal{H}_M$ is rank deficient. Fig. 5 shows the results of the same experiment as in Fig. 4 as a function of $M[\mathbf{h}(1)]$. There is no obvious dependence between $M[\mathbf{h}(1)]$ and the error and we can conclude that the critical factor is indeed the magnitude of the first channel coefficient. Multi-step prediction based equalization is intended to be used when $M[\mathbf{h}(0)]$ is small [10]. It seems controversial that under such conditions, the error at the output of the multi-step PEF becomes large. We suspect that with small $M[\mathbf{h}(0)]$, the violation of the *length and zero condition* of assumption A4 is approached.

5.2 Performance of the blind equalizers

Computing the predictor coefficients from (6) with the noise component $\sigma_v^2\mathbf{I}_{MP}$ present results in biased estimators [4]. In Fig. 6 we study how subtracting the noise component from (6) affects the equalization performance. We assume exact knowledge of σ_v^2 to obtain the unbiased equalizer. The MSE is plotted as a function of the prediction order M using channel 1 and zero-delay equalization, so that the FF-PEF and FB-PEF are identical. The SNR is 15 dB and sample size $N = 2000$. It is clear from the figure that the sensitivity of the unbiased approach would require accurate channel order estimation, as opposed to the biased approach. Potentially leading to a lower MSE and lacking the complexity of estimating the noise statistics, the biased estimators seem more practical for applications.

We evaluate the performance of the FF-PEF and FB-PEF blind equalizers considering the results of section 5.1. The MSE of the equalizers for channel 1 and $k = 1$ is shown in the top plot of Fig. 7 as a function of the SNR. A large sample size $N = 25000$ is used for illustrative purposes. The analytical results are on par with the simulated ones at low SNRs. At high SNRs the finite sample size causes an error floor in the simulation results. The FB-PEF is a serial system of two predictors and thus suffers from propagation of errors in the sample statistics. This is especially evident at high SNRs, where the FF-PEF achieves a performance significantly closer to its asymptotic MSE than the FB-PEF. The bottom plot of Fig. 7 presents the simulated convergence curves of the two equalizers as a function of the sample size N for SNR 15 dB. The SNR is now relatively low, so that the errors in the sample statistics are not dominant, as can be verified from the top plot. Therefore, the simulated results are consistent with the analytical ones, in that the FB-PEF MSE is slightly lower than the FF-PEF MSE.

In section 5.1, it was observed that the residual ISI component at the output of the truncated channel is large when the first channel coefficient is small in magnitude. Not surprisingly, the property carries over to the performance of the blind equalizers, as well. The MSE of the FF-PEF and the FB-PEF for channel 2 is shown in Fig. 8. Despite the fact that a k that produces the smallest MSE is chosen, both equalizers perform very poorly even for high SNRs. The analytical MSE of the FF-PEF and FB-PEF is plotted in Fig. 9

as a function of $M[\mathbf{h}(0)]$ and $M[\mathbf{h}(1)]$ for random tap channels. The same procedure as in section 5.1 was used in the Monte-Carlo simulations, except that a k resulting in the smallest MSE was chosen for each realization of the random channel. Similarly, the last two rows of Table 1 present the $M[\mathbf{h}(0)]$ results when the randomly weighted two-ray raised cosine channel defined in section 5.1 was used. From Fig. 9 and Table 1, the performance of both equalizers can be seen to depend on $M[\mathbf{h}(0)]$. Interestingly for the FB-PEF in Table 1, the MSE seems to experience a slight increase with $M[\mathbf{h}(0)]$ first, until larger values of $M[\mathbf{h}(0)]$. In general, the FB-PEF outperforms the FF-PEF.

The choice of k in the multi-step prediction controls the delay of the equalizers. The analytical MSE of the blind equalizers for channel 1 is shown in Fig. 10 for $k = 1, \dots, 4$. The FB-PEF is very sensitive to a proper choice of k, which is a significant drawback in practical systems. The FF-PEF performs consistently with all the considered values of k, although it does not beat the FB-PEF when optimum values of k are used.

6 Conclusions

The performance of two previously suggested prediction-based blind equalizers — and the limitations of a model they are based on — were analyzed. The assumption of zero noise that is used in deriving the algorithms is not valid for real communication channels, which introduces errors at the output of the equalizers. An analytical lower bound for the MSE of the equalizers in the presence of noise was derived. The MSE was evaluated for empirically measured and random channels. Simulations were used to verify the results. In general, the FB-PEF performs better than the FF-PEF. However, with high SNRs or non-ideal choice of k, the FF-PEF has the advantage.

The rationale behind using multi-step prediction in blind equalization lies in the need to control the delay of the equalizer. It is generally acknowledged that the delay is especially needed when the channel energy is not concentrated in the beginning of the impulse response. Our results suggest that the equalizer delay alone is not sufficient to ensure a good performance; the magnitude of the first channel coefficient should also be sufficiently large. Further analysis on the effect of the $\mathbf{h}(0)$ magnitude on the estimation error is needed.

How other second-order statistics based equalizers using assumption A4 can deal with the situation is another interesting research topic.

