Signal Processing 82 (2002) 1233–1253
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Performance analysis and comparison of blind to non-blind least-squares equalization with respect to effective channel overmodeling

A.D. Beikos[a,*,1], A.P. Liavas[b,2]

[a]Department of Informatics and Telecommunications, University of Athens, Panepistimioupolis, Ilissia 15784, Athens, Greece
[b]Department of Mathematics, University of the Aegean, Karlovassi 83200, Samos, Greece

Received 2 July 2001

Abstract

The object of this work is the study of a direct blind equalization algorithm which appeared recently in the literature. It is a least-squares (LS) equalization method in the blind context, assuming a linear FIR communication channel and a linear equalizer. If channel order is known, blind LS equalizers can be constructed that entirely suppress intersymbol interference in noiseless signal transmission. In practice, though, channels may be comprised of a few "big" consecutive taps, which we call "significant part", surrounded by a lot of smaller leading and/or trailing "tail" terms. In such an environment, channel order is harder to define while the value used by the algorithm is critical to its performance. We carry out both theoretical analysis, making use of perturbation theory arguments, and simulations for the cases where channel order determination procedure has yielded an estimate greater than ("effective overmodeling") or equal to the order of the significant part. Our purpose is to compare the performance of blind LS algorithm with that of its non-blind counterpart. We conclude that (a) when channel does not possess leading tail terms, blind LS is robust to effective overmodeling, meaning that it behaves very much like non-blind LS, and (b) when leading tail terms are present, blind LS will generally not work satisfactorily in the effective overmodeling scenario. In either case, when the order of the significant part is identified correctly and the actual significant parts of subchannels are sufficiently diverse, the algorithm behaves well. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Blind channel equalization; SOS equalization; Least-squares equalization; Effective overmodeling; Perturbation theory

1. Introduction

Intersymbol interference (ISI) is one of the main factors obstructing reliable digital communications. It indicates the spreading in time of the transmitted symbols by the propagation medium and may be destructive at high enough symbol rates. In order to remove the corrupting effects of ISI, a special device is employed at the receiver called an *equalizer*.

* Corresponding author. Tel.: +30-10-7275328; fax: +30-10-7275214.
E-mail addresses: abeikos@di.uoa.gr (A.D. Beikos), liavas@aegean.gr (A.P. Liavas).

[1] Supported by PENED-99 "ISODIA" (99$E\Delta$83), a program of the Greek General Secretariat for Research and Technology.
[2] Supported by EPET-II, a program of the Greek General Secretariat for Research and Technology.

0165-1684/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0165-1684(02)00246-3

Conventional equalizer design techniques rely on the periodic transmission of data already known to the receiver, referred to as *training sequences*. A priori knowledge of such data allows for either the direct computation of the equalizer or the computation of the channel coefficients as a preliminary step before equalizer determination [9].

Training sequences clearly result in a waste of some of the channel's capacity. In order to allocate the maximum possible transmitting capacity to the users, *blind equalization* algorithms have received extensive attention. They do *not* make use of training sequences but rely solely on the output of the communication channel to achieve the desired equalization task.

The more traditional of these techniques sample the output of the channel at the baud rate. Therefore, they inevitably use higher-than second-order statistics (HOS) of the sampled symbols since only in this way is it possible to retrieve channel phase information. Nonetheless, this characteristic is responsible for two important disadvantages, namely: need for large sample sizes and potential capture in undesirable local minima.

In the pioneering work of [13] it is proved that channel phase information is present in channel output second-order statistics (SOS) if the input is observed through more than one sufficiently diverse channels. This amounts to oversampling the channel output and/or using several antennas at the receiver. These implementations are equivalent at a higher level of abstraction since they can all be modeled as a number of separate "virtual" channels driven by the same input. This setting has been named SIMO in the literature after "single-input multiple-output". Hence, the equalizers follow a "multiple-input single-output" (MISO) setting, in that they exploit each of the multiple virtual channel outputs to yield the equalized output for the actual channel.

SOS techniques alleviate the problems of HOS techniques and are therefore advantageous. Several algorithms have been developed with the SOS-SIMO setting in mind that either directly estimate the equalizer [2,3,8,11] or estimate the channel [1,7,14] at an initial step. What will ultimately determine the usefulness of these techniques is their robustness to real-world conditions which, more often than not, stray from theoretical assumptions. It is well known that most blind channel identification methods are very sensitive to channel overmodeling. Direct blind equalization algorithms were developed in the hope of overcoming this kind of sensitivity. It remains to be studied, however, if this is really the case.

One representative of the class of blind SOS algorithms that directly compute the equalizer is described in [10]. It is the blind analog of non-blind LS equalization. Briefly put, if the order of the channel is M and its output is oversampled by a factor of p, then an equalizer of order $L_{\mathrm{eq}} \geqslant M/(p-1)-1$ can be found that will *entirely* suppress the ISI introduced in the noiseless transmission of a white input sequence. In a typical implementation, a channel order determination procedure is employed to furnish an estimate that is subsequently fed into the algorithm.

Of particular interest is the case where a rather long channel of order M is incorrectly detected to be of order $L+1$ where $L+1<M$. Long channels appear in the context of microwave radio links [5,6] and they are usually comprised of a few big consecutive taps (with, probably, some small intermediate taps), called "significant part" throughout the paper and whose order we symbolize by L^*+1, while the rest of them are rather small leading and trailing terms and are referred to as "tails".

In this work, we attempt to examine the robustness properties of [10] in the 1-input/2-output channel context when the channel has a total order of M and the equalizer order is $L<M-1$, that is, shorter than required for perfect input reconstruction. Furthermore, we assume that all participating statistical quantities are known with infinite precision and the system is noiseless. Our aim is to unveil potential sensitivity of the algorithm to the channel-order mismatch. Statistical inaccuracies and additive channel noise are naturally expected to deteriorate system performance. It is interesting to remark at this point that the effect of long, small tail terms is equivalent with the presence of coloured noise in the system.

The rest of our paper is organized as follows: In Section 2, we present the channel model used and we review the algorithm developed in [10]. Section 3 is devoted to our contribution, i.e., the performance analysis of blind LS. In Sections 3.1 and 3.2, we decompose the "equalization" of the Mth-order channel by the Lth-order equalizer into an "ideal" part plus a perturbation that is sufficiently small. The ideal part conforms to the theoretical assumptions of [10]. The perturbation causes the violation of these assumptions. We perform a first-order perturbation analysis and we outline the key results of it. In Section 3.3 we use the results of the perturbation analysis and the findings of [4] to relate the performance of blind LS to that of non-blind LS. In particular, we compute the order of magnitude of the Euclidean distance of the combined responses (i.e., the cascade of channel + equalizer) furnished by each algorithm. This distance, along with the already known behaviour of non-blind LS [4], allows anticipation of performance of blind LS. The conclusions drawn boil down to the following points:

Trailing tails only: In the somewhat unrealistic case where the channel does *not* possess any leading small terms, we prove that blind LS algorithm is *robust* to "effective channel overmodeling". That is, when the estimated channel order is greater than the order of the *actual* significant part of the channel, in symbols when $L+1 > L^*+1$, blind LS algorithm will produce well-behaving equalizers, in the sense that their performance will be close to that of their non-blind counterparts. More specifically, if the *actual* significant parts of subchannels are sufficiently diverse, then blind LS equalizers for $L \geqslant L^*$ and delays $d = 0,\ldots,L+(L^*+1)$ will offer good equalization performance while for $d = L+(L^*+1)+1,\ldots,2L+1$ performance will be generally poor.

Leading and trailing tails: If the channel possesses small tail terms *surrounding* its significant part, the behaviour of blind LS algorithm is different depending on the accuracy of the channel order determination procedure: In the case of effective channel overmodeling, poor equalization performance should generally be expected for *every* delay. In the "exact order case", however, (i.e., when $L+1=L^*+1$) if the *actual* significant parts of subchannels are sufficiently diverse, then blind equalizer performance is satisfactory for *each* possible delay.

2. Least-squares equalization

2.1. Channel model

We adopt the baseband SIMO FIR channel model. The transmitted (scalar) sequence $s(n)$ is filtered by p FIR linear filters, yielding p (scalar) sequences $x_i(n)$ $(i = 1,\ldots,p)$ at the output of the communication system. Each FIR filter is of order M, i.e., it consists of $M+1$ taps, and will be called "subchannel" in the sequel. It proves handy to view the output of the system as the vector $\mathbf{x}(n) = [x_1(n)\cdots x_p(n)]^{\mathrm{T}}$. Now, let $\mathbf{h}_i$ be the p-component vector containing the ith tap of each subchannel. The output is then given by the convolution: $\mathbf{x}(n) = \sum_{i=0}^{M} \mathbf{h}_i s_{n-i}$. By stacking the $L+1$ most recent outputs, we construct the vector $\mathbf{X}_L(n) = [\mathbf{x}(n)^{\mathrm{T}}\cdots\mathbf{x}(n-L)^{\mathrm{T}}]^{\mathrm{T}}$. This vector can be expressed as: $\mathbf{X}_L(n) = \mathscr{T}_L(\mathbf{H}_M)\mathbf{s}_{L+M}(n)$ where

$$\mathscr{T}_L(\mathbf{H}_M) = \begin{bmatrix} \mathbf{h}_0 & \cdots\cdots & \mathbf{h}_M & & \\ & \ddots & & \ddots & \\ & & \mathbf{h}_0 & \cdots\cdots & \mathbf{h}_M \end{bmatrix}$$

is a $p(L+1)\times(L+M+1)$ generalized Sylvester matrix, $\mathbf{s}_{L+M}(n) = [s(n)\cdots s(n-L-M)]^{\mathrm{T}}$ is an $(L+M)$th order vector grouping past input samples and $\mathbf{H}_M = [\mathbf{h}(0)^{\mathrm{T}}\cdots\mathbf{h}(M)^{\mathrm{T}}]^{\mathrm{T}}$ is the vector grouping all subchannels taps. If $p(L+1) \geqslant L+M+1$ and the subchannels posses no common zeros, then matrix $\mathscr{T}_L(\mathbf{H}_M)$ is of *full-column rank*. These are the conditions which virtually all SOS-based blind identification/equalization algorithms build on and are commonly referred to as the *zero-forcing conditions*. The full-column rank property of $\mathscr{T}_L(\mathbf{H}_M)$ ensures that it is *left invertible*, i.e., it holds $\mathscr{T}_L^{\sharp}(\mathbf{H}_M)\mathscr{T}_L(\mathbf{H}_M) = \mathbf{I}$, where $\sharp$ denotes Moore–Penrose pseudoinversion and $\mathbf{I}$ the identity matrix. The left invertibility of the matrix is synonymous with the *equalizability* of the system.

2.2. Blind, non-blind MMSE equalizers

In [10], a novel algorithm for blind channel equalization using SOS is presented. Under the assumption that the zero-forcing conditions hold, it computes an equalizer $\mathbf{g}_{L,i}$ that minimizes the mean square error in the estimation of the input symbols. In other words, if the estimated symbol at the output of $\mathbf{g}_{L,i}$ is $z(n)=\mathbf{g}_{L,i}^{\mathrm{H}}\mathbf{X}_L(n)$ then $E\{\|z(n)-s(n-i)\|_2^2\}$ is minimized, in the presence of uncorrelated with the input signal, additive, white noise at the output of each subchannel. When noise is absent, the equalizer computed this way will exactly recover the input sequence $s(n-i)$. The desired Lth order, i-delay equalizer is found by solving the Wiener–Hopf equation:

$$E\{\mathbf{X}_L(n)\mathbf{X}_L^{\mathrm{H}}(n)\}\mathbf{g}_{L,i}=E\{\mathbf{X}_L(n)s_{n-i}^*\},$$

where $\{\cdot\}^*$ denotes complex conjugation and $\{\cdot\}^{\mathrm{H}}$ denotes Hermitian transposition. In the following, we will use the notation $\mathbf{R}_{L,0}(\mathbf{H}_M)\triangleq E\{\mathbf{X}_L(n)\mathbf{X}_L^{\mathrm{H}}(n)\}$. Assuming a white, unit-variance input sequence there hold the following in the noiseless case:

$$\mathbf{R}_{L,0}(\mathbf{H}_M)=\mathscr{T}_L(\mathbf{H}_M)\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M), E\{\mathbf{X}_L(n)s_{n-i}^*\}=\mathscr{T}_L(\mathbf{H}_M)(:,i+1)\triangleq\hat{\mathbf{H}}_i, \quad (1)$$

where in (1) we have made use of standard Matlab notation $\mathbf{A}(:,i+1)$ to denote the $(i+1)$st column of matrix $\mathbf{A}$, in this particular case $\mathbf{A}$ being $\mathscr{T}_L(\mathbf{H}_M)$.

