OFDM channel estimation by singular value decomposition

Edfors, Ove; Sandell, Magnus; van de Beek, Jan-Jaap; Wilson, Sarah Kate; Börjesson, Per Ola

1996

OFDM channel estimation by singular value decomposition*

Ove Edfors† Magnus Sandell† Jan-Jaap van de Beek†
Sarah Kate Wilson‡ Per Ola Börjesson†

† Division of Signal Processing
Luleå University of Technology
S-971 87 Luleå
SWEDEN

‡ School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907
U.S.A.

* This work has been presented in part at the 1996 Vehicular Technology Conference (VTC'96) in Atlanta, Georgia, Apr 28–May 1, 1996, pp 923-927.

Abstract

In this paper we present and analyse low-rank channel estimators for orthogonal frequency-division multiplexing (OFDM) using the frequency correlation of the channel. Low-rank approximations based on the discrete Fourier transform (DFT) have been proposed but they suffer from poor performance when the channel is not sample-spaced. We apply the theory of optimal rank-reduction to linear minimum mean-squared error (LMMSE) estimators and show that these estimators, when using a fixed design, are robust to changes in channel correlation and signal-to-noise ratio (SNR). The performance is presented in terms of uncoded symbol-error rate (SER) for a system using 16-QAM.

Contents

Chapter 1

Introduction

Wireless digital communication systems using multi-amplitude modulation schemes, such as quadrature amplitude modulation (QAM), require estimation and tracking of the fading channel. In general, this means a more complex receiver than for differential modulation schemes, such as differential phase-shift keying (DPSK), where the receivers operate without a channel estimate [1].

In orthogonal frequency-division multiplexing (OFDM) systems, DPSK is appropriate for relatively low data rates, such as in the European digital-audio broadcast (DAB) system [2]. However, for more spectrally-efficient OFDM systems, coherent modulation is more appropriate.

The structure of OFDM signalling allows a channel estimator to use both time and frequency correlation. Such a two-dimensional estimator structure is generally too complex for a practical implementation. To reduce the complexity, separating the use of time and frequency correlation has been proposed [3]. This combined scheme uses two separate FIR-Wiener-filters, one in the frequency direction and the other in the time direction.

In this paper we present and analyse a class of block-oriented channel estimators for OFDM, where only the frequency correlation of the channel is used in the estimation. Whatever their level of performance, it may be improved with the addition of a second filter using the time correlation [3, 4].

Though a linear minimum mean-squared error (LMMSE) estimator using only frequency correlation has lower complexity than one using both time and frequency correlation, it still requires a large number of operations. We introduce a low-complexity approximation to a frequency-based LMMSE estimator that uses the theory of optimal rank reduction. Other types of low-rank approximations, based on the discrete-time Fourier transform (DFT), have been proposed for OFDM systems before [5, 6, 7]. The work presented in this paper was inspired by the observations in [7], where it is shown that DFT-based low-rank channel estimators have limited performance for non-sample-spaced channels and high SNRs.

After presenting the OFDM system model and our scenario in Section 2, we introduce the estimators and derive their complexities in Section 3. We analyse the symbol-error rate (SER) performance in Section 4 where we also discuss design considerations. The proposed low-rank estimator is compared to other estimators in Section 5 and a summary and concluding remarks appear in Section 6.

Chapter 2

System description

2.1 System model

Figure 2.1 displays the OFDM base-band model used in this paper. We assume that the use of a cyclic prefix (CP) [8] both preserves the orthogonality of the tones and eliminates inter-symbol interference (ISI) between consecutive OFDM symbols. Further, the channel $g(t, \tau)$ is assumed to be slowly fading, so it is considered to be constant during one OFDM symbol. The number of tones in the system is N, and the length of the cyclic prefix is L samples.

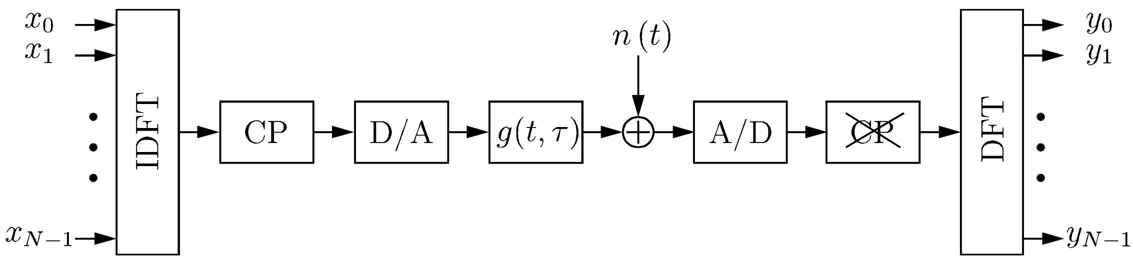

Figure 2.1: Base band model of an OFDM system. 'CP' denotes the cyclic prefix.