A Appendix

Proof of Theorem 1: When the forward prediction-error filter is optimum in the mean-square sense, Wiener filtering theory states that the filter input and the output error are orthogonal [6]. In the context of prediction-error filtering, the orthogonality condition can be written as

$$\mathrm{E}\{\mathbf{y}_M(n-k)\mathbf{f}_{M,k}^H(n)\} = \mathbf{0}_{MP\times P} \tag{38}$$

We extend the proof provided in [4] by considering the additive noise as well. Substituting for $\mathbf{f}_{M,k}(n)$ from (8) results in

$$\mathrm{E}\{\mathbf{y}_M(n-k)\mathbf{y}_{M+k}^H(n)\}\mathbf{A}_{M,k}^H = \mathbf{0}_{MP\times P} \tag{39}$$

Using (4), we get

$$\mathrm{E}\{[\mathcal{H}_M\mathbf{x}_{L_h+M-1}(n-k) + \mathbf{v}_M(n-k)][\mathcal{H}_{M+k}\mathbf{x}_{L_h+M+k-1}(n) + \mathbf{v}_{M+k}(n)]^H\}\mathbf{A}_{M,k}^H = \mathbf{0}_{MP\times P} \tag{40}$$

An $i\times j$ matrix of zeros is denoted $\mathbf{0}_{i\times j}$ and an $i\times i$ unity matrix $\mathbf{I}_i$. The input sequence and the additive noise are assumed uncorrelated, zero-mean and white, so that

$$\mathcal{H}_M[\mathbf{0}_{(L_h+M-1)\times k}\ \mathbf{I}_{L_h+M-1}]\mathcal{H}_{M+k}^H\mathbf{A}_{M,k}^H + \sigma_v^2[\mathbf{0}_{MP\times Pk}\ \mathbf{I}_{MP}]\mathbf{A}_{M,k}^H = \mathbf{0}_{MP\times P} \tag{41}$$

Premultiplying by $\mathcal{H}_M^\dagger$ results in

$$[\mathbf{0}_{(L_h+M-1)\times k}\ \mathbf{I}_{L_h+M-1}]\mathcal{H}_{M+k}^H\mathbf{A}_{M,k}^H + \sigma_v^2\mathcal{H}_M^\dagger[\mathbf{0}_{MP\times Pk}\ \mathbf{I}_{MP}]\mathbf{A}_{M,k}^H = \mathbf{0}_{(L_h+M-1)\times P} \tag{42}$$

$$[\mathbf{0}_{(L_h+M-1)\times k}\ \mathbf{I}_{L_h+M-1}]\mathcal{H}_{M+k}^H\mathbf{A}_{M,k}^H - \sigma_v^2\mathcal{H}_M^\dagger\mathbf{w}_{M,k}^H = \mathbf{0}_{(L_h+M-1)\times P} \tag{43}$$

where $\mathcal{H}_M^\dagger\mathcal{H}_M = \mathbf{I}$, when $MP \geq L_h+M-1$, i.e. $\mathcal{H}_M$ has more rows than columns. Let us denote

$$\mathcal{H}_{M+k}^H \triangleq \begin{bmatrix}\mathcal{H}^k \\ \mathcal{H}^f\end{bmatrix} \tag{44}$$

where $\mathcal{H}^k$ contains the first k rows of $\mathcal{H}_{M+k}^H$ (we assume that $k \leq L_h - 1$),

$$\mathcal{H}^k \triangleq \begin{bmatrix} \mathbf{h}^H(0) & \mathbf{0}_{(M+k-1)\times P} & & & \\ \mathbf{h}^H(1) & \mathbf{h}^H(0) & \mathbf{0}_{(M+k-2)\times P} & & \\ \vdots & \ddots & & & \\ \mathbf{h}^H(k-1) & \dots & & \mathbf{h}^H(0) & \mathbf{0}_{MP\times P} \end{bmatrix} \tag{45}$$

and $\mathcal{H}^f$ contains the remaining rows. The effect of the matrix multiplier $[\mathbf{0}_{(L_h+M-1)\times k}\ \mathbf{I}_{L_h+M-1}]$ is to remove $\mathcal{H}^k$ from $\mathcal{H}_{M+k}^H$. Consequently, we can write (43) as

$$\mathcal{H}^f \mathbf{A}_{M,k}^H = \sigma_v^2 \mathcal{H}_M^\dagger \mathbf{w}_{M,k}^H \tag{46}$$

Realizing that $\mathbf{A}_{M,k} = [\mathbf{I}_P\ \mathbf{0}_{P\times P(k-1)}\ \ -\mathbf{w}_{M,k}]$, inserting $\mathcal{H}^k$ on top of $\mathcal{H}^f$ in (46) results in

$$\begin{bmatrix} \mathcal{H}^k \\ \mathcal{H}^f \end{bmatrix} \mathbf{A}_{M,k}^H = \begin{bmatrix} \mathbf{h}^H(0) \\ \vdots \\ \mathbf{h}^H(k-1) \\ \sigma_v^2 \mathcal{H}_M^\dagger \mathbf{w}_{M,k}^H \end{bmatrix} \tag{47}$$

which equals

$$\mathcal{H}_{M+k}^H \mathbf{A}_{M,k}^H = \begin{bmatrix} \mathbf{h}^H(0) \\ \vdots \\ \mathbf{h}^H(k-1) \\ \sigma_v^2 \mathcal{H}_M^\dagger \mathbf{w}_{M,k}^H \end{bmatrix} \tag{48}$$

and further

$$\mathbf{A}_{M,k} \mathcal{H}_{M+k} = [\mathbf{h}(0), \dots, \mathbf{h}(k-1), \sigma_v^2 \mathbf{w}_{M,k} \mathcal{H}_M^{\dagger H}] \tag{49}$$