Thus, the Wiener–Hopf equation transforms into

$$(\mathscr{T}_L(\mathbf{H}_M)\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))\mathbf{g}_{L,i}=\hat{\mathbf{H}}_i. \quad (2)$$

Since the output SOS can be computed without knowledge of the input, it is easily seen that what hinders the solution of (2) in the blind context is the *unknown* $\hat{\mathbf{H}}_i$. The main contribution of [10] lies into furnishing a method of acquiring $\hat{\mathbf{H}}_i$ without training sequences under the sole assumption that the zero-forcing conditions hold.

In the sequel, we present the steps employed to recover $\hat{\mathbf{H}}_i$. Let $\mathbf{J}^i$ be the Jordan matrix. This is a square matrix with zero entries except for the ith lower subdiagonal, where its entries equal to 1. For $i=0$, $\mathbf{J}^0=\mathbf{I}$. Let us define

$$\mathbf{R}_{L,i}(\mathbf{H}_M)\triangleq E\{\mathbf{X}_L(n+i)\mathbf{X}_L^{\mathrm{H}}(n)\}.$$

Then

$$\mathbf{R}_{L,i}(\mathbf{H}_M)=\mathscr{T}_L(\mathbf{H}_M)\mathbf{J}_{L+M+1}^i\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M),$$

where $\mathbf{J}_{L+M+1}^i$ is the respective $(L+M+1)\times(L+M+1)$ Jordan matrix.

In [10] it is proved that if we define

$$\mathbf{D}_i\triangleq\mathbf{R}_{L,i}(\mathbf{H}_M)\mathbf{R}_{L,0}^{\sharp}(\mathbf{H}_M)\mathbf{R}_{L,i}^{\mathrm{H}}(\mathbf{H}_M), \quad (3)$$

then

$$\Delta\mathbf{D}_i\triangleq\mathbf{D}_i-\mathbf{D}_{i+1}=\hat{\mathbf{H}}_i\hat{\mathbf{H}}_i^{\mathbf{H}},$$

that is, $\Delta\mathbf{D}_i$ is a rank-one non-negative definite matrix. Its non-zero eigenvalue is $\lambda_i=\|\hat{\mathbf{H}}_i\|_2^2$ and the respective eigenspace is spanned by any non-zero multiple of $\hat{\mathbf{H}}_i$. By means of the EVD, we can compute $\hat{\mathbf{H}}_i$ within a sign ambiguity, i.e., $\hat{\mathscr{H}}_i\triangleq\sqrt{\lambda_i}\mathbf{v}=\pm\hat{\mathbf{H}}_i$, where $\mathbf{v}$ a unit 2-norm vector in the range space of $\Delta\mathbf{D}_i$. Using $\hat{\mathscr{H}}_i$ to solve the Wiener–Hopf equation, we can identify the input also within a sign ambiguity, i.e., $z(n)=\pm s(n-i)$. This is the best we can hope for any algorithm of the blind SOS class.

3. Performance analysis of the *L*th-order blind LS equalizer

Recalling the zero-forcing conditions, we note that the equalizer must be of *at least* a minimum order, directly related to the subchannels order and number. Therefore, as a first step towards the implementation of the algorithm, a (sub-)*channel-order* determination procedure must be employed, in order to allow for the suitable choice of the *equalizer order*. However, it may happen that the order M of the subchannels of $\mathbf{H}_M$ is rather large, while our procedure comes up with an order estimation of $L+1$, where $L+1 < M$. In the sequel, we will consider the behaviour of the Lth-order equalizer, in the two subchannel case, i.e., when $p = 2$.

We note that for two Mth-order subchannels, the zero forcing conditions demand the use of an equalizer with order greater than or equal to $M-1$. However, since the estimation procedure has yielded the misleading output that the subchannels are of order $L+1$, we are justified to use an equalizer of order L, as this is the shortest possible prescribed by the zero forcing conditions for two subchannels of order $L+1$ each. Nevertheless, since in reality $L < M-1$ we unintentionally violate the zero forcing conditions and it is worthwhile to examine the behaviour of the blind LS algorithm in this scenario. In order to study the behaviour of the Lth-order blind LS algorithm when applied to the Mth-order channel $\mathbf{H}_M$, we find it convenient to introduce the following partition [4–6]:

$$\mathbf{H}_M = \mathbf{H}^z_{L+1} + \mathbf{D}^z_{L+1},$$

where

$$\mathbf{H}^z_{L+1} = \Big[\underbrace{\mathbf{0}^\mathrm{T}\cdots\mathbf{0}^\mathrm{T}}_{m_1}\underbrace{\mathbf{h}^\mathrm{T}_{m_1}\cdots\mathbf{h}^\mathrm{T}_{m_2}}_{L+2}\underbrace{\mathbf{0}^\mathrm{T}\cdots\mathbf{0}^\mathrm{T}}_{M-m_2}\Big]^\mathrm{T}, \qquad \mathbf{D}^z_{L+1} = \Big[\underbrace{\mathbf{h}^\mathrm{T}_0\cdots\mathbf{h}^\mathrm{T}_{m_1-1}}_{m1}\underbrace{\mathbf{0}^\mathrm{T}\cdots\mathbf{0}^\mathrm{T}}_{L+2}\underbrace{\mathbf{h}^\mathrm{T}_{m_2+1}\cdots\mathbf{h}^\mathrm{T}_M}_{M-m_2}\Big]^\mathrm{T} \tag{4}$$

and $0 \leqslant m_1 < m_2 \triangleq m_1 + L + 1 \leqslant M$. Here, $\mathbf{H}^z_{L+1}$ groups the $L+2$ consecutive block-terms of $\mathbf{H}_M$ having the largest energy while replacing all the rest by zeros. We will call $\mathbf{H}^z_{L+1}$ the "($L+1$)st-order zero-padded significant part" of $\mathbf{H}_M$. On the other hand, $\mathbf{D}^z_{L+1}$ is the complement of $\mathbf{H}^z_{L+1}$ and will simply be referred to as the "unmodeled tails". We also define an additional vector containing exclusively the significant part as $\mathbf{H}_{L+1} \triangleq [\mathbf{h}^\mathrm{T}_{m_1}\cdots\mathbf{h}^\mathrm{T}_{m_2}]^\mathrm{T}$. Without loss of generality, we assume that $\|\mathbf{H}_M\|_2 = 1$ and $\|\mathbf{D}^z_{L+1}\|_2 \triangleq \varepsilon$. Our interest will be focused on the cases where ε is sufficiently smaller than 1, i.e., when $\varepsilon \ll 1$. In these cases, the size of the unmodeled part is "small" and a first-order perturbation analysis describes the algorithm behaviour well. In the cases where the size of the unmodeled tails ε is not sufficiently smaller than 1 we do not expect the algorithm to perform well, due to large undermodeling error. Furthermore, in these cases, a first-order perturbation analysis cannot describe the algorithm behaviour well.

To sum up, we will strive to determine what happens when an Lth-order equalizer attempts to equalize our *real-world* Mth-order channel $\mathbf{H}_M$ with $L < M-1$. We will attack the problem in three steps.

3.1. First step: ideal conditions

We assume for a moment an ideal situation whereby our real-world subchannels are described by $\mathbf{H}_{L+1}$. We can then obtain a zero-forcing equalizer of order L for delay i, by applying the already familiar steps of the algorithm:

$$\mathbf{X}_L(n) = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{s}_{2L+1}(n), \qquad \mathbf{R}_{L,i}(\mathbf{H}_{L+1}) = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}^\mathrm{H}_L(\mathbf{H}_{L+1}),$$
$$\mathbf{D}_i = \mathbf{R}_{L,i}(\mathbf{H}_{L+1})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathbf{R}^\mathrm{H}_{L,i}(\mathbf{H}_{L+1}), \tag{5}$$

$$\Delta\mathbf{D}_i = \mathbf{D}_i - \mathbf{D}_{i+1} = \mathbf{H}_i\mathbf{H}^\mathrm{H}_i. \tag{6}$$

A brief comment with respect to (5) is in order. On comparison with (3), we note that pseudoinversion has changed into inversion. This is because $\mathscr{T}_L(\mathbf{H}_{L+1})$ is now *square* and *full-column rank* (under the hypothesis that the subchannels do not share common zeros), thus it is *invertible*. Hence, $\mathbf{R}_{L,0}(\mathbf{H}_{L+1}) = \mathscr{T}_L(\mathbf{H}_{L+1})\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_{L+1})$ is *invertible*, as well. Similar to the earlier definition of $\hat{\mathbf{H}}_i$ in (1), $\mathbf{H}_i$ in (6) is defined as

$$\mathbf{H}_i \triangleq \mathscr{T}_L(\mathbf{H}_{L+1})(:, i+1) \tag{7}$$

and denotes the $(i+1)$st column of matrix $\mathscr{T}_L(\mathbf{H}_{L+1})$.

Following the discussion of Section 2.2, EVD of $\boldsymbol{\Delta}\mathbf{D}_i$ will reveal vector $\mathscr{H}_i = \pm\mathbf{H}_i$, that eventually leads to the desired equalizer:

$$\mathbf{g}_{L,i} = \mathbf{R}_{L,0}^{-1}(\mathbf{H}_{L+1})\mathscr{H}_i, \quad i = 0, \ldots, 2L+1.$$

When we put the equalizer into effect, we get the desired $(2L+1)$st-order combined response:

$$\mathbf{c}_{2(L+1),i} = \mathbf{g}_{L,i}^{\mathrm{H}}\mathscr{T}_L(\mathbf{H}_{L+1}) = \pm\,\mathbf{e}_{i+1}, \quad i = 0, \ldots, 2L+1,$$

where $\mathbf{c}_{2(L+1),i}$ is a vector with $2(L+1)$ elements. We note that if we replace subchannels $\mathbf{H}_{L+1}$ with the zero padded ones $\mathbf{H}_{L+1}^z$, we get

$$\mathbf{c}_{L+M+1,i} = \mathbf{g}_{L,i}^{\mathrm{H}}\mathscr{T}_L(\mathbf{H}_{L+1}^z) = \pm\,\mathbf{e}_{m_1+i+1}, \quad i = 0, \ldots, 2L+1,$$

where $\mathbf{c}_{L+M+1,i}$ is a vector with $L+M+1$ elements.