Under these assumptions we can describe the system as a set of parallel Gaussian channels, shown in Figure 2.2, with correlated attenuations h_k. The attenuations on each tone are given by

$$h_k = G\left(\frac{k}{NT_s}\right), \; k = 0 \ldots N-1,$$

where $G(\cdot)$ is the frequency response of the channel $g(t, \tau)$ during the OFDM symbol, and T_s is the sampling period of the system. In matrix notation we describe the OFDM system as

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{y}$ is the received vector, $\mathbf{X}$ is a matrix containing the transmitted signalling points on its diagonal, $\mathbf{h}$ is a channel attenuation vector, and $\mathbf{n}$ is a vector of i.i.d. complex, zero-mean, Gaussian noise with variance σ_n^2.

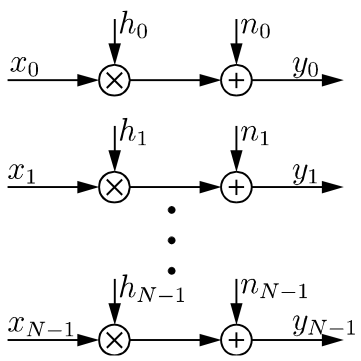

Figure 2.2: The OFDM system, described as a set of parallel Gaussian channels with correlated attenuations.

2.2 Channel model

We are using a fading multi-path channel model [1], consisting of M impulses

$$g(\tau) = \sum_{k=0}^{M-1} \alpha_k \delta(\tau - \tau_k T_s), \tag{2.2}$$

where α_k are zero-mean, complex Gaussian, random variables, with a power-delay profile $\theta(\tau_k)$. In this paper we have used $M = 5$ impulses and two versions of this channel model:

- **Synchronized channel**. This is a model of a perfectly time-synchronized OFDM system, where the first fading impulse always has a zero-delay, $\tau_0 = 0$, and other fading impulses have delays that are uniformly and independently distributed over the length of the cyclic prefix. The impulse power-delay profile, $\theta(\tau_k) = Ce^{-\tau_k/\tau_{rms}}$, decays exponentially [9].

- **Uniform channel.** All impulses have the same average power and their delays are uniformly and independently distributed over the length of the cyclic prefix.

2.3 Scenario

Our scenario consists of a wireless 16-QAM OFDM system, designed for an outdoor environment, that is capable of carrying digital video. The system operates at 500 kHz bandwidth and is divided into 64 tones with a total symbol period of 136 μs, of which 8 μs is the cyclic prefix. One OFDM symbol thus consists of 68 samples ($N + L = 68$), four of which are contained in the cyclic prefix ($L = 4$). The uncoded data rate of the system is 1.9 MBit/sec. We assume that $\tau_{rms} = 1$ sample for the synchronized channel.

Chapter 3

Linear channel estimation across tones

In the following we present the LMMSE estimate of the channel attenuations $\mathbf{h}$ from the received vector $\mathbf{y}$ and the transmitted data $\mathbf{X}$. We assume that the received OFDM symbol contains data known to the estimator – either training data or receiver decisions.

The complexity reduction of the LMMSE estimator consists of two separate steps. In the first step we modify the LMMSE by averaging over the transmitted data, obtaining a simplified estimator. In the second step we reduce the number of multiplications required by applying the theory of optimal rank-reduction [10].

3.1 LMMSE estimation

The LMMSE estimate of the channel attenuations $\mathbf{h}$, in (2.1), given the received data $\mathbf{y}$ and the transmitted symbols $\mathbf{X}$ is [11]

$$\widehat{\mathbf{h}}_{lmmse} = \mathbf{R}_{hh}\left(\mathbf{R}_{hh} + \sigma_n^2\left(\mathbf{X}\mathbf{X}^H\right)^{-1}\right)^{-1}\widehat{\mathbf{h}}_{ls} \tag{3.1}$$

where

$$\widehat{\mathbf{h}}_{ls} = \mathbf{X}^{-1}\mathbf{y} = \left[\frac{y_0}{x_0}\ \frac{y_1}{x_1}\ \cdots\ \frac{y_{N-1}}{x_{N-1}}\right]^T \tag{3.2}$$

is the least-squares (LS) estimate of $\mathbf{h}$, σ_n^2 is the variance of the additive channel noise and $\mathbf{R}_{hh} = E\left\{\mathbf{h}\mathbf{h}^H\right\}$ is the channel autocorrelation. The superscript $(\cdot)^H$ denotes Hermitian transpose. In the following we assume, without loss of generality, that the variances of the channel attenuations in $\mathbf{h}$ are normalized to unity, *i.e.* $E\left\{|h_k|^2\right\} = 1$.