Postmultiplying by $\mathbf{x}_{L_h+M+k-1}(n)$ and adding a noise term on both sides, we get

$$\begin{aligned} &\mathbf{A}_{M,k} \mathcal{H}_{M+k} \mathbf{x}_{L_h+M+k-1}(n) + \mathbf{A}_{M,k} \mathbf{v}_{M+k}(n) \\ &= [\mathbf{h}(0), \dots, \mathbf{h}(k-1), \sigma_v^2 \mathbf{w}_{M,k} \mathcal{H}_M^{\dagger H}] \mathbf{x}_{L_h+M+k-1}(n) + \mathbf{A}_{M,k} \mathbf{v}_{M+k}(n) \end{aligned} \tag{50}$$

and

$$\mathbf{f}_{M,k}(n) = [\mathbf{h}(0), \dots, \mathbf{h}^H(k-1), \sigma_v^2 \mathbf{w}_{M,k} \mathcal{H}_M^{\dagger H}] \mathbf{x}_{L_h+M+k-1}(n) + \mathbf{A}_{M,k} \mathbf{v}_{M+k}(n) \tag{51}$$

$$\mathbf{f}_{M,k}(n) = \sum_{j=0}^{k-1} \mathbf{h}(j) x(n-j) + \sigma_v^2 \mathbf{w}_{M,k} \mathcal{H}_M^{\dagger H} \mathbf{x}_{L_h+M-1}(n-k) + \mathbf{A}_{M,k} \mathbf{v}_{M+k}(n) \tag{52}$$

which is the desired result. $\square$

Proof of equation (33): The backward prediction error is given by (31) and (32) as $\mathbf{b}_{M_b}(n) = \mathbf{h}(k-1)x(n-k-M_b+1) + \boldsymbol{\varepsilon}(n)$. When $\mathbf{P_b} = \mathrm{E}\{\mathbf{b}_{M_b}(n)\mathbf{b}_{M_b}^H(n)\}$, we can write

$$
\begin{aligned}
\mathbf{P_b} =& \mathbf{h}(k-1)\mathbf{h}^H(k-1) + \mathbf{B}_{M_b}^0 \Gamma_{M_b+1,k}\Gamma_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} \\
&+\sigma_v^2 \mathbf{B}_{M_b}^0 \mathcal{A}_{M_b+1,k}\mathcal{A}_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} + \mathbf{B}_{M_b}^{\mathrm{err}} \mathrm{E}\{\tilde{\mathbf{y}}_{M_b+1}(n)\tilde{\mathbf{y}}_{M_b+1}^H(n)\}\mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{h}(k-1)\mathrm{E}\{x(n-k-M_b+1)\mathbf{x}_{L_h+M+M_b-1}^H(n-k)\}\Gamma_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} \\
&+\mathbf{B}_{M_b}^0 \Gamma_{M_b+1,k}\mathrm{E}\{\mathbf{x}_{L_h+M+M_b-1}(n-k)x(n-k-M_b+1)\}\mathbf{h}^H(k-1) \\
&+\mathbf{h}(k-1)\mathrm{E}\{x(n-k-M_b+1)\tilde{\mathbf{y}}_{M_b+1}^H(n)\}\mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}\mathrm{E}\{\tilde{\mathbf{y}}_{M_b+1}(n)x(n-k-M_b+1)\}\mathbf{h}^H(k-1) \\
&+\mathbf{B}_{M_b}^0 \Gamma_{M_b+1,k}\mathrm{E}\{\mathbf{x}_{L_h+M+M_b-1}(n-k)\tilde{\mathbf{y}}_{M_b+1}^H(n)\}\mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}\mathrm{E}\{\tilde{\mathbf{y}}_{M_b+1}(n)\mathbf{x}_{L_h+M+M_b-1}^H(n-k)\}\Gamma_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} \\
&+\mathbf{B}_{M_b}^0 \mathcal{A}_{M_b+1,k}\mathrm{E}\{\mathbf{v}_{M+k+M_b}(n)\tilde{\mathbf{y}}_{M_b+1}^H(n)\}\mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}\mathrm{E}\{\tilde{\mathbf{y}}_{M_b+1}(n)\mathbf{v}_{M+k+M_b}^H(n)\}\mathcal{A}_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H}
\end{aligned}
\tag{53}
$$

where we have used assumptions A1 — A3. Substituting for $\tilde{\mathbf{y}}_{M_b+1}(n)$ from (22) and using assumptions A1 — A3 again, we get