3.2. Second step: realistic conditions

After dealing with the ideal case, we proceed to examine the real-world situation. We still seek an Lth-order equalizer, only this time we attempt to equalize the subchannels $\mathbf{H}_M$. Re-applying the first steps of the algorithm, we get

$$\tilde{\mathbf{X}}_L(n) = \mathscr{T}_L(\mathbf{H}_M)\mathbf{s}_{L+M}(n), \qquad \mathbf{R}_{L,i}(\mathbf{H}_M) = \mathscr{T}_L(\mathbf{H}_M)\mathbf{J}_{L+M+1}^i\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M),$$

$$\tilde{\mathbf{D}}_i = \mathbf{R}_{L,i}(\mathbf{H}_M)\mathbf{R}_{L,0}^{-1}(\mathbf{H}_M)\mathbf{R}_{L,i}^{\mathrm{H}}(\mathbf{H}_M), \qquad \widetilde{\boldsymbol{\Delta}\mathbf{D}}_i = \tilde{\mathbf{D}}_i - \tilde{\mathbf{D}}_{i+1}.$$

The tilde over some of the preceding variables denotes a quantity that is *perturbed* with respect to its ideal counterpart. In other words: $\tilde{\mathbf{D}}_i = \mathbf{D}_i + \mathscr{E}(\mathbf{D}_i)$, $\widetilde{\boldsymbol{\Delta}\mathbf{D}}_i = \boldsymbol{\Delta}\mathbf{D}_i + \mathscr{E}(\boldsymbol{\Delta}\mathbf{D}_i)$, where the $\mathscr{E}$ terms are perturbation terms, resulting from the presence of the tails in the realistic case. Matrix $\mathbf{R}_{L,i}(\mathbf{H}_M)$ may also be considered as a perturbation of the ideal matrix $\mathbf{R}_{L,i}(\mathbf{H}_{L+1})$. Therefore, we write $\mathbf{R}_{L,i}(\mathbf{H}_M) = \mathbf{R}_{L,i}(\mathbf{H}_{L+1}) + \mathscr{E}(\mathbf{R}_{L,i})$ where for simplicity we have dispensed with the tilde in $\mathbf{R}_{L,i}(\mathbf{H}_M)$ since the same meaning is conveyed by the argument $\mathbf{H}_M$. For the same reason, we symbolize the perturbation term as $\mathscr{E}(\mathbf{R}_{L,i})$ instead of the more complicated $\mathscr{E}(\mathbf{R}_{L,i}(\mathbf{H}_{L+1}))$.

We should note that, unlike the ideal case, $\widetilde{\boldsymbol{\Delta}\mathbf{D}}_i$ need *not* be a rank-one matrix. By means of the EVD, computation of $\tilde{\lambda}_i$, the *maximum* eigenvalue of matrix $\widetilde{\boldsymbol{\Delta}\mathbf{D}}_i$, and $\tilde{\mathbf{H}}_i$, the perturbed version of $\mathbf{H}_i$ in (6), is possible. In practice though, there is always a sign ambiguity as we can only compute $\tilde{\mathscr{H}}_i = \pm\tilde{\mathbf{H}}_i$. Therefore, the perturbed equalizer is given by $\tilde{\mathbf{g}}_{L,i} = \tilde{\mathbf{R}}_{L,0}^{-1}(\mathbf{H}_M)\tilde{\mathscr{H}}_i$ and the respective combined response by $\tilde{\mathbf{c}}_{L+M+1,i} = \tilde{\mathbf{g}}_{L,i}^{\mathrm{H}}\mathscr{T}_L(\mathbf{H}_M)$.

In the next subsection, we will present the key results of a first-order perturbation analysis with respect to the size of the unmodeled tails ε. The goal of the analysis is to determine how the real-world problem variables differ from their ideal counterparts, due to the presence of the unmodeled tails.

3.2.1. Perturbation analysis results

Before we outline the results of our perturbation analysis, we need to introduce some extra notation. In particular, let $\tilde{\mathbf{l}}_i \triangleq \tilde{\mathbf{H}}_i/\|\tilde{\mathbf{H}}_i\|_2$, $\mathbf{l}_i \triangleq \mathbf{H}_i/\|\mathbf{H}_i\|_2$, $\mathscr{E}(\mathbf{l}_i) \triangleq \tilde{\mathbf{l}}_i - \mathbf{l}_i$, $\mathscr{E}(\mathbf{H}_i) \triangleq \tilde{\mathbf{H}}_i - \mathbf{H}_i$. We also remind that $\tilde{\lambda}_i$

symbolizes the largest eigenvalue of $\widetilde{\Delta\mathbf{D}}_i$ and that λ_i is the only non-zero eigenvalue of $\Delta\mathbf{D}_i$ and define $\mathscr{E}(\lambda_i) \triangleq \tilde{\lambda}_i - \lambda_i$. Finally, $\mathbf{P}_i^{\perp} = \mathbf{I} - \mathbf{l}_i\mathbf{l}_i^{\mathrm{H}}$ is the projector onto the orthogonal complement of the space spanned by $\mathbf{H}_i$. For each of the $\mathscr{E}$-terms the following approximations hold *to the first order*

Approximation 1:

$$\mathscr{E}(\mathbf{D}_i) = \mathbf{T}_{i,1} + \mathbf{T}_{i,2} + \mathbf{T}_{i,3} + \mathbf{T}_{i,4} + \mathbf{T}_{i,5} + \mathbf{T}_{i,6},$$

where

$$\mathbf{T}_{i,1} = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{D}^z_{L+1})\mathbf{J}^{-i}_{L+M+1}\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}^z_{L+1}),$$

$$\mathbf{T}_{i,2} = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{H}^z_{L+1})\mathbf{J}^{-i}_{L+M+1}\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1}),$$

$$\mathbf{T}_{i,5} = -\mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{H}^z_{L+1})\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1})\mathscr{T}_L^{-\mathrm{H}}(\mathbf{H}_{L+1})\mathbf{J}^{-i}_{2(L+1)}\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_{L+1})$$

and $\mathbf{T}_{i,3} = \mathbf{T}_{i,2}^{\mathrm{H}}$, $\mathbf{T}_{i,4} = \mathbf{T}_{i,1}^{\mathrm{H}}$, $\mathbf{T}_{i,6} = \mathbf{T}_{i,5}^{\mathrm{H}}$.

Approximation 2:

$$\mathscr{E}(\Delta\mathbf{D}_i) = \Delta\mathbf{T}_{i,1} + \Delta\mathbf{T}_{i,2} + \Delta\mathbf{T}_{i,3} + \Delta\mathbf{T}_{i,4} + \Delta\mathbf{T}_{i,5} + \Delta\mathbf{T}_{i,6}$$

where $\Delta\mathbf{T}_{i,j} \triangleq \mathbf{T}_{i,j} - \mathbf{T}_{i+1,j}$.

Approximation 3:

$$\mathscr{E}(\mathbf{l}_i) = \lambda_i^{-1}\mathbf{P}_i^{\perp}\mathscr{E}(\Delta\mathbf{D}_i)\mathbf{l}_i.$$

Approximation 4:

$$\mathscr{E}(\lambda_i) = \mathbf{l}_i^{\mathrm{H}}\mathscr{E}(\Delta\mathbf{D}_i)\mathbf{l}_i.$$

Approximation 5:

$$\mathscr{E}(\mathbf{H}_i) = \frac{1}{\lambda_i}\mathbf{P}_i^{\perp}\mathscr{E}(\Delta\mathbf{D}_i)\mathbf{H}_i + \frac{1}{2\lambda_i^2}(\mathbf{H}_i^H\mathscr{E}(\Delta\mathbf{D}_i)\mathbf{H}_i)\mathbf{H}_i.$$

The detailed proofs of all of the preceding results are found in the appendix.

3.3. Third step: blind LS vs. non-blind LS algorithm

By now we have developed the tools to assist us in assessing the behaviour of blind LS algorithm. We will do so by comparing its combined response with the corresponding one of non-blind LS algorithm.

One preliminary word is in order at this point: In the case under investigation, i.e., when $L < M - 1$, matrix $\mathscr{T}_L(\mathbf{H}_M)$ is “fat” and, with *probability* 1, it is a *full row-rank* matrix. Therefore, the product $\mathscr{T}_L(\mathbf{H}_M)\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M)$ is an *invertible* square matrix.

Non-blind LS equalizer is found by solving (2)

$$\hat{\mathbf{g}}_{L,i} = (\mathscr{T}_L(\mathbf{H}_M)\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{-1}\hat{\mathbf{H}}_i.$$

Blind LS equalizer is found in exactly the same way, only now the *estimate* $\tilde{\mathbf{H}}_i$ is used instead of the true column vector $\hat{\mathbf{H}}_i$:

$$\tilde{\mathbf{g}}_{L,i} = (\mathscr{T}_L(\mathbf{H}_M)\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{-1}\tilde{\mathbf{H}}_i.$$

Thus, the combined response in the non-blind case is found as

$$\hat{\mathbf{c}}_{L+M+1,i} = \hat{\mathbf{g}}_{L,i}^{\mathrm{H}} \mathscr{T}_L(\mathbf{H}_M) = \hat{\mathbf{H}}_i^{\mathrm{H}} (\mathscr{T}_L(\mathbf{H}_M)\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{-1} \mathscr{T}_L(\mathbf{H}_M) \;\Leftrightarrow$$

$$\hat{\mathbf{c}}_{L+M+1,i} = \hat{\mathbf{H}}_i^{\mathbf{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp}, \tag{8}$$

where we have used the property that for a full column-rank matrix $\mathbf{A}$ it holds $\mathbf{A}^{\sharp} = (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}$. In an identical way, the combined response of the blind equalizer is found to be

$$\tilde{\mathbf{c}}_{L+M+1,i} = \tilde{\mathbf{H}}_i^{\mathrm{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp}. \tag{9}$$

As can be seen from (8) and (9), and was expected anyway, $\hat{\mathbf{H}}_i$ and $\tilde{\mathbf{H}}_i$ are responsible for the differences in the combined responses. We note that in this theoretic treatment we have dispensed with the sign ambiguity that is a triviality of the implementation.

We will proceed using the perturbation analysis as an intermediate step to unveil instances where the combined responses of the algorithms are very close to each other, subject to the magnitude of the unmodeled tails ε. We distinguish the following cases based on the *actual* shape of the subchannels:

3.3.1. Trailing tails only

Suppose that the subchannels begin with an *actual* significant part of order L^*+1. That is to say, terms $\mathbf{h}_0$ and $\mathbf{h}_{L^*+1}$ are O(1) in magnitude (some intermediate terms may be smaller) while terms $\mathbf{h}_k$ for $k = L^*+2\ldots M$ are $O(\varepsilon)$ with $\varepsilon \ll 1$. As seen in Section 3.2.1, for $\mathbf{H}_i$ and $\tilde{\mathbf{H}}_i$ it holds

$$\tilde{\mathbf{H}}_i = \mathbf{H}_i + \mathscr{E}(\mathbf{H}_i), \quad i = 0, \ldots, 2L+1. \tag{10}$$

In addition, the filtering matrix $\mathscr{T}_L(\mathbf{H}_M)$ can be written as

$$\mathscr{T}_L(\mathbf{H}_M) = \mathscr{T}_L(\mathbf{H}_{L+1}^z) + \mathscr{T}_L(\mathbf{D}_{L+1}^z).$$

By defining

$$\mathbf{d}_i \triangleq \mathscr{T}_L(\mathbf{D}_{L+1}^z)(:, i+1) \tag{11}$$

and taking into account (1), there holds the following for each column of the participating matrices:

$$\hat{\mathbf{H}}_i = \mathscr{T}_L(\mathbf{H}_{L+1}^z)(:, i+1) + \mathbf{d}_i, \quad i = 0, \ldots, L+M. \tag{12}$$

We also remark that for $i = 0, \ldots, 2L+1$ it holds

$$\mathscr{T}_L(\mathbf{H}_{L+1}^z)(:, i+1) = \mathscr{T}_L(\mathbf{H}_{L+1})(:, i+1). \tag{13}$$

Thus, utilizing (7), (12), (13) we have

$$\hat{\mathbf{H}}_i = \mathbf{H}_i + \mathbf{d}_i, \quad i = 0, \ldots, 2L+1. \tag{14}$$

Hence, combining (10) and (14), we come up with

$$\tilde{\mathbf{H}}_i = \hat{\mathbf{H}}_i - \mathbf{d}_i + \mathscr{E}(\mathbf{H}_i), \quad i = 0, \ldots, 2L+1. \tag{15}$$