The LMMSE estimator (3.1) is of considerable complexity, since a matrix inversion is needed every time the training data in $\mathbf{X}$ changes. We reduce the complexity of this estimator by averaging over the transmitted data [1], *i.e.* we replace the term $(\mathbf{X}\mathbf{X}^H)^{-1}$ in (3.1) with its expectation $E\left\{(\mathbf{X}\mathbf{X}^H)^{-1}\right\}$. Assuming the same signal constellation on all tones and equal probability on all constellation points, we have $E\left\{(\mathbf{X}\mathbf{X}^H)^{-1}\right\} = E\left\{|1/x_k|^2\right\}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Defining the average signal-to-noise ratio as $\text{SNR} = E\left\{|x_k|^2\right\}/\sigma_n^2$, we obtain a simplified estimator

$$\widehat{\mathbf{h}} = \mathbf{R}_{hh}\left(\mathbf{R}_{hh} + \frac{\beta}{\text{SNR}}\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_{ls}, \tag{3.3}$$

where

$$\beta = E\left\{|x_k|^2\right\} E\left\{|1/x_k|^2\right\}$$

is a constant depending on the signal constellation. In the case of 16-QAM transmission, $\beta = 17/9$. Because $\mathbf{X}$ is no longer a factor in the matrix calculation, the inversion of $\mathbf{R}_{hh} + \frac{\beta}{\text{SNR}}\mathbf{I}$ does not need to be calculated each time the transmitted data in $\mathbf{X}$ changes. Furthermore, if $\mathbf{R}_{hh}$ and SNR are known beforehand or are set to fixed nominal values, the matrix $\mathbf{R}_{hh}(\mathbf{R}_{hh} + \frac{\beta}{\text{SNR}}\mathbf{I})^{-1}$ needs to be calculated only once. Under these conditions the estimation requires N multiplications per tone. To further reduce the complexity of the estimator, we proceed with low-rank approximations below.

3.2 Optimal low-rank approximations

Optimal rank reduction is achieved by using the singular value decomposition (SVD) [10]. The SVD of the channel autocovariance matrix is

$$\mathbf{R}_{hh} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \tag{3.4}$$

where $\mathbf{U}$ is a unitary matrix containing the eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix, containing the singular values $\lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_{N-1}$ on its diagonal[1]. In Appendix A it is shown that the optimal rank-p estimator is

$$\widehat{\mathbf{h}}_p = \mathbf{U}\mathbf{\Delta_p}\mathbf{U}^H\widehat{\mathbf{h}}_{ls} \tag{3.5}$$

where $\mathbf{\Delta}_p$ is a diagonal matrix with the values

$$\delta_k = \begin{cases} \frac{\lambda_k}{\lambda_k + \frac{\beta}{\text{SNR}}} & k = 0, 1, \ldots, p-1 \\ 0 & k = p, \ldots, N-1 \end{cases} \tag{3.6}$$

Viewing the orthonormal matrix $\mathbf{U}^H$ as a transform[2], the singular value λ_k of $\mathbf{R}_{hh}$ is the channel power (variance) contained in the k^{th} transform coefficient after transforming the LS estimate $\widehat{\mathbf{h}}_{ls}$. Since $\mathbf{U}$ is unitary, this transformation can be viewed as rotating the vector $\widehat{\mathbf{h}}_{ls}$ so that all its components are uncorrelated [10]. The dimension of the space of essentially time- and band-limited signals leads us to the rank needed in the low-rank estimator. In [12] it is shown that this dimension is about $2BT + 1$ where B is the one-sided bandwidth and T is the time interval of the signal. Accordingly, the magnitude of the singular values of $\mathbf{R}_{hh}$ should drop rapidly after about $L + 1$ large values, where L is the length of the cyclic prefix ($2B = 1/T_s$, $T = LT_s$ and $2BT + 1 = L + 1$).

We present the channel power contained in the first 15 coefficients in Figure 3.1. The calculations are based on our scenario and the two channel models: the synchronized and the uniform. The magnitude of the channel power drops rapidly after about $k = 4$, *i.e.* 5 coefficients, which is consistent with the observation that the dimension of the space spanned by $\mathbf{R}_{hh}$ is approximately $L + 1$, that is, $4 + 1 = 5$ in this case.

A block diagram of the rank$-p$ estimator in (3.5) is shown in Figure 3.2, where the LS-estimate is calculated from $\mathbf{y}$ by multiplying by $\mathbf{X}^{-1}$.

[1] Since we are dealing with Hermitian matrices the λ_ks are also eigenvalues. However, we use the terminology of the SVD since it is more general and can be used in optimal rank reduction of non-Hermitian matrices.

[2] The transform in this special case of low-rank approximation is the Karhunen-Loeve (a.k.a. Hotelling) transform of $\mathbf{h}$.