$$
\begin{aligned}
\mathbf{P_b} =& \mathbf{h}(k-1)\mathbf{h}^H(k-1) \\
&+\mathbf{B}_{M_b}^0 \Gamma_{M_b+1,k}\Gamma_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} + \sigma_v^2 \mathbf{B}_{M_b}^0 \mathcal{A}_{M_b+1,k}\mathcal{A}_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}(\Pi_{M_b+1}\Pi_{M_b+1}^H + \sigma_v^2 \mathcal{A}_{M_b+1,k}\mathcal{A}_{M_b+1,k}^H)\mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{h}(k-1)\mathbf{j}_{M_b}^H \Gamma_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} + \mathbf{B}_{M_b}^0 \Gamma_{M_b+1,k}\mathbf{j}_{M_b}\mathbf{h}^H(k-1) \\
&+\mathbf{h}(k-1)\mathrm{E}\{x(n-k-M_b+1)\mathbf{x}_{L_h+M+k+M_b-1}^H(n)\}\Pi_{M_b+1}^H \mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}\Pi_{M_b+1}\mathrm{E}\{\mathbf{x}_{L_h+M+k+M_b-1}(n)x(n-k-M_b+1)\}\mathbf{h}^H(k-1) \\
&+\mathbf{B}_{M_b}^0 \Gamma_{M_b+1,k}\mathrm{E}\{\mathbf{x}_{L_h+M+M_b-1}(n-k)\mathbf{x}_{L_h+M+k+M_b-1}^H(n)\}\Pi_{M_b+1}^H \mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}\Pi_{M_b+1}\mathrm{E}\{\mathbf{x}_{L_h+M+k+M_b-1}(n)\mathbf{x}_{L_h+M+M_b-1}^H(n-k)\}\Gamma_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H} \\
&+\mathbf{B}_{M_b}^0 \mathcal{A}_{M_b+1,k}\mathrm{E}\{\mathbf{v}_{M+k+M_b}(n)\mathbf{v}_{M+k+M_b}^H(n)\}\mathcal{A}_{M_b+1,k}^H \mathbf{B}_{M_b}^{\mathrm{err}H} \\
&+\mathbf{B}_{M_b}^{\mathrm{err}}\mathcal{A}_{M_b+1,k}\mathrm{E}\{\mathbf{v}_{M+k+M_b}(n)\mathbf{v}_{M+k+M_b}^H(n)\}\mathcal{A}_{M_b+1,k}^H \mathbf{B}_{M_b}^{0H}
\end{aligned}
\tag{54}
$$

where $\mathbf{j}_{M_b}$ is defined as in (33). Using assumptions A1 and A2 for the last time leads to the desired result given in (33). $\square$

References

[1] K. Abed-Meraim, E. Moulines, and P. Loubaton. Prediction error method for second-order blind identification. *IEEE Tr. Signal Proc.*, 45(3):694–705, March 1997.

[2] Z. Ding and Y. Li. On channel identification based on second-order cyclic spectra. *IEEE Tr. Signal Proc.*, 42(5):1260–1264, May 1994.

[3] W. A. Gardner. A new method of channel identification. *IEEE Tr. Commun.*, 39(6):813–817, June 1991.

[4] D. Gesbert and P. Duhamel. Robust blind channel identification and equalization based on multi-step predictors. In *IEEE ICASSP*, pages 3621–3624, 1997.

[5] G. B. Giannakis and S. D. Halford. Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions. *IEEE Tr. Signal Proc.*, 45(9):2277–2292, September 1997.

[6] S. Haykin. *Adaptive Filter Theory.* Prentice-Hall, Englewood Cliffs, NJ, 3rd edition, 1996.

[7] J. Mannerkoski and D. P. Taylor. Blind equalization using least-squares lattice prediction. *IEEE Tr. Signal Proc.*, 47(3):630–640, March 1998.

[8] C. B. Papadias and D. T. M. Slock. Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction. *IEEE Tr. Signal Proc.*, 47(3):641–653, March 1999.

[9] D. T. Slock. Blind fractionally-spaced equalization, perfect reconstruction filter banks and multichannel linear prediction. In *IEEE ICASSP*, pages IV585–IV588, Adelaide, Australia, April 1994.

[10] D. T. Slock and C. B. Papadias. Further results on blind identification and equalization of multiple FIR channels. In *IEEE ICASSP*, pages 1964–1967, Detroit, MI, May 1995.

[11] L. Tong, G. Xu, and T. Kailath. Blind identification and equalization based on second-order statistics: A time-domain approach. *IEEE Tr. Info. Theory*, 40(2):340–349, March 1994.

[12] L. Tong, G. Xu, and T. Kailath. Blind identification based on second-order statistics: A frequency-domain approach. *IEEE Tr. Info. Theory*, 41(1):329–334, January 1995.

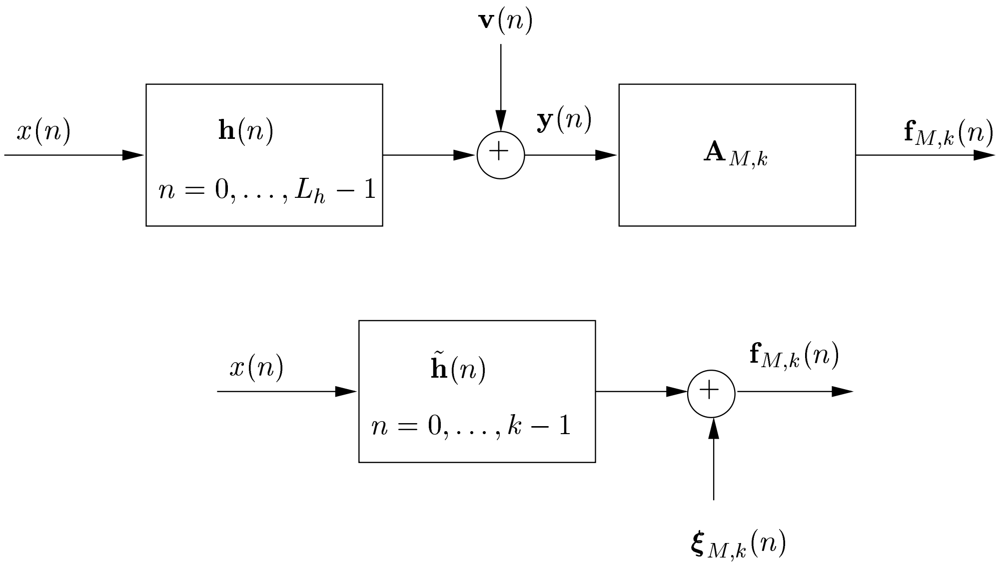

Figure 1: The output of a k-step PEF as an output of a truncated channel.