Plugging (15) into (9), we get the following:

$$\tilde{\mathbf{c}}_{L+M+1,i} = (\hat{\mathbf{H}}_i - \mathbf{d}_i + \mathscr{E}(\mathbf{H}_i))^{\mathrm{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp} \overset{(8)}{\Leftrightarrow}$$

$$\tilde{\mathbf{c}}_{L+M+1,i} = \hat{\mathbf{c}}_{L+M+1,i} + (-\mathbf{d}_i + \mathscr{E}(\mathbf{H}_i))^{\mathrm{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp} \;\Leftrightarrow$$

$$\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i} = (-\mathbf{d}_i + \mathscr{E}(\mathbf{H}_i))^{\mathrm{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp} \overset{(A.21)}{\Leftrightarrow}$$

$$\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i} = (-\mathbf{d}_i + \mathbf{P}_i^{\perp}\mathbf{d}_i)^{\mathrm{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp} \;\Leftrightarrow$$

$$\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i} = \left(-\frac{1}{\lambda_i}\mathbf{H}_i\mathbf{H}_i^{\mathrm{H}}\mathbf{d}_i\right)^{\mathrm{H}} (\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp}, \quad i = 0, \ldots, 2L+1. \tag{16}$$

As $i = 0, \ldots, 2L+1$ vectors $\mathbf{H}_i$ and $\mathbf{d}_i$ in (16) run through the following sequences of values, respectively:

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{h}_0 \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_0 \\ \mathbf{0} \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{h}_{L+1} \\ \mathbf{h}_L \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{h}_1 \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_{L+1} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{h}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{h}_{L+1} \\ \vdots \\ \vdots \\ \mathbf{h}_1 \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{h}_{L+1} \end{bmatrix},$$

$$\mathbf{d}_i = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{h}_{L+2} \\ \mathbf{0} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{h}_{L+3} \\ \mathbf{h}_{L+2} \\ \mathbf{0} \\ \vdots \\ \vdots \\ \mathbf{0} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{h}_{2L+1} \\ \mathbf{h}_{2L} \\ \mathbf{h}_{2L-1} \\ \vdots \\ \mathbf{h}_{L+2} \\ \mathbf{0} \end{bmatrix},$$

where $\mathbf{h}_k = 0$ if $k > M$.

It is therefore apparent that $\mathbf{H}_i^{\mathrm{H}}\mathbf{d}_i = 0$ for *every* $i = 0, \ldots, 2L+1$. Therefore, the difference of $\tilde{\mathbf{c}}_{L+M+1,i}$ from $\hat{\mathbf{c}}_{L+M+1,i}$ *vanishes* to the *first order*. Thus it holds

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i}\|_2 = \mathrm{O}(\varepsilon^2)\|\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M)^{\#}\|_2, \quad i = 0, \ldots, 2L+1.$$

In addition, $\|(\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\#}\|_2 = \|(\mathscr{T}_L(\mathbf{H}_M))^{\#}\|_2 = 1/\sigma_{\min}$ where $\sigma_{\min}$ is the minimum non-zero singular value of matrix $\mathscr{T}_L(\mathbf{H}_M)$. So we can write

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i}\|_2 = \mathrm{O}(\varepsilon^2)\frac{1}{\sigma_{\min}}, \quad i = 0, \ldots, 2L+1. \tag{17}$$

This finding has the interpretation that if $\sigma_{\min}$ is not very small, then $\tilde{\mathbf{c}}_{L+M+1,i}$ will be close to $\hat{\mathbf{c}}_{L+M+1,i}$, for *every* delay $i = 0, \ldots, 2L+1$ and hence blind LS equalizer performance is close to its non-blind counterpart.

Effective overmodeling: Of particular interest is the case of effective overmodeling. What we mean by that is the case where $L+1 > L^*+1$, i.e., our channel-order determination procedure has yielded an estimate larger than the order of the *actual* significant part of the channel.

What remains unknown in (17) is the order of magnitude of $\sigma_{\min}$. To this end, we can write $\mathscr{T}_L(\mathbf{H}_M) = \mathscr{T}_L(\mathbf{H}^z_{L^*+1}) + \mathscr{T}_L(\mathbf{D}^z_{L^*+1})$ where $\mathbf{H}^z_{L^*+1}$ is defined in the exact same spirit as $\mathbf{H}^z_{L+1}$ in (4) with the exception that it groups the L^*+2 terms of the *actual* significant part. $\mathbf{D}^z_{L^*+1}$ is the complement of $\mathbf{H}^z_{L^*+1}$ in the sense that $\mathbf{D}^z_{L^*+1} = \mathbf{H}_M - \mathbf{H}^z_{L^*+1}$.

In [12, p. 204] it is proved that if $\mathbf{E}$ is a perturbation on a matrix $\mathbf{A}$, then for each singular value σ of $\mathbf{A}$ it holds $|\sigma - \tilde{\sigma}| \leqslant \|E\|_2$, where $|\cdot|$ denotes absolute value and $\tilde{\sigma}$ the corresponding singular value of the perturbed matrix $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$. (Actually, a more general form of this statement is proved but for our purpose this special case suffices).

Considering $\mathscr{T}_L(\mathbf{H}^z_{L^*+1})$ as $\mathbf{A}$, $\mathscr{T}_L(\mathbf{D}^z_{L^*+1})$ as the perturbation and $\mathscr{T}_L(\mathbf{H}_M)$ as $\tilde{\mathbf{A}}$, application of the preceding theorem is straightforward. Moreover, the minimum singular value of matrix $\mathscr{T}_L(\mathbf{H}^z_{L^*+1})$ is *known* to be equal to zero. Thus $|0 - \sigma_{\min}| \leqslant \|\mathscr{T}_L(\mathbf{D}^z_{L^*+1})\|_2$ meaning that $\sigma_{\min}$ is an $\mathrm{O}(\varepsilon)$ quantity. When $0 \ll \sigma_{\min} \approx \varepsilon$, meaning that $\sigma_{\min}$ is much closer to ε than to 0, we can write from (17)

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i}\|_2 = \mathrm{O}(\varepsilon), \quad i = 0, \ldots, 2L+1. \tag{18}$$

Based on the findings of [4], if the *actual* significant parts of subchannels are sufficiently diverse, then non-blind LS equalizers for delays $i = 0, \ldots, L + (L^* + 1)$ exhibit satisfactory performance. Due to (18), the behaviour of blind LS equalizers for the same delays is roughly the same. Therefore, blind LS equalizers are expected to behave satisfactorily for $i = 0, \ldots, L + (L^* + 1)$.

For the rest of the delays, i.e., $i = L + (L^* + 1) + 1, \ldots, 2L + 1$, non-blind LS equalizers are known to generally behave poorly [4]. Therefore, blind LS equalizers will similarly exhibit poor behaviour as they approach the already erroneous performance of their non-blind counterparts.

Exact order case: It may happen of course that our channel order determination procedure identifies the *exact* order value of the *actual* significant part, i.e., $L + 1 = L^* + 1$. Two things differentiate this setting from effective overmodeling, the first being that now *all* of the $2(L+1)$ non-blind equalizers are known to behave satisfactorily, if the *actual* significant parts of the subchannels are sufficiently diverse [4]. The second is the order of magnitude of $\sigma_{\min}$.

Same as before, $\mathscr{T}_L(\mathbf{H}_M)$ may be written as $\mathscr{T}_L(\mathbf{H}_M) = \mathscr{T}_L(\mathbf{H}^z_{L^*+1}) + \mathscr{T}_L(\mathbf{D}^z_{L^*+1})$ where $2(L+1) \times (L+M+1)$ "fat" matrix $\mathscr{T}_L(\mathbf{H}^z_{L^*+1})$ has $L + (L^*+1) + 1$ non-zero singular values. Since $L = L^*$, *all* of its $2(L+1)$ singular values are non-zero. Thus, if $s_{\min}$ is the minimum singular value of $\mathscr{T}_L(\mathbf{H}^z_{L^*+1})$, then $|s_{\min} - \sigma_{\min}| \leqslant \|\mathscr{T}_L(\mathbf{D}^z_{L^*+1})\|_2$ meaning that $s_{\min} - \|\mathscr{T}_L(\mathbf{D}^z_{L^*+1})\|_2 \leqslant \sigma_{\min} \leqslant s_{\min} + \|\mathscr{T}_L(\mathbf{D}^z_{L^*+1})\|_2$.

It is seen that the order of magnitude of $\sigma_{\min}$ is drastically determined by $s_{\min}$, which is a measure of diversity of the *actual* significant part of the channel [5]. If $\varepsilon \ll s_{\min} \approx 1$, which means that the *actual* significant subchannels are sufficiently diverse, then

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i}\|_2 = \mathrm{O}(\varepsilon^2), \quad i = 0, \ldots, 2L+1.$$

This has the implication that *all* blind LS equalizers demonstrate good performance since their combined responses are extremely close to those of the non-blind LS equalizers which are known to behave well.

3.3.2. Subchannels with leading and trailing tails

Now, we suppose that each subchannel is comprised by a significant part of order $L^* + 1$ surrounded by leading and (possibly) trailing $\mathrm{O}(\varepsilon)$ tail terms. The total order of each subchannel is assumed to be M, the convention used throughout the paper.

Let us define $\mathbf{d}_{i+m_1} \triangleq \mathscr{T}_L(\mathbf{D}^z_{L+1})(:, i + m_1 + 1)$, $\hat{\mathbf{H}}_{i+m_1} \triangleq \mathscr{T}_L(\mathbf{H}_M)(:, i + m_1 + 1)$. These definitions are identical in spirit with those of $\mathbf{d}_i$, $\hat{\mathbf{H}}_i$ made earlier in (11) and (1). Working along the same lines as before

$$\mathscr{T}_L(\mathbf{H}_M) = \mathscr{T}_L(\mathbf{H}^z_{L+1}) + \mathscr{T}_L(\mathbf{D}^z_{L+1})$$

from which it follows that

$$\hat{\mathbf{H}}_{i+m_1} = \mathscr{T}_L(\mathbf{H}^z_{L+1})(:, i + m_1 + 1) + \mathbf{d}_{i+m_1}, \quad i = 0, \ldots, 2L+1. \tag{19}$$

In addition it holds

$$\mathscr{T}_L(\mathbf{H}^z_{L+1})(:, i + m_1 + 1) = \mathbf{H}_i, \quad i = 0, \ldots, 2L+1. \tag{20}$$

where $\mathbf{H}_i$ was defined in (7).

Combining (19), (20) we deduce the following

$$\hat{\mathbf{H}}_{i+m_1} = \mathbf{H}_i + \mathbf{d}_{i+m_1}, \quad i = 0, \ldots, 2L+1. \tag{21}$$

Eq. (10) is a *generic* result of our perturbation analysis and we repeat it here for completeness:

$$\tilde{\mathbf{H}}_i = \mathbf{H}_i + \mathscr{E}(\mathbf{H}_i), \quad i = 0, \ldots, 2L+1. \tag{22}$$

Combining (21) with (22) we get

$$\tilde{\mathbf{H}}_i = \hat{\mathbf{H}}_{i+m_1} - \mathbf{d}_{i+m_1} + \mathscr{E}(\mathbf{H}_i), \quad i = 0, \ldots, 2L+1. \tag{23}$$

Using (23) together with (9) yields

$$\tilde{\mathbf{c}}_{L+M+1,i} = (\hat{\mathbf{H}}_{i+m_1} - \mathbf{d}_{i+m_1} + \mathscr{E}(\mathbf{H}_i))^{\mathrm{H}}(\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp} \overset{(8)}{\Leftrightarrow}$$

$$\tilde{\mathbf{c}}_{L+M+1,i} = \hat{\mathbf{c}}_{L+M+1,i+m_1} + (-\mathbf{d}_{i+m_1} + \mathscr{E}(\mathbf{H}_i))^{\mathrm{H}}(\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp} \quad \Leftrightarrow$$

$$\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i+m_1} = (-\mathbf{d}_{i+m_1} + \mathscr{E}(\mathbf{H}_i))^{\mathrm{H}}(\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_M))^{\sharp}, \quad i = 0, \ldots, 2L+1.$$

By definition, $\mathbf{d}_{i+m_1}$ is an $\mathrm{O}(\varepsilon)$ quantity. The size of term $\mathscr{E}(\mathbf{H}_i)$ (see Approximation (5)) can be assessed by a perturbation bound. However, due to the complex form of $\mathscr{E}(\mathbf{H}_i)$, such a bound will be based on successive application of the triangle and submultiplicative matrix norm inequalities and will be quite loose, in general. (The usefulness of such bounds lies more in revealing potential instability points of the algorithm than in providing accurate perturbation size assessment.) In our study, since we are interested mostly in the order of magnitude of the difference between the combined responses of blind and non-blind LS equalizers, we may say that $\mathscr{E}(\mathbf{H}_i)$ is an $\mathrm{O}(\varepsilon)$ quantity, since every bound we could derive for this quantity is of the form $\alpha\varepsilon$, where α is a constant. Hence, we may write

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i+m_1}\|_2 = \mathrm{O}(\varepsilon)\ \frac{1}{\sigma_{\min}}, \quad i = 0, \ldots, 2L+1, \tag{24}$$

where $\sigma_{\min}$ symbolizes, again, the minimum singular value of $\mathscr{T}_L(\mathbf{H}_M)$.