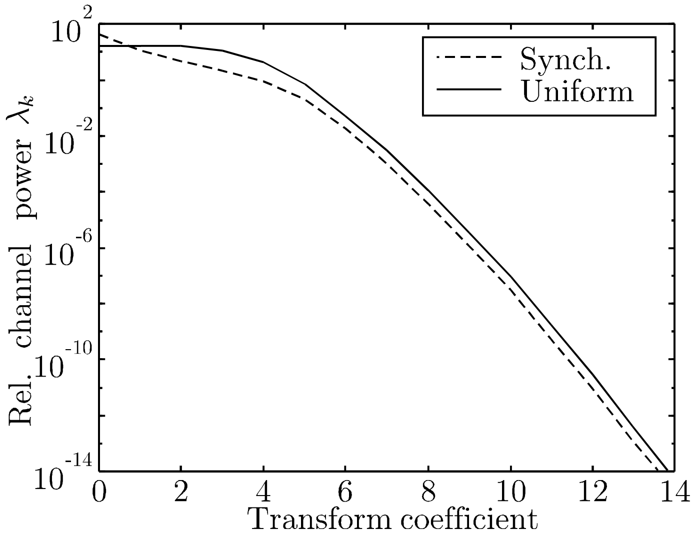

Figure 3.1: Relative channel power, $\lambda_k/\mathrm{E}|h_k|^2$, in the transform coefficients for the two example channels.

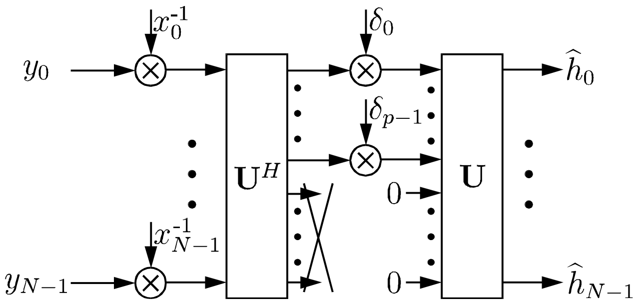

Figure 3.2: Block diagram of the rank$-p$ channel estimator.

3.3 Estimator complexity

The limiting factor of the rank$-p$ estimators is an error floor, see Section 4. To eliminate this error floor up to a given SNR we need to make sure our estimator rank is large enough. This prompts an analysis of the computational complexity of the rank$-p$ estimator. The implementation we have chosen is based on writing (3.5) as a sum of rank-1 matrices, which gives us the expression

$$\widehat{\mathbf{h}}_p = \left(\sum_{k=0}^{p-1} \delta_k \mathbf{u}_k \mathbf{u}_k^H \right) \widehat{\mathbf{h}}_{ls} = \sum_{k=0}^{p-1} \mathbf{q}_k \left\langle \mathbf{u}_k, \widehat{\mathbf{h}}_{ls} \right\rangle \tag{3.7}$$

where $\mathbf{q}_k = \delta_k \mathbf{u}_k$ and $\left\langle \mathbf{u}_k, \widehat{\mathbf{h}}_{ls} \right\rangle = \mathbf{u}_k^H \widehat{\mathbf{h}}_{ls}$ is the Euclidian inner product. The linear combination of p vectors of length N also requires pN multiplications. The estimation thus requires $2pN$ multiplications and the total number of multiplications per tone becomes $2p$. In comparison with the full estimator (3.3), we have managed to reduce the number of multiplications from N to $2p$ per tone. The smaller p is, the lower the computational complexity, but the larger the approximation error becomes. Following the analysis in Section 3.2, we can expect a good approximation when p is in the range of samples in the cyclic prefix, which is usually much smaller than the number of tones, N.

A legitimate question at this point is what happens for a system with many tones and many samples in the cyclic prefix. The number of calculations per tone can be considerable if a rank$-p$ estimator is used directly on all tones in the system. One solution to this problem is a partitioning of the tones into reasonable sized blocks and, at a certain performance loss, perform the estimation independently in these blocks. By dividing the channel attenuations into K equally sized blocks, the bandwidth in each block is reduced by a factor K. Referring again to the dimension of the space of essentially time- and bandlimited signals [12], the expected number of essential base vectors is reduced from $L + 1$ to $L/K + 1$. Hence the complexity of the estimator decreases accordingly.

To illustrate the idea, let us assume we have a system with $N = 1024$ tones and a $L = 64$ sample cyclic prefix. The uniform channel correlation between the attenuations h_m and h_n in this system is, see Appendix B,

$$r_{m,n} = \begin{cases} 1 & \text{if} \quad m = n \\ \frac{1-e^{-j2\pi L \frac{m-n}{N}}}{j2\pi L \frac{m-n}{N}} & \text{if} \quad m \neq n \end{cases} .$$

This only depends on the distance between the tones, $m - n$, and the ratio between the length of the cyclic prefix and the number of tones, L/N. The 1024 tone system can be described by

$$\begin{bmatrix} \mathbf{y}^{(1)} \\ \vdots \\ \mathbf{y}^{(16)} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} & & \\ & \ddots & \\ & & \mathbf{X}^{(16)} \end{bmatrix} \begin{bmatrix} \mathbf{h}^{(1)} \\ \vdots \\ \mathbf{h}^{(16)} \end{bmatrix} + \begin{bmatrix} \mathbf{n}^{(1)} \\ \vdots \\ \mathbf{n}^{(16)} \end{bmatrix},$$

that is, as 16 parallel 64-tone systems,

$$\mathbf{y}^{(k)} = \mathbf{X}^{(k)}\mathbf{h}^{(k)} + \mathbf{n}^{(k)}, \;\; k = 1, 2, \ldots 16.$$

We have the same channel correlation in each subsystem as we have in the 64-tone scenario in this paper ($L/N = 4/64 = 64/1024$). By estimating the channel attenuations $\mathbf{h}^{(k)}$ in each sub-system independently, we neglect the correlation between tones in different sub-systems, but obtain the same MSE performance as in our 64-tone scenario.