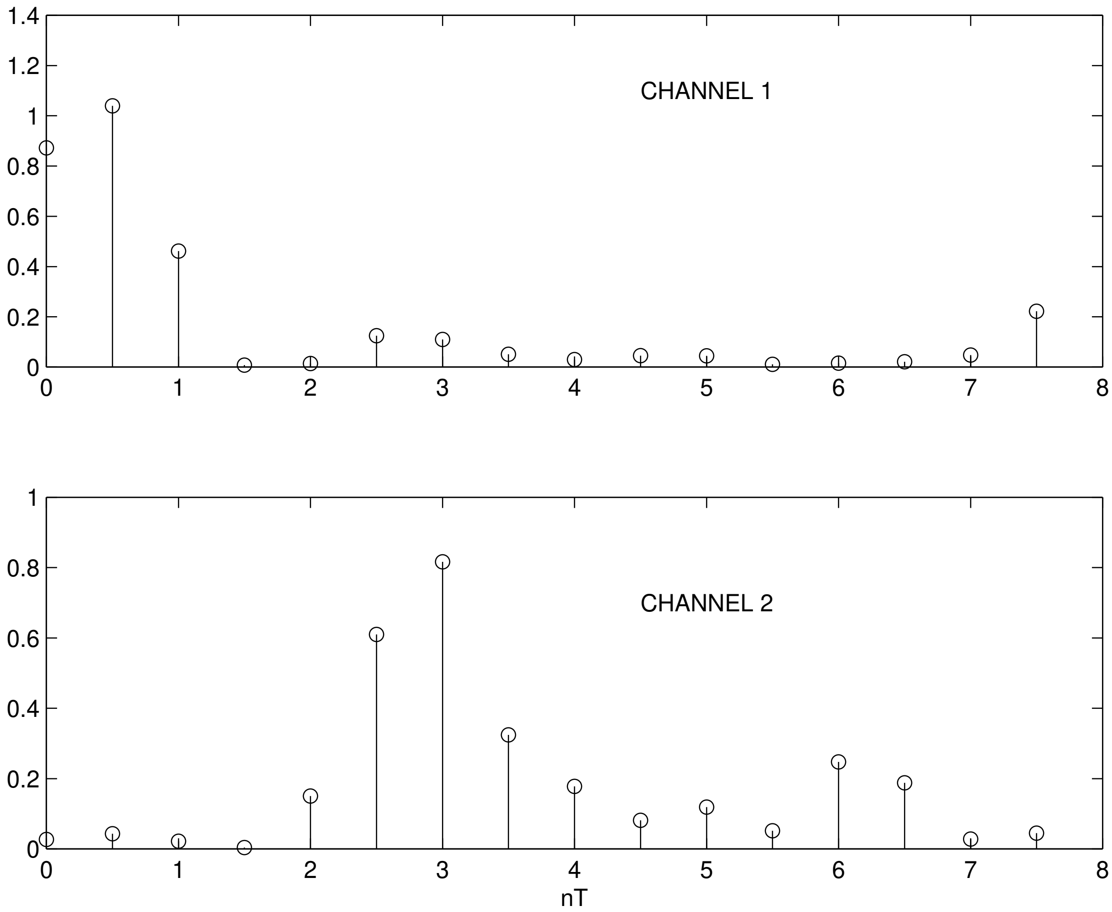

Figure 2: Magnitude of channels 1 and 2.

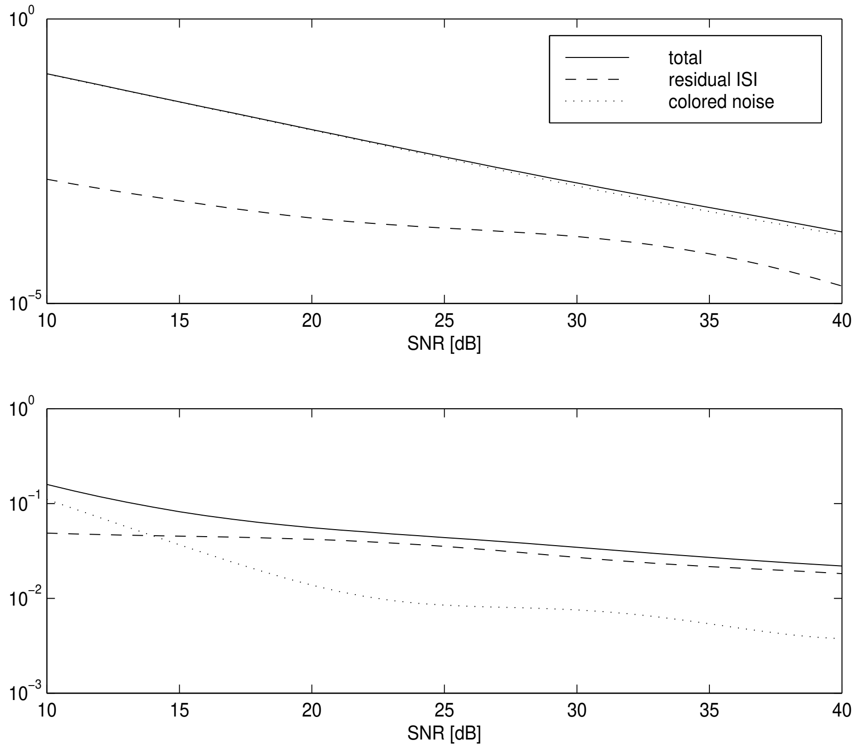

Figure 3: Error at the output of the partial channel, channels 1 ($k = 1$) and 2 ($k = 4$).