Effective overmodeling: Like before, we assume that the estimate for the subchannels order exceeds the value of the actual significant part order, i.e., $L+1 > L^*+1$. All of the discussion made earlier about the order of magnitude of $\sigma_{\min}$ in the case of effective overmodeling is valid, so $\sigma_{\min}$ proves to be an $\mathrm{O}(\varepsilon)$ quantity. When $0 \ll \sigma_{\min} \approx \varepsilon$, we can write from (24)

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i+m_1}\|_2 = \mathrm{O}(1), \quad i = 0, \ldots, 2L+1. \tag{25}$$

In other words, the compared combined responses may differ by a *lot* at *worst*, as (25) suggests. They may be closer to each other for certain delays but, in general, they diverge. Consequently, even when non-blind LS equalizers are satisfactory, blind LS equalizers will generally be poor in performance. This fact will be more pronounced as the amount of overmodeling increases, as we will observe in the simulations section. Therefore, the case under investigation is recognized to be one where the algorithm does *not* exhibit robustness, in the sense that it usually performs poorly and the occasions where performance may be good cannot be predicted beforehand.

Exact order case: Situation improves in the exact order case ($L+1 = L^*+1$). Everything that has already been stated about the order of magnitude of $\sigma_{\min}$ is valid. When $\varepsilon \ll \sigma_{\min} \approx 1$, i.e., when the actual significant parts of subchannels are sufficiently diverse, we can write from (24)

$$\|\tilde{\mathbf{c}}_{L+M+1,i} - \hat{\mathbf{c}}_{L+M+1,i+m_1}\|_2 = \mathrm{O}(\varepsilon), \quad i = 0, \ldots, 2L+1.$$

In addition, *all* of the $2(L+1)$ non-blind LS equalizers are known to perform well, when the *actual* significant parts of the subchannels are sufficiently diverse [4]. Thus, *all* blind LS equalizers are expected to have satisfactory performance.

Recapitulating, we need to emphasize the critical dependence of algorithm behaviour on channel shape. When subchannels begin with “big” taps, blind LS performance is impervious to effective overmodeling. Blind LS equalizers will share the same performance pattern with their non-blind counterparts. Situation is radically different when subchannels incorporate leading “small” terms. In this case, effective overmodeling should be avoided as it generally results in poor equalization performance.

4. Simulations

We begin by constructing a significant part of order 3 for each of two subchannels. We are already familiar with the single vector representation: $\mathbf{H}_3 = [0.1949 \quad 0.7626 \quad -0.2186 \quad -0.3478 \quad 0.2766 \quad 0.0136 \quad -0.3465 \quad 0.1224]^{\mathrm{T}}$. In other words, the significant parts of the subchannels are: $\mathbf{H}_3^{(1)} = [0.1949 \quad -0.2186 \quad 0.2766 \quad -0.3465]^{\mathrm{T}}$, $\mathbf{H}_3^{(2)} = [0.7626 \quad -0.3478 \quad 0.0136 \quad 0.1224]^{\mathrm{T}}$. Their taps are drawn from a uniform distribution in the interval $[-1, 1]$ and $\|\mathbf{H}_3\|_2 = 1$. We will add tail terms to these significant parts such that the total order of each subchannel equals to 33. In the familiar notation, the total channel is represented by vector $\mathbf{H}_{33}$ and the zero padded significant part as $\mathbf{H}_{33}^z$. The magnitude of the tails is chosen such that: $10\log_{10}(\|\mathbf{H}_3\|_2^2/\|\mathbf{H}_{33} - \mathbf{H}_{33}^z\|_2^2) = 40\,\mathrm{dB}$. To illustrate the differences in algorithm behaviour depending on channel shape we add the tail terms in two ways: (1) We add all 30 tail terms *after* the significant part of each subchannel (2) We add 7 tail terms *before* the significant part of each subchannel and we leave the rest 23 *after* it. In both cases we use the *same* tail terms originally acquired from a random generator such that the aforementioned magnitude constraint is met.

For each channel we run blind LS algorithm for varying equalizer orders covering both effective overmodeling and exact order case. We compute blind LS equalizers for every possible delay. Simultaneously, we run non-blind LS algorithm for the same equalizer orders and all possible delays. As a stand alone measure for the equalization quality of blind LS algorithm we use open eye measure (OEM) defined for a vector $\mathbf{c}$ as:

$$\mathrm{OEM}(\mathbf{c}) = \left(\sum_i |c_i| - \max_i |c_i|\right) \Big/ \max_i |c_i|.$$

The smaller the OEM the better the equalization performance. We support our theoretic treatment by means of three graphs for the experiments made on each channel. We will comment on these graphs in the sequel.

Trailing tails only: Fig. 1 illustrates an effectively overmodeled case and consists of two subgraphs. The first one depicts the Euclidean distance of blind LS combined response from the closest non-blind LS combined response, as a function of the delay parameter for a fixed value of equalizer order $L = 6$. In it there are two horizontal lines that correspond to quantities $\varepsilon/\sigma_{\min}$ and $\varepsilon^2/\sigma_{\min}$ presented in the theoretic exposition. The second subgraph depicts OEM values of blind LS, again as a function of the delay parameter for the same equalizer order $L = 6$.

Our analysis predicts that the distance of the combined responses will be an $\mathrm{O}(\varepsilon^2/\sigma_{\min})$ quantity. Our prediction is consistent with the first subgraph of Fig. 1. Moreover, we assert that the equalizers corresponding to delay parameters from $i = 0, \ldots, L + (L^* + 1) = 0, \ldots, 9$ will perform well. This fact is verified by the second subgraph.

In Fig. 2, we present the distance of combined responses vs. the delay parameter for varying values of equalizer orders. We begin at $L = 2$ (exact order case) and we proceed up to $L = 6$. Theory claims this distance to be on the order of $\mathrm{O}(\varepsilon^2/\sigma_{\min})$ for every delay and equalizer order. What plays prominent role in the setting, though, is the value of $\sigma_{\min}$, which is *larger* for the exact order case. Thus blind equalizers for this case will be closer to the non-blind ones than in the effectively overmodeled cases. The effect of $\sigma_{\min}$ is verified by Fig. 2.

Fig. 3 sketches the OEM of blind LS vs. the delay parameter for the same equalizer orders as in Fig. 2. We are able to reconfirm our assertion that for a specific order L, the equalizers corresponding to delay parameters $i = 0, \ldots, L + (L^* + 1)$ perform satisfactorily. We also notice that among those equalizers who work well the ones of larger order behave better (smaller OEM). This is a known result in the context of non-blind LS equalization and since blind LS equalizers are close to their non-blind counterparts, they behave similarly.

Leading and trailing tails: Fig. 4 is the corresponding of Fig. 1 and likewise examines an effectively overmodeled case. The horizontal lines depict once more the values $\varepsilon/\sigma_{\min}$, $\varepsilon^2/\sigma_{\min}$ as discussed in the theoretic section. The order of the equalizer used is again $L = 6$. Based on our theoretic findings we expect

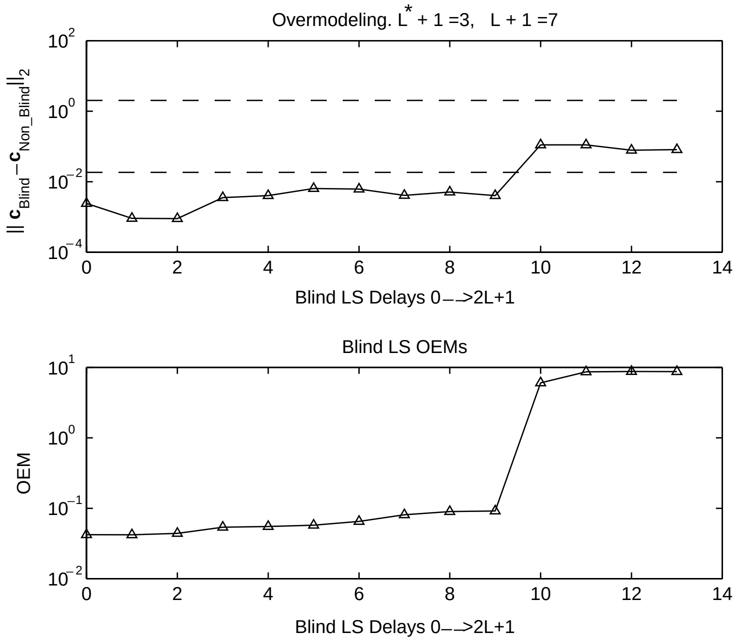

Fig. 1. Subchannels with trailing tail terms only, actual significant part order $L^*+1=3$, equalizer order $L=6$, i.e. effective overmodeling. Upper subplot: Euclidean distance of blind and non-blind LS combined responses vs. blind LS delay parameter. Lower subplot: Open eye measure for blind LS vs. blind LS delay parameter.

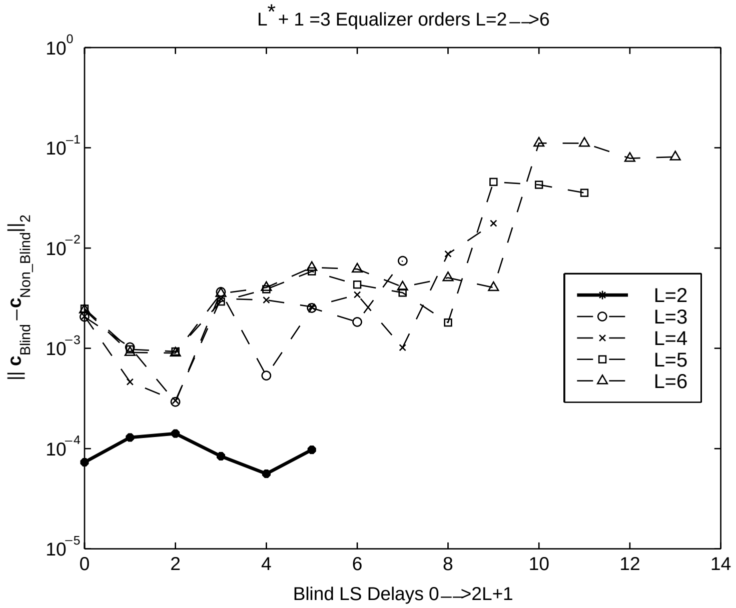

Fig. 2. Subchannels with trailing tail terms only, actual significant part order $L^* + 1 = 3$, varying equalizer order L (exact order case: $L = 2$). Euclidean distance of blind and non-blind LS combined responses vs. blind LS delay parameter.