Chapter 4

Estimator performance and design

We propose a generic low-rank frequency-based channel estimator, *i.e.* the estimator is designed for fixed, nominal values of SNR and channel correlation. Hence, we need to analyse how the rank, channel correlation and SNR should be chosen for this estimator so that it is robust to variations in the channel statistics, *i.e.* mismatch. As a performance measure, we use uncoded symbol-error rate (SER) for 16-QAM signalling. The SER in this case can be calculated from the mean-squared error (MSE) with the formulae in [13].

4.1 Rank reduction

The mean-squared error, relative to the channel power $E\left\{|h_k|^2\right\}$, of the rank$-p$ estimator is mainly determined by the channel power contained in the transform coefficients and can be expressed, see Appendix C,

$$\mathrm{mse}\,(p) = \frac{1}{N}\sum_{k=0}^{p-1}\left(\lambda_k\,(1-\delta_k)^2 + \frac{\beta}{\mathrm{SNR}}\delta_k^2\right) + \frac{1}{N}\sum_{k=p}^{N-1}\lambda_k \tag{4.1}$$

where λ_k and δ_k are given by (3.4) and (3.6) respectively. The MSE (4.1) is a monotonically decreasing function of SNR and can be bounded from below by the last term,

$$\underline{\mathrm{mse}}(p) = \frac{1}{N}\sum_{k=p}^{N-1}\lambda_k \leq \mathrm{mse}\,(p)\,, \tag{4.2}$$

which is the sum of the channel power in the transform coefficients not used in the estimate. This MSE-floor, $\underline{\mathrm{mse}}(p)$, will give rise to a error floor in the symbol-error rates.

The error floor is the main limitation on the complexity reduction achieved by optimal rank reduction. As an illustration, Figure 4.1 displays the SER relative to the channel variance, for three different ranks, as a function of the SNR. The ranks chosen are $p = 5$, 6 and 7, and the channel used in the example is the synchronized channel. The corresponding SER-floors are shown as horizontal lines. For $p = 7$, the SER-floor is relatively small, and the SER of the rank-7 estimator is comparable to the original, full-rank estimator (3.3) in the range 0 to 30 dB in SNR. By choosing the appropriate rank on the estimator, we can essentially avoid the impact from the SER-floor up to a given SNR. When we have full rank, $p = N$, no SER-floor exists.

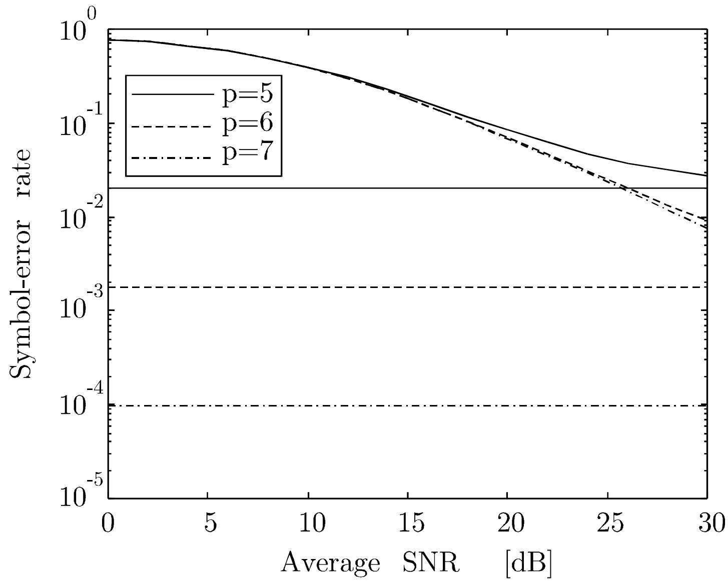

Figure 4.1: Low-rank estimator symbol-error rate as a function of SNR, with ranks $p = 5$, 6 and 7. Corresponding SER-floors shown as horizontal lines. (Synchronized channel)

Based on the channel powers presented in Figure 3.1, we show the corresponding SER-floors, relative to the channel variance, in Figure 4.2. After about rank−4 the SER-floor decreases rapidly. We are therefore able to obtain a good estimator approximation with a relatively low rank.