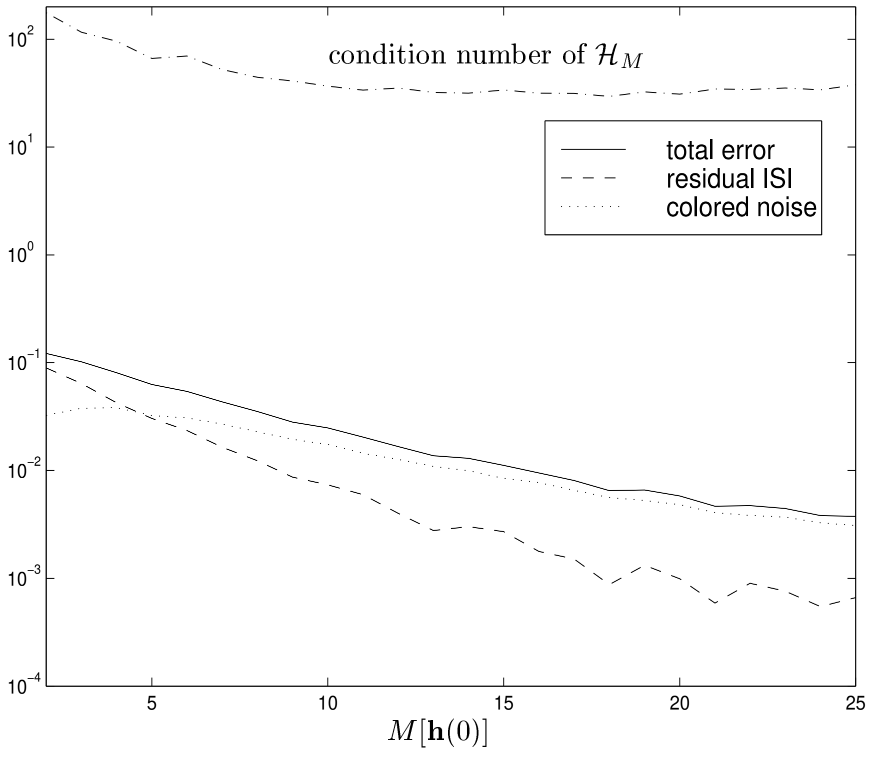

Figure 4: Error dependence on the $\mathbf{h}(0)$ magnitude (37), random channels, SNR of 30 dB.

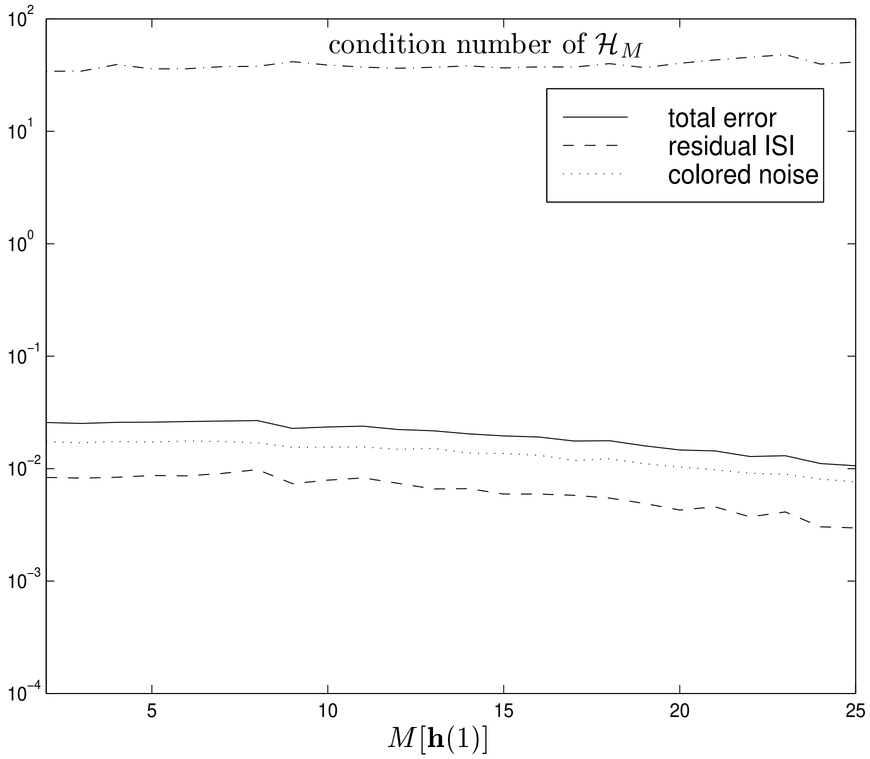

Figure 5: Error dependence on the $\mathbf{h}(1)$ magnitude (37), random channels, SNR of 30 dB.