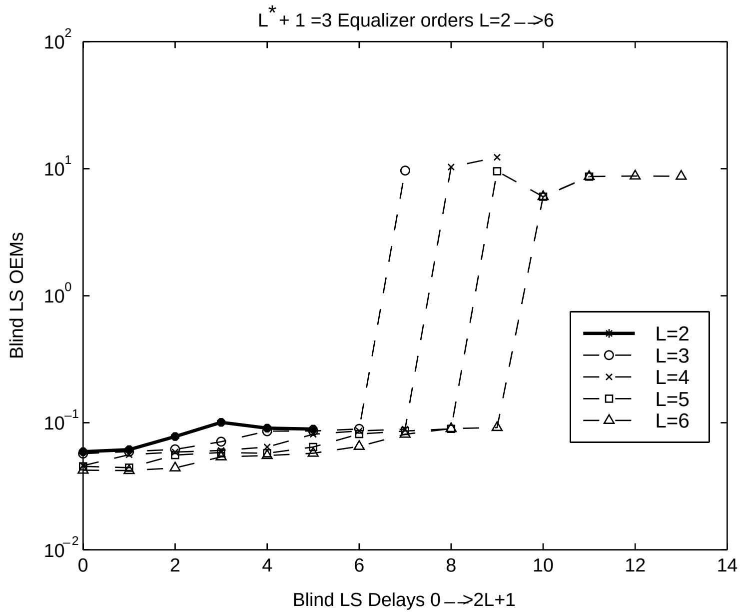

Fig. 3. Subchannels with trailing tail terms only, actual significant part order $L^* + 1 = 3$, varying equalizer order L (exact order case: $L = 2$). Open eye measure for blind LS vs. blind LS delay parameter.

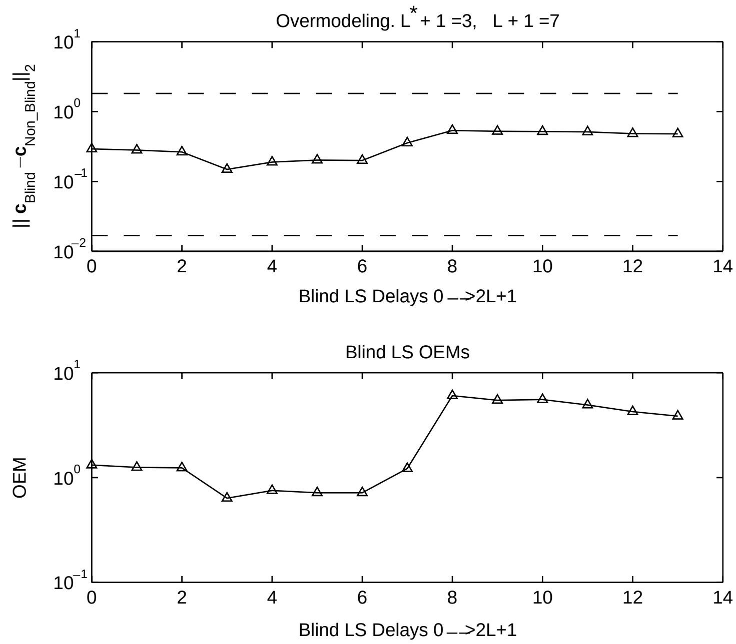

Fig. 4. Subchannels with leading and trailing tail terms, actual significant part order $L^* + 1 = 3$, equalizer order $L = 6$, i.e. effective overmodeling. Upper subplot: Euclidean distance of blind and non-blind LS combined responses vs. blind LS delay parameter. Lower subplot: Open eye measure for blind LS vs. blind LS delay parameter.

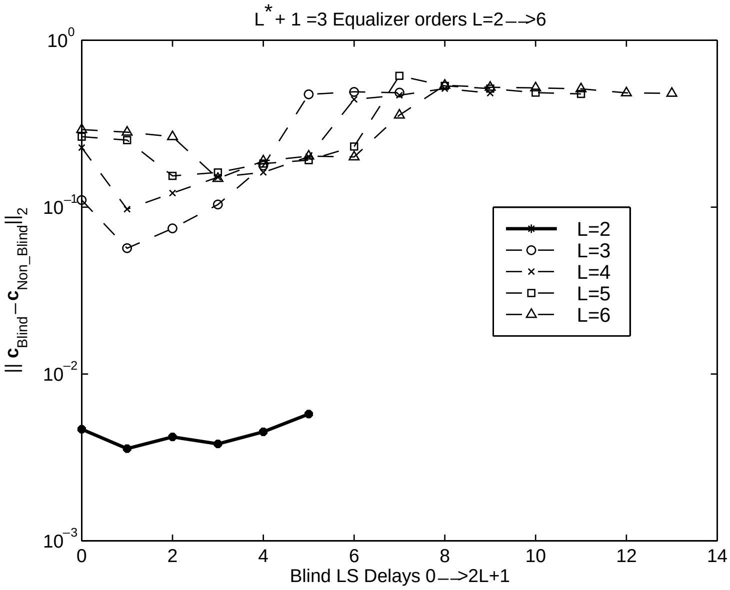

Fig. 5. Subchannels with leading and trailing tail terms, actual significant part order $L^* + 1 = 3$, varying equalizer order L (exact order case: $L = 2$). Euclidean distance of blind and non-blind LS combined responses vs. blind LS delay parameter.

the distance of the combined responses to be on the order of $O(\varepsilon/\sigma_{\min})$, which is roughly $O(1)$ for effective overmodeling. We generally expect non-satisfactory behaviour for every blind LS equalizer. Both of our expectations are confirmed in Fig. 4.

Fig. 5 is the counterpart of Fig. 2. We compute the distance of the combined responses for equalizer orders ranging from $L=2$ to 6 and every possible value of delay parameter thereof. Our analysis predicts this distance to be on the order of $O(\varepsilon/\sigma_{\min})$. Again $\sigma_{\min}$ plays a critical role in differentiating the exact order from the effectively overmodeled cases. Thus for the former the distance is roughly $O(\varepsilon)$ while for the latter is $O(1)$. This is exactly the situation depicted in Fig. 5.

Finally, Fig. 6 is the corresponding of Fig. 3. Here we can see the OEM values for blind LS equalizers of order $L = 2, \ldots, 6$ for every possible value of the delay parameter. Exact order case equalizers ($L = 2$) perform well for every delay, as we originally expected. The rest of the equalizers, all corresponding to effectively overmodeled cases, perform worse. The deterioration in performance is more severe as the order of the equalizer used is increased. We also notice that for the effectively overmodeled case where $L = 3$ and $i = 1$ the equalizer performance is rather good. We have come across this situation in our experiments in cases where we have effective overmodeling by a small number of taps. In these cases, some equalizers may perform favourably. The fact is, though, that equalizers will be unable to cope with their prescribed task once overmodeling gets larger, which indicates the sensitivity of the algorithm to this setting. To sum up, situation in Fig. 6 is perfectly in tune with our theoretic expectations.

5. Conclusions

We studied a blind LS algorithm for the construction of linear FIR equalizers in the FIR-SIMO channel setting. Our aim was to highlight instances where algorithm's erroneous perception of channel characteristics has a substantial impact on its performance. As errors in channel order determination are inevitable, information

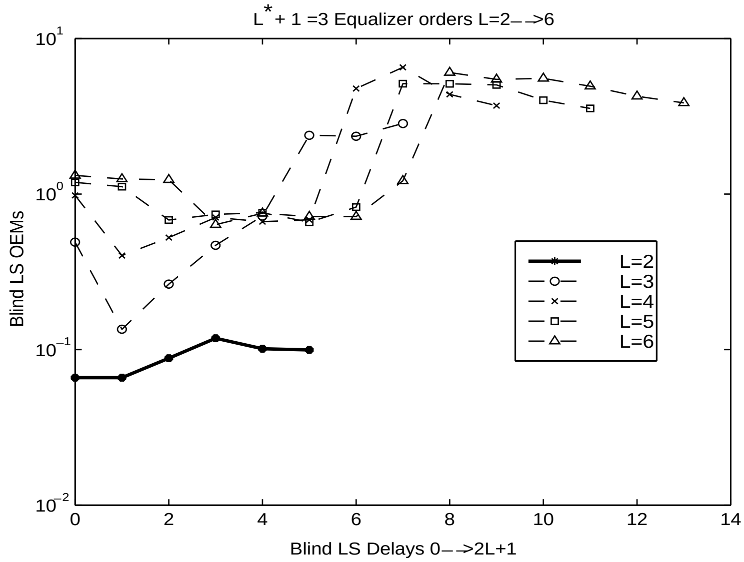

Fig. 6. Subchannels with leading and trailing tail terms, actual significant part order $L^* + 1 = 3$, varying equalizer order L (exact order case: $L = 2$). Open eye measure for blind LS vs. blind LS delay parameter.

about their influence on output quality is of significant practical interest. Microwave channels are usually comprised by a few large taps, which we called significant part and whose order we symbolized by $L^* + 1$, while the rest of them are smaller in magnitude. In this work, we examined the cases where channel order estimation led to the use of equalizers of order $L \geqslant L^*$. The analysis culminated in the performance comparison with non-blind LS. Our conclusions, verified by simulations, are summarized in the following: If channel does *not* contain any leading small terms then blind LS will behave similarly to non-blind LS irrespective of potential channel order estimation error. On the contrary, when channel possesses leading small terms, satisfactory results are generally expected only in the case where significant part order $L^* + 1$ is accurately identified. If order estimate is larger, performance of blind LS will generally be poor.

Acknowledgement

Authors are grateful to Professor Sergios Theodoridis for his support throughout the development of this work.

Appendix A

A.1. First-order perturbation expansion of $\mathbf{R}_{L,i}(\mathbf{H}_M)$

To start off with our perturbation analysis we note a useful equality with respect to the correlation matrix:

$$\begin{aligned}\mathbf{R}_{L,i}(\mathbf{H}_{L+1}) &= \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}^{\mathrm{H}}_L(\mathbf{H}_{L+1})\\ &= \mathscr{T}_L(\mathbf{H}^z_{L+1})\mathbf{J}^i_{L+M+1}\mathscr{T}^{\mathrm{H}}_L(\mathbf{H}^z_{L+1}) = \mathbf{R}_{L,i}(\mathbf{H}^z_{L+1}).\end{aligned} \quad \text{(A.1)}$$

That is to say, the correlation matrix is invariable to the zero-padding operation of its argument. The true correlation matrix can be analyzed as follows:

$$\begin{aligned}\mathbf{R}_{L,i}(\mathbf{H}_M) &= \mathscr{T}_L(\mathbf{H}_M)\mathbf{J}^i_{L+M+1}\mathscr{T}^{\mathrm{H}}_L(\mathbf{H}_M)\\ &= \{\mathscr{T}_L(\mathbf{H}^z_{L+1}) + \mathscr{T}_L(\mathbf{D}^z_{L+1})\}\mathbf{J}^i_{L+M+1}\{\mathscr{T}_L(\mathbf{H}^z_{L+1}) + \mathscr{T}_L(\mathbf{D}^z_{L+1})\}^{\mathrm{H}}.\end{aligned}$$

By omitting the second-order perturbation term $\mathscr{T}_L(\mathbf{D}^z_{L+1})\mathbf{J}^i_{L+M+1}\mathscr{T}^{\mathrm{H}}_L(\mathbf{D}^z_{L+1})$ we reach the first-order expansion:

$$\mathbf{R}_{L,i}(\mathbf{H}_M) = \mathbf{R}_{L,i}(\mathbf{H}^z_{L+1}) + \mathscr{E}(\mathbf{R}_{L,i}), \tag{A.2}$$

where

$$\mathscr{E}(\mathbf{R}_{L,i}) = \mathscr{T}_L(\mathbf{H}^z_{L+1})\mathbf{J}^i_{L+M+1}\mathscr{T}^{\mathrm{H}}_L(\mathbf{D}^z_{L+1}) + \mathscr{T}_L(\mathbf{D}^z_{L+1})\mathbf{J}^i_{L+M+1}\mathscr{T}^{\mathrm{H}}_L(\mathbf{H}^z_{L+1}). \tag{A.3}$$