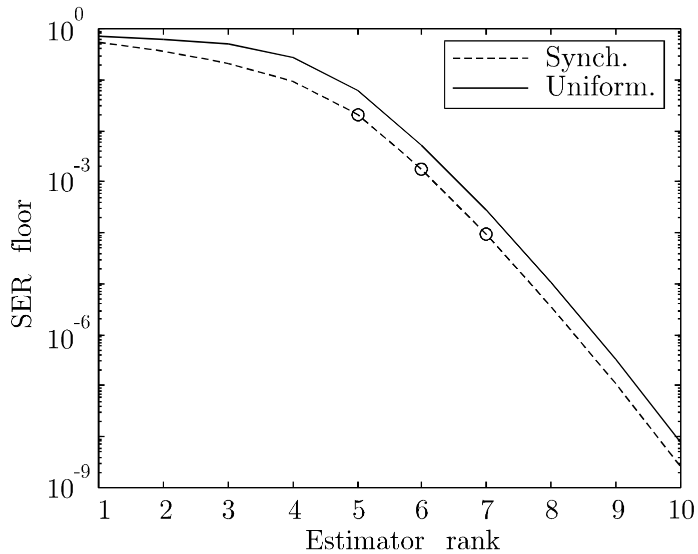

Figure 4.2: Estimator SER-floor as a function of estimator rank. Circles show the SER-floors appearing in Figure 4.1.

4.2 SER performance under mismatch

In practice, the true channel correlation and SNR are not known. To get a general expression for the estimator SER, we derive it under the assumption that the estimator is designed for correlation $\mathbf{R}_{hh}$ and signal-to-noise ratio SNR, but the true values are $\mathbf{R}_{\widetilde{h}\widetilde{h}}$ and $\widetilde{\mathrm{SNR}}$, respectively, where $\widetilde{\mathbf{h}}$ denotes a channel with different statistics than $\mathbf{h}$. This allows us to analyse this estimator's sensitivity to design errors. Under these assumptions, the relative MSE of the rank$-p$ estimate (3.5) becomes, see Appendix C,

$$\mathrm{mse}\,(p) = \frac{1}{N}\sum_{k=0}^{p-1}\left[\mu_k\left(1-\delta_k\right)^2 + \frac{\beta}{\widetilde{\mathrm{SNR}}}\delta_k^2\right] + \frac{1}{N}\sum_{k=p}^{N-1}\mu_k \tag{4.3}$$

where μ_k is the k^{th} diagonal element of $\mathbf{U}^H\mathbf{R}_{\widetilde{h}\widetilde{h}}\mathbf{U}$, cf. (3.4). It can be interpreted as the variance of the transformed channel, $\mathbf{U}^H\widetilde{\mathbf{h}}$ under correlation mismatch since

$$E\left\{\left(\mathbf{U}^H\widetilde{\mathbf{h}}\right)\left(\mathbf{U}^H\widetilde{\mathbf{h}}\right)^H\right\} = \mathbf{U}^H\mathbf{R}_{\widetilde{h}\widetilde{h}}\mathbf{U}$$

It should be noted that the elements of $\mathbf{U}^H\widetilde{\mathbf{h}}$ are no longer uncorrelated. However due to the fact that the power-delay profile is short compared to the OFDM symbol, the first p elements can be expected to contain most of the power. This property will ensure only a small performance loss when the estimator is design for wrong channel statistics.

If rank$-p$ estimators are used in a real system, the sensitivity to mismatch in both channel correlation and SNR are important. We will show that a rank$-p$ estimator based on the uniform channel model and a nominal SNR can be used as fixed generic estimator with only a small loss in average performance. We divide the mismatch analysis into two parts: first we analyse the SER when we have a mismatch in channel correlation and later we analyse the SER when we have a mismatch in SNR.

4.2.1 Incorrect channel correlation

From (4.3), with no SNR mismatch ($\mathrm{SNR} = \widetilde{\mathrm{SNR}}$), but incorrect channel correlation, ($\mathbf{R}_{hh} \neq \mathbf{R}_{\widetilde{h}\widetilde{h}}$), we obtain the performance for the correlation mismatch cases. We compared the performance of our channel estimator in two mismatch situations: i) using the a uniform channel when the true channel model was the synchronous channel and ii) using the synchronous channel when the true channel model was the uniform channel. The resulting channel estimates that were used in the detection of the data produced no noticable difference in symbol error rates – less than 0.1 dB change in effective SNR for an average SNR up to 20 dB. However, when both the channel SNR and the channel correlation matrix are mismatched, the nominal design SNR becomes more important. This can be seen in Figure 8, where we present the resulting symbol error rate for rank-8 estimators. For the mismatched cases, marked with 'o', the uniform design is more robust, *i.e.* the error in case of mismatch is lower. With the restriction that the true channel has a power-delay profile shorter than the cyclic prefix, designing for a uniform power-delay profile can be seen as a minimax design.