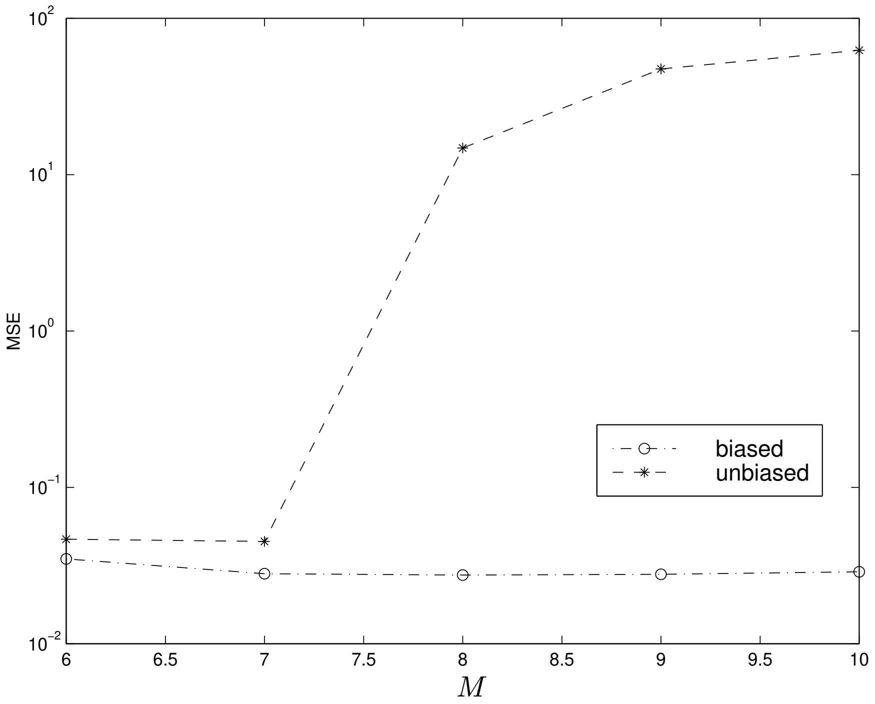

Figure 6: MSE of the zero-delay equalizer as a function of the predictor order M, channel 1, SNR 15 dB, sample size $N = 2000$.

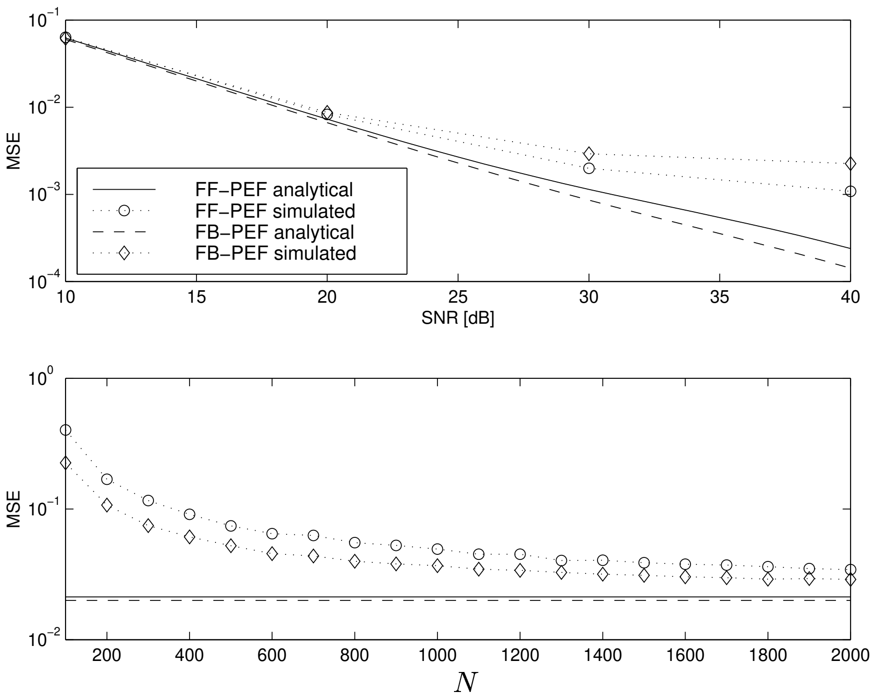

Figure 7: MSE of the FF-PEF and FB-PEF, channel 1, $k = 1$. Top; $N = 25000$. Bottom; SNR=15 dB