Combining (A.1) and (A.2), we deduce the first-order perturbation expansion connecting the correlation matrices of real and ideal cases:

$$\mathbf{R}_{L,i}(\mathbf{H}_M) = \mathbf{R}_{L,i}(\mathbf{H}_{L+1}) + \mathscr{E}(\mathbf{R}_{L,i}). \tag{A.4}$$

A.2. First-order perturbation expansion of $\tilde{\mathbf{D}}_i$

$\tilde{\mathbf{D}}_i$ is given by

$$\tilde{\mathbf{D}}_i = \mathbf{R}_{L,i}(\mathbf{H}_M)\mathbf{R}^{-1}_{L,0}(\mathbf{H}_M)\mathbf{R}^{\mathrm{H}}_{L,i}(\mathbf{H}_M). \tag{A.5}$$

For $\mathbf{R}_{L,0}(\mathbf{H}_M)$ it holds to the first-order $\mathbf{R}_{L,0}(\mathbf{H}_M) = \mathbf{R}_{L,0}(\mathbf{H}_{L+1}) + \mathscr{E}(\mathbf{R}_{L,0})$ from (A.4). The first-order perturbation in its inverse is thus found to be [12, p. 130]

$$\mathbf{R}^{-1}_{L,0}(\mathbf{H}_M) = \mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1}) - \mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathscr{E}(\mathbf{R}_{L,0})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1}). \tag{A.6}$$

Plugging (A.4) and (A.6) into (A.5) and ignoring higher-order error terms, we get

$$\tilde{\mathbf{D}}_i = \mathbf{R}_{L,i}(\mathbf{H}_{L+1})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathbf{R}^{\mathrm{H}}_{L,i}(\mathbf{H}_{L+1}) + \mathbf{T}_{i,A} + \mathbf{T}_{i,B} + \mathbf{T}_{i,C} \quad \Leftrightarrow$$

$$\tilde{\mathbf{D}}_i = \mathbf{D}_i + \mathbf{T}_{i,A} + \mathbf{T}_{i,B} + \mathbf{T}_{i,C},$$

where terms $\mathbf{T}_{i,A}$, $\mathbf{T}_{i,B}$, $\mathbf{T}_{i,C}$ are analyzed as

$$\mathbf{T}_{i,A} = \mathbf{R}_{L,i}(\mathbf{H}_{L+1})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathscr{E}(\mathbf{R}_{L,i})^{\mathrm{H}}$$

$$\mathbf{T}_{i,B} = -\mathbf{R}_{L,i}(\mathbf{H}_{L+1})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathscr{E}(\mathbf{R}_{L,0})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathbf{R}^{\mathrm{H}}_{L,i}(\mathbf{H}_{L+1}),$$

$$\mathbf{T}_{i,C} = \mathscr{E}(\mathbf{R}_{L,i})\mathbf{R}^{-1}_{L,0}(\mathbf{H}_{L+1})\mathbf{R}^{\mathrm{H}}_{L,i}(\mathbf{H}_{L+1}) = \mathbf{T}^{\mathrm{H}}_{i,A}.$$

By substituting for $\mathbf{R}_{L,i}(\mathbf{H}_{L+1})$ and $\mathbf{R}_{L,0}(\mathbf{H}_{L+1})$ and employing (A.3) for $\mathscr{E}(\mathbf{R}_{L,i})$ and $\mathscr{E}(\mathbf{R}_{L,0})$, we can further expand $\mathbf{T}_{i,A}$ (hence $\mathbf{T}_{i,C}$), $\mathbf{T}_{i,B}$ as follows:

$$\mathbf{T}_{i,A} = \mathbf{T}_{i,1} + \mathbf{T}_{i,2}, \qquad \mathbf{T}_{i,C} = \mathbf{T}_{i,3} + \mathbf{T}_{i,4} = \mathbf{T}^{\mathrm{H}}_{i,2} + \mathbf{T}^{\mathrm{H}}_{i,1}, \qquad \mathbf{T}_{i,B} = \mathbf{T}_{i,5} + \mathbf{T}_{i,6},$$

where

$$\mathbf{T}_{i,1} = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{D}^z_{L+1})\mathbf{J}^{-i}_{L+M+1}\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}^z_{L+1}),$$

$$\mathbf{T}_{i,2} = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{H}^z_{L+1})\mathbf{J}^{-i}_{L+M+1}\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1}),$$

$$\mathbf{T}_{i,3} = \mathbf{T}^{\mathrm{H}}_{i,2}, \qquad \mathbf{T}_{i,4} = \mathbf{T}^{\mathrm{H}}_{i,1},$$

$$\mathbf{T}_{i,5} = -\mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{H}^z_{L+1})\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1})\mathscr{T}_L^{-\mathrm{H}}(\mathbf{H}_{L+1})\mathbf{J}^{-i}_{2(L+1)}\mathscr{T}_L^{\mathrm{H}}(\mathbf{H}_{L+1}),$$

$$\mathbf{T}_{i,6} = \mathbf{T}^{\mathrm{H}}_{i,5}. \tag{A.7}$$

That is to say, the first-order perturbation expansion for $\tilde{\mathbf{D}}_i$, with the aid of the terms above, shapes into the following:

$$\tilde{\mathbf{D}}_i = \mathbf{D}_i + \mathscr{E}(\mathbf{D}_i), \tag{A.8}$$

where

$$\mathscr{E}(\mathbf{D}_i) = \mathbf{T}_{i,1} + \mathbf{T}_{i,2} + \mathbf{T}_{i,3} + \mathbf{T}_{i,4} + \mathbf{T}_{i,5} + \mathbf{T}_{i,6}.$$

A.3. First-order perturbation expansion of $\widetilde{\Delta\mathbf{D}}_i$

Since $\tilde{\mathbf{D}}_i$ is as in (A.8), $\widetilde{\Delta \mathrm{D}}_i$ will equal to

$$\widetilde{\Delta \mathrm{D}}_i = \Delta\mathbf{D}_i + \mathscr{E}(\Delta\mathbf{D}_i), \tag{A.9}$$

where

$$\mathscr{E}(\Delta\mathbf{D}_i) = \mathscr{E}(\mathbf{D}_i) - \mathscr{E}(\mathbf{D}_{i+1}) = \Delta\mathbf{T}_{i,1} + \Delta\mathbf{T}_{i,2} + \Delta\mathbf{T}_{i,3} + \Delta\mathbf{T}_{i,4} + \Delta\mathbf{T}_{i,5} + \Delta\mathbf{T}_{i,6},$$

where $\Delta\mathbf{T}_{i,j} \triangleq \mathbf{T}_{i,j} - \mathbf{T}_{i+1,j}$.

A.3.1. Trailing tail terms only

We will first present the case where the tails *strictly follow* the significant part. For this special case, a lot of cancellations take place and this has a pronounced impact on the subsequent stages of the analysis.

Knowing $\mathbf{T}_{i,2}$ from (A.7), we can further simplify as follows:

$$\mathbf{T}_{i,2} = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)}\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})[\mathscr{T}_L(\mathbf{H}_{L+1}) \quad \mathbf{0}_{2(L+1)\times(M-L-1)}]\mathbf{J}^{-i}_{L+M+1}\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1})$$

$$= [\mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)} \quad \mathbf{0}_{2(L+1)\times(M-L-1)}]\mathbf{J}^{-i}_{L+M+1}\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1}) \tag{A.10}$$

$$= [\mathbf{0}_{2(L+1)\times i} \quad \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}^i_{2(L+1)} \quad \mathbf{0}_{2(L+1)\times(M-L-1-i)}]\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1}). \tag{A.11}$$

In the transition from (A.10) to (A.11) we note that right multiplication by matrix $\mathbf{J}^{-i}_{L+M+1}$ merely equals to a shift of the preceding matrix i columns to the right.

Upon a little reflection we deduce that

$$\Delta\mathbf{T}_{i,2} = \mathbf{T}_{i,2} - \mathbf{T}_{i+1,2}$$

$$= [\mathbf{0}_{2(L+1)\times i} \quad \mathscr{T}_L(\mathbf{H}_{L+1})(:, i+1) \quad \mathbf{0}_{2(L+1)\times(L+M-i)}]\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1})$$

$$= [\mathbf{0}_{2(L+1)\times i} \quad \mathbf{H}_i \quad \mathbf{0}_{2(L+1)\times(L+M-i)}]\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}^z_{L+1}), \tag{A.12}$$

where we have made use of the shorter notation $\mathbf{H}_i = \mathscr{T}_L(\mathbf{H}_{L+1})(:, i+1)$.

In addition, it turns out that $\mathbf{T}_{i,1} = -\mathbf{T}_{i,6}$ and $\mathbf{T}_{i,4} = -\mathbf{T}_{i,5}$ for all i, meaning that

$$\mathbf{T}_{i,1} + \mathbf{T}_{i,4} + \mathbf{T}_{i,5} + \mathbf{T}_{i,6} = 0, \qquad \mathbf{T}_{i+1,1} + \mathbf{T}_{i+1,4} + \mathbf{T}_{i+1,5} + \mathbf{T}_{i+1,6} = 0$$

and hence

$$\Delta\mathbf{T}_{i,1} + \Delta\mathbf{T}_{i,4} + \Delta\mathbf{T}_{i,5} + \Delta\mathbf{T}_{i,6} = 0,$$

which leaves us with nothing but the very elegant

$$\widetilde{\Delta\mathbf{D}}_i = \Delta\mathbf{D}_i + \Delta\mathbf{T}_{i,2} + \mathbf{T}_{i,2}^{\mathrm{H}}. \tag{A.13}$$

In order to prove that $\mathbf{T}_{i,4} = -\mathbf{T}_{i,5}$ it suffices to prove that

$$\mathscr{T}_L(\mathbf{H}_{L+1}^z)\mathbf{J}_{L+M+1}^i = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}_{2(L+1)}^i\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})\mathscr{T}_L(\mathbf{H}_{L+1}^z). \tag{A.14}$$

This claim should be evident upon inspecting $\mathbf{T}_{i,4}, \mathbf{T}_{i,5}$. Thus, from (A.14), we have equivalently:

$$[\mathscr{T}_L(\mathbf{H}_{L+1}) \quad \mathbf{0}_{2(L+1)\times(M-L-1)}]\mathbf{J}_{L+M+1}^i = \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}_{2(L+1)}^i\mathscr{T}_L^{-1}(\mathbf{H}_{L+1})[\mathscr{T}_L(\mathbf{H}_{L+1}) \quad \mathbf{0}_{2(L+1)\times(M-L-1)}] \Leftrightarrow$$

$$[\mathscr{T}_L(\mathbf{H}_{L+1}) \quad \mathbf{0}_{2(L+1)\times(M-L-1)}]\mathbf{J}_{L+M+1}^i = [\mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}_{2(L+1)}^i \quad \mathbf{0}_{2(L+1)\times(M-L-1)}]. \tag{A.15}$$

The truth of (A.15) should be obvious. Right multiplication by matrix $\mathbf{J}_{L+M+1}^i$ amounts to a left shift of the preceding matrix by i columns. Due to the special form of matrix $\mathbf{A} \triangleq [\mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{0}]$, it does not make any difference whether we shift it as a whole (left-hand side of (A.15)) or just shift its leftmost submatrix (right-hand side of (A.15)).

Having proved that $\mathbf{T}_{i,4} = -\mathbf{T}_{i,5}$ it is straightforward that $\mathbf{T}_{i,1} = -\mathbf{T}_{i,6}$ holds, as $\mathbf{T}_{i,4} = \mathbf{T}_{i,1}^{\mathrm{H}}$, $\mathbf{T}_{i,5} = \mathbf{T}_{i,6}^{\mathrm{H}}$.