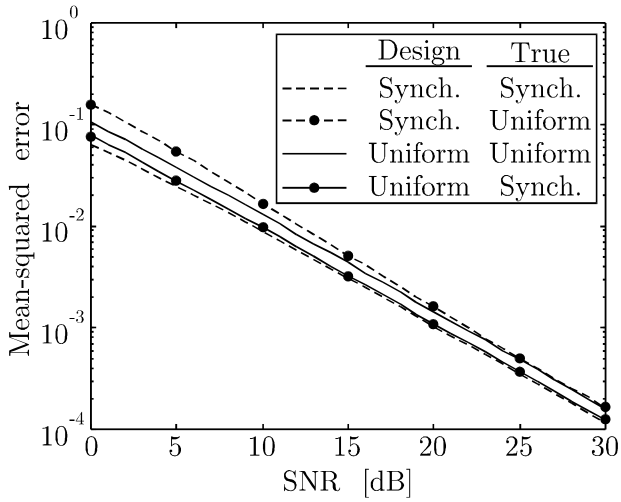

Figure 4.3: MSE for correct and mismatched design. The latter is marked with circles (◯).

4.2.2 Incorrect SNR

Finally we evaluate the sensitivity to mismatch in design SNR for a rank-8 estimator. When there is no mismatch in channel correlation, and nominal SNRs of 10, 20 and 30 dB are used in the design, the sensitivity to SNR mismatch is not that large. However, in Figure 4.4, we present the SER for the same rank-8 estimators, but with the difference that the true channel correlation is mismatched with the design correlation. In this second case, there is a clear difference between the two designs: the higher the nominal design SNR, the better the overall performance of the estimator in the range 0 to 30 dB in SNR. It should be noted that a LMMSE-estimator designed for a large SNR approaches the LS-estimator.

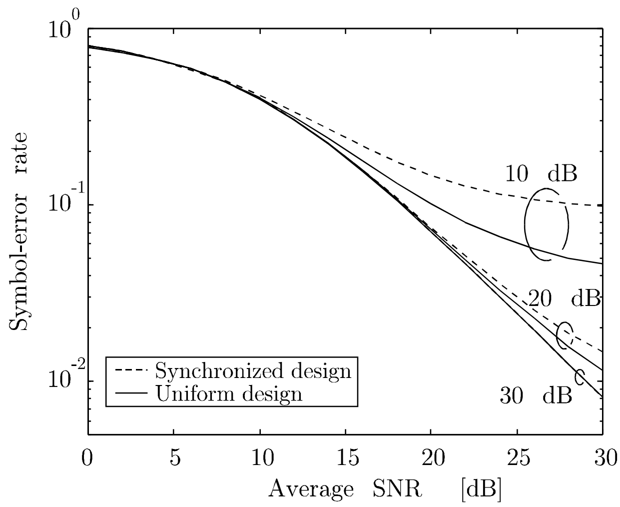

Figure 4.4: Rank-8 estimator SER when SNRs of 10, 20 and 30 dB are used in the design. The estimators are designed for incorrect channel correlation.

Chapter 5

Generic low-rank estimator

If we want a robust generic channel estimator design for OFDM systems, of the low-rank type, the analysis in the previous section suggests the use of the uniform channel correlation and a relatively high SNR as nominal design parameters. The design of such an estimator only requires knowledge about the length of the cyclic prefix, the number of tones in the system and the target range of SNRs for the application. If the receiver cannot afford an estimator that includes tracking of channel correlation and SNR, this channel estimator works reasonably well for fixed SNR and channel correlation.

5.1 Performance gain

For the scenario used in this paper, Sec. 2.3, we choose a rank-8 estimator with uniform design and SNR = 30 dB. The performance of this estimator is presented in Fig. 5.1, where the SER for the LS-estimate (3.2) and known channel are also shown. As can be seen, the low-rank estimator is 3.5 dB better than the LS-estimator and less than 1 dB from the known channel.

5.2 Comparison to FIR-filters

An alternative to using low-rank estimators to smooth the channel estimates is to use a FIR-filter instead. Hence we will compare our proposed low-rank estimators to FIR-filters of the same complexity. The FIR-filters are $2p$-taps Wiener filters [10], *i.e.* $2p$ multiplications per tone that are designed for the same channel correlation and SNR as the low-rank estimators. Figure 5.2 shows the SER for rank$-p$ estimators in comparison with FIR-filters of the same computational complexity. When the complexity is 16 multiplications per tone (A) the rank$-p$ estimator has about 0.2 dB advantage in SNR over the FIR-filter in the range of SNRs shown. When the number of calculations goes down to 12 multiplications per tone (B) the SER-floor of the rank$-p$ estimator becomes visible and the FIR-filter performs better at SNRs above 20 dB.