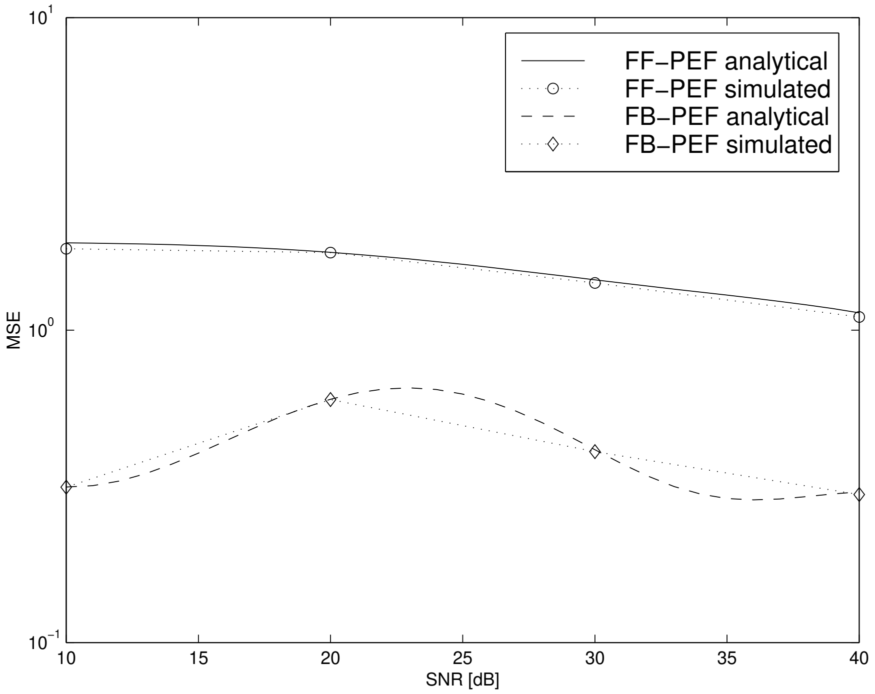

Figure 8: Minimum MSE of the FF-PEF ($k = 3$) and FB-PEF ($k = 4$), channel 2, $N = 25000$.

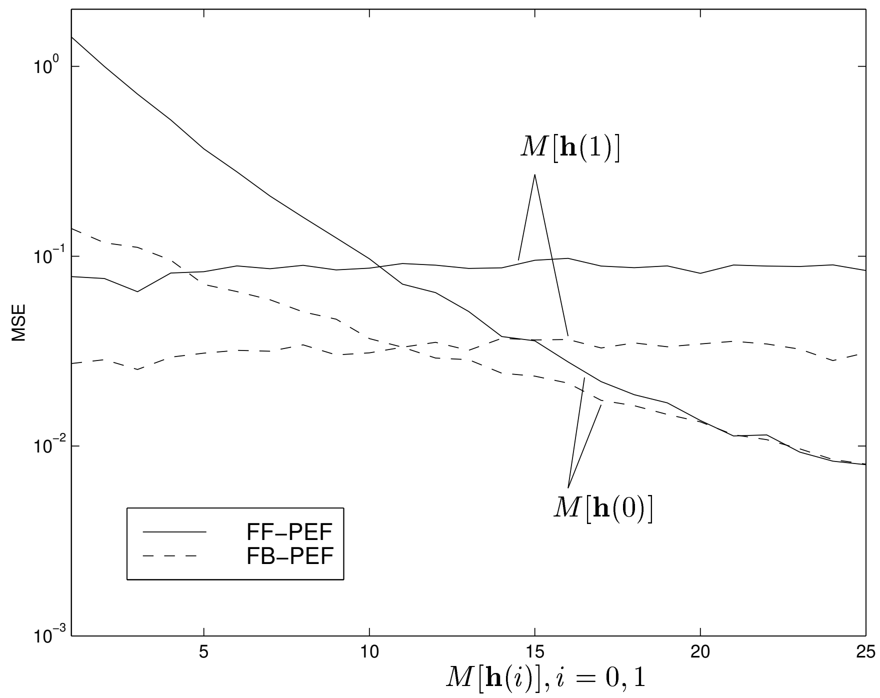

Figure 9: Analytical MSE of the FF-PEF and FB-PEF, random channels, SNR = 30 dB.

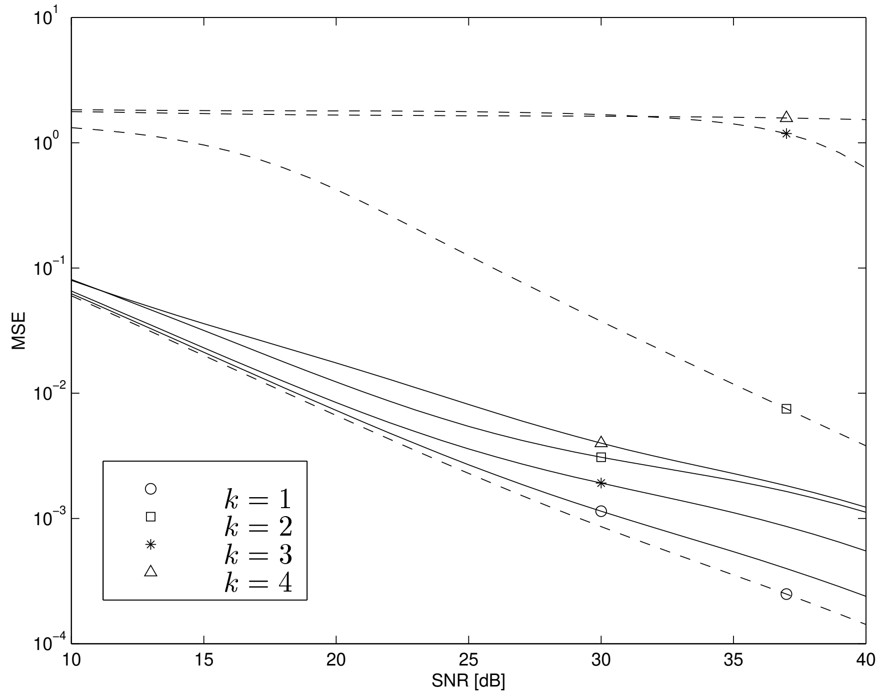

Figure 10: MSE of the blind equalizers for $k = 1, \ldots, 4$, channel 1. The solid line for the FF-PEF and the dashed line for the FB-PEF.