A.3.2. General case

In the general case, things are pretty much the same as long as $\Delta\mathbf{T}_{i,2}$ is concerned. Working along the exact same lines as before we can find that

$$\Delta\mathbf{T}_{i,2} = [\mathbf{0}_{2(L+1)\times m_1} \quad \mathbf{0}_{2(L+1)\times i} \quad \mathbf{H}_i \quad \mathbf{0}_{2(L+1)\times(L+M-m_1-i)}]\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}_{L+1}^z). \tag{A.16}$$

This is analogous to (A.12), only now there are m_1 extra zero columns in the leftmost part of $\Delta\mathbf{T}_{i,2}$ due to the presence of the leading tails in the subchannels.

For the rest of the terms, however, the situation is substantially different. Equations $\mathbf{T}_{i,1} = -\mathbf{T}_{i,6}$, $\mathbf{T}_{i,4} = -\mathbf{T}_{i,5}$ *no longer* hold for an i other than $i = 0$. The counterpart of (A.15) now becomes

$$[\mathbf{0}_{2(L+1)\times m_1} \quad \mathscr{T}_L(\mathbf{H}_{L+1}) \quad \mathbf{0}_{2(L+1)\times(M-L-m_1-1)}]\mathbf{J}_{L+M+1}^i$$
$$= [\mathbf{0}_{2(L+1)\times m_1} \quad \mathscr{T}_L(\mathbf{H}_{L+1})\mathbf{J}_{2(L+1)}^i \quad \mathbf{0}_{2(L+1)\times(M-L-m_1-1)}],$$

which is *not* a valid equality for $i \neq 0$ due to the presence of the leading zero submatrices. Therefore, we are now deprived of the nullification of $\mathbf{T}_{i,1} + \mathbf{T}_{i,4} + \mathbf{T}_{i,5} + \mathbf{T}_{i,6}$ for $i \neq 0$. As a consequence, there is little to be done to simplify $\mathscr{E}(\Delta\mathbf{D}_i)$. Grouping of some terms can result in a different form for $\mathscr{E}(\Delta\mathbf{D}_i)$, but this has mainly to do with aesthetics and not the essence of a true simplification. As a result, for the general case we just stick to (A.9) in its original form.

A.4. First-order perturbation in the largest eigenvalue/eigenvector of $\widetilde{\Delta\mathbf{D}}_i$

The next step of the algorithm amounts to obtaining $\tilde{\lambda}_i$, the largest eigenvalue of $\widetilde{\Delta\mathbf{D}}_i$, along with the corresponding eigenvector. In the ideal case, $\Delta\mathbf{D}_i = \mathbf{H}_i\mathbf{H}_i^{\mathrm{H}}$, $\lambda_i = \|\mathbf{H}_i\|_2^2$, the corresponding eigenvector is $\mathbf{H}_i$ and its unit-norm counterpart is $\mathbf{l}_i = \mathbf{H}_i/\|\mathbf{H}_i\|_2$. In [12, p. 240, Section 2.3] we find that to the first order it holds

$$\mathscr{E}(\mathbf{l}_i) \triangleq \tilde{\mathbf{l}}_i - \mathbf{l}_i = \lambda_i^{-1}\mathbf{P}_i^{\perp}\mathscr{E}(\Delta\mathbf{D}_i)\mathbf{l}_i, \tag{A.17}$$

where $\tilde{\mathbf{l}}_i$ is the perturbed version of $\mathbf{l}_i$, $\mathscr{E}(\mathbf{\Delta D}_i)$ is the (first-order) perturbation in $\widetilde{\mathbf{\Delta D}}_i$ and $\mathbf{P}_i^{\perp}$ is the projector onto the complement of the subspace produced by $\mathbf{l}_i$. Therefore for $\mathbf{P}_i^{\perp}$ it holds: $\mathbf{P}_i^{\perp} = \mathbf{I}_{2(L+1)} - \mathbf{l}_i\mathbf{l}_i^{\mathrm{H}}$. We trivially note that in the ideal case $\mathbf{H}_i = \|\mathbf{H}_i\|_2\mathbf{l}_i = \sqrt{\lambda_i}\mathbf{l}_i$. Thus, in the real case we can compute $\tilde{\mathbf{H}}_i$ as: $\tilde{\mathbf{H}}_i = \sqrt{\tilde{\lambda}_i}\tilde{\mathbf{l}}_i$.

In addition, in [12, p. 183] we find that

$$\mathscr{E}(\lambda_i) \triangleq \tilde{\lambda}_i - \lambda_i = \mathbf{l}_i^{\mathrm{H}}\mathscr{E}(\mathbf{\Delta D}_i)\mathbf{l}_i = \frac{\mathbf{H}_i^{\mathrm{H}}}{\|\mathbf{H}_i\|_2}\mathscr{E}(\mathbf{\Delta D}_i)\frac{\mathbf{H}_i}{\|\mathbf{H}_i\|_2} = \frac{1}{\lambda_i}\mathbf{H}_i^{\mathrm{H}}\mathscr{E}(\mathbf{\Delta D}_i)\mathbf{H}_i. \tag{A.18}$$

With the aid of (A.17) and (A.18), we can compute a first-order approximation for $\mathscr{E}(\mathbf{H}_i) \triangleq \tilde{\mathbf{H}}_i - \mathbf{H}_i$. As a preliminary step we need to know a first-order approximation for $\sqrt{\tilde{\lambda}_i} - \sqrt{\lambda_i}$. A first-order expression for this quantity is found by means of the Taylor series for the function $f(x) = \sqrt{x}$. Keeping only the first-order terms we have

$$\sqrt{x} = \sqrt{x_0} + \frac{1}{2\sqrt{x_0}}(x - x_0)$$

from which we deduce

$$\sqrt{\tilde{\lambda}_i} - \sqrt{\lambda_i} = \frac{1}{2\sqrt{\lambda_i}}\mathscr{E}(\lambda_i).$$

Thus to the first order,

$$\mathscr{E}(\mathbf{H}_i) = \tilde{\mathbf{H}}_i - \mathbf{H}_i = \sqrt{\tilde{\lambda}_i}\tilde{\mathbf{l}}_i - \sqrt{\lambda_i}\mathbf{l}_i = \left(\sqrt{\lambda_i} + \frac{1}{2\sqrt{\lambda_i}}\mathscr{E}(\lambda_i)\right)(\mathbf{l}_i + \mathscr{E}(\mathbf{l}_i)) - \sqrt{\lambda_i}\mathbf{l}_i$$

$$= \sqrt{\lambda_i}\mathscr{E}(\mathbf{l}_i) + \frac{1}{2\sqrt{\lambda_i}}\mathscr{E}(\lambda_i)\mathbf{l}_i. \tag{A.19}$$

Combining (A.17), (A.18) and (A.19), we reach the following first-order result

$$\mathscr{E}(\mathbf{H}_i) = \tilde{\mathbf{H}}_i - \mathbf{H}_i = \frac{1}{\lambda_i}\mathbf{P}_i^{\perp}\mathscr{E}(\mathbf{\Delta D}_i)\mathbf{H}_i + \frac{1}{2\lambda_i^2}(\mathbf{H}_i^H\mathscr{E}(\mathbf{\Delta D}_i)\mathbf{H}_i)\mathbf{H}_i. \tag{A.20}$$

A.4.1. Trailing tail terms only

We can take advantage of the simple form $\mathscr{E}(\mathbf{\Delta D}_i)$ has and obtain a simpler form for (A.20). We remind that $\mathscr{E}(\mathbf{\Delta D}_i)$ for this particular case is given by (A.12) and (A.13). Thus, we may decompose the first term of (A.20) as

$$\frac{1}{\lambda_i}\mathbf{P}_i^{\perp}\mathscr{E}(\mathbf{\Delta D}_i)\mathbf{H}_i = \frac{1}{\lambda_i}\mathbf{P}_i^{\perp}(\mathbf{\Delta T}_{i,2} + \mathbf{\Delta T}_{i,2}^{\mathrm{H}})\mathbf{H}_i = \frac{1}{\lambda_i}\mathbf{P}_i^{\perp}\mathbf{\Delta T}_{i,2}^{\mathrm{H}}\mathbf{H}_i$$

$$= \mathbf{P}_i^{\perp}\mathscr{T}_L(\mathbf{D}_{L+1}^z)\mathbf{e}_{i+1} = \mathbf{P}_i^{\perp}\mathbf{d}_i.$$

The second term of (A.20) *vanishes* as subsequent calculations suggest

$$\mathbf{H}_i^{\mathrm{H}}\mathscr{E}(\mathbf{\Delta D}_i)\mathbf{H}_i = \mathbf{H}_i^{\mathrm{H}}\mathbf{\Delta T}_{i,2}\mathbf{H}_i + (\mathbf{H}_i^{\mathrm{H}}\mathbf{\Delta T}_{i,2}\mathbf{H}_i)^{\mathrm{H}},$$

where

$$\mathbf{H}_i^{\mathrm{H}}\mathbf{\Delta T}_{i,2}\mathbf{H}_i = \lambda_i\mathbf{e}_{i+1}^{\mathrm{H}}\mathscr{T}_L^{\mathrm{H}}(\mathbf{D}_{L+1}^z)\mathbf{H}_i = \lambda_i\mathbf{d}_i^{\mathrm{H}}\mathbf{H}_i.$$

However, as we explained in detail in Section 3.3.1, $\mathbf{d}_i^{\mathrm{H}}\mathbf{H}_i = 0$ for all $i = 0, \ldots, 2L+1$.

Consequently, (A.20) assumes the simple form

$$\mathscr{E}(\mathbf{H}_i) = \mathbf{P}_i^{\perp}\mathbf{d}_i. \tag{A.21}$$

References

[1] K. Abed-Meraim, E. Moulines, P. Loubaton, Prediction error method for second-order blind identification, IEEE Trans. Signal Proc. 45 (3) (March 1997) 694–705.
[2] D. Gesbert, P. Duhamel, S. Mayrargue, On-line blind multichannel equalization based on mutually referenced filters, IEEE Trans. Signal Proc. 45 (9) (September 1997) 2307–2317.
[3] G.B. Giannakis, S.D. Halford, Blind fractionally spaced equalization of noisy FIR channels: direct and adaptive solutions IEEE Trans. Signal Proc. 45 (9) (September 1997) 2277–2292.
[4] A.P. Liavas, Least-squares channel equalization performance versus equalization delay, IEEE Trans. Signal Proc. 48 (6) (June 2000) 1832–1835.
[5] A.P. Liavas, P.A. Regalia, J.-P. Delmas, Robustness of least-squares and subspace methods for blind channel identification/equalization with respect to effective channel undermodeling/overmodeling, IEEE Trans. Signal Proc. 47 (6) (June 1999) 1636–1645.
[6] A.P. Liavas, P.A. Regalia, J.-P. Delmas, On the robustness of the linear prediction method for blind channel identification with respect to effective channel undermodeling/overmodeling, IEEE Trans. Signal Proc. 48 (5) (May 2000) 1477–1481.
[7] E. Moulines, P. Duhamel, J-F Cardoso, S. Mayrargue, Subspace methods for the blind identification of multichannel FIR filters, IEEE Trans. Signal Proc. 43 (2) (February 1995) 516–525.
[8] C.B. Papadias, D.T.M. Slock, Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction, IEEE Trans. Signal Proc. 47 (3) (March 1999) 641–654.
[9] J.G. Proakis, Digital Communications, 3rd Edition, McGraw-Hill, New York, 1995.
[10] J. Shen, Z. Ding, Direct blind MMSE channel equalization based on second-order statistics, IEEE Trans. Signal Proc. 48 (4) (April 2000) 1015–1022.
[11] D.T.M. Slock, Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction, Proceedings of the International Conference on Acoustics Speech and Signal Processing, Adelaide, Australia, Vol. IV, April 19–22, 1994, pp. 585–588.
[12] G. Stewart, J. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.
[13] L. Tong, G. Xu, T. Kailath, Blind identification and equalization based on second-order statistics: a time domain approach IEEE Trans. Inform. Theory 40 (2) (March 1994) 340–349.
[14] G. Xu, H. Liu, L. Tong, T. Kailath, A least-squares approach to blind channel identification, IEEE Trans. Signal Proc. 43 (12) (December 1995) 2982–2993.