However, it should be noted that the performance of the low-rank estimators depend heavily of the size of the cyclic prefix. If the cyclic prefix were to be decreased (relative to the OFDM symbol), the low-rank estimator would increase its performance. This is due to the fact that the "dimension" of the channel (whose duration is assumed to be shorter than the cyclic prefix) decreases and can thus be represented with fewer coefficients. On the other hand, if the cyclic

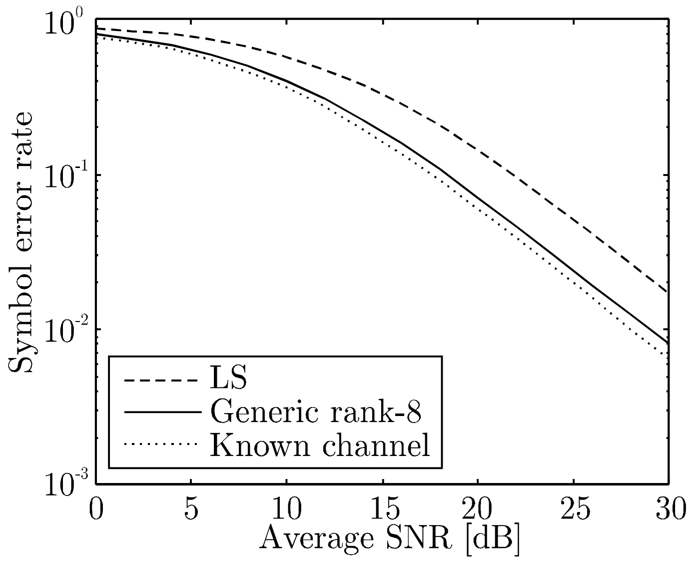

Figure 5.1: SER for 16-QAM training data and a synchronized channel. The generic rank−8 estimator, designed for a uniform channel and 30 dB in SNR, is compared to the LS estimator and known channel at the receiver.

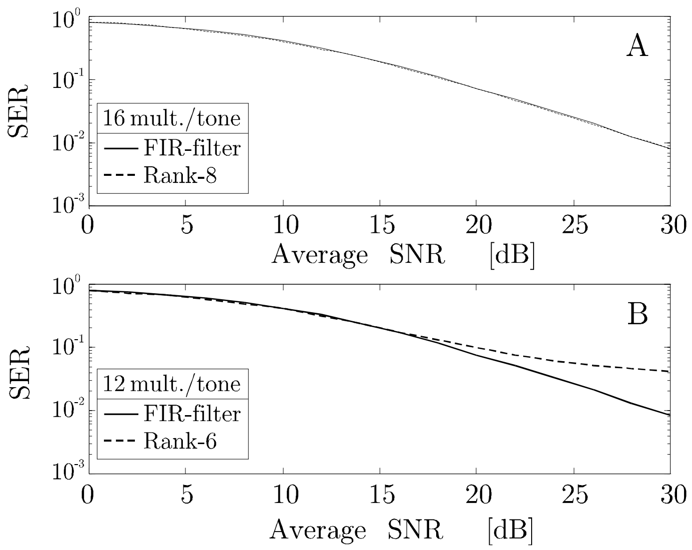

Figure 5.2: SER comparison between the rank−p estimators and FIR Wiener-filters of the same complexity. Both estimators are designed for the uniform channel and 30 dB SNR. A) 16 multiplications per tone and B) 12 multiplications per tone.

prefix increases in size, more coefficients are needed to avoid large approximation errors. Hence, whether or not the low-rank estimator is better than the FIR-filter depends on the relative size of the cyclic prefix and the allowed complexity.

5.3 The use of time correlation

The low-rank estimator presented in this paper is based on frequency correlation only but the time-correlation of the channel can also be used. The two-dimensional LMMSE estimator can be simplified using the same technique with rank reduction as described here. However, in [14] it is shown that such an estimator gives an inferior performance for a fixed complexity. Hence, it seems that separating the use of frequency- and time-correlation is the most efficient way of estimating the channel.

Other approaches to use the time-correlation is *e.g.* to use a decision-directed scheme [13] or FIR-filters [3, 15]. The former can be used in a slow-fading environment, where it offers good performance for a minimal complexity and the latter is preferred in case of fast fading. It is possible to use a bank of FIR-filters and choose the most appropriate according the estimated Doppler frequency [16].

Chapter 6

Conclusions

We have investigated low-complexity low-rank approximations of the LMMSE channel estimator for non-sample-spaced channels. The investigation shows that an estimator error-floor, inherent in the low-rank approximation, is the significant limitation to the achieved complexity reduction. We showed that a generic low-rank estimator design, based on the uniform channel correlation and a nominal SNR, can be used in our 64-tone scenario. Compared with the full LMMSE (3.3), there is only a small loss in performance up to a SNR of 30 dB but a reduction in complexity with a factor $N/2p = 4$. For systems with more subchannels this gain is even larger. The generic estimator design only requires knowledge about the length of the cyclic prefix, the number of tones in the system and the target range of SNRs for the application.

We also compared low-rank estimators to FIR-filters across the tones. The comparison showed that at low complexities and high SNRs the FIR-filters is the preferable choice, due to the error floor in the low-rank approximation. However, if we can allow up to 16 multiplications per tone in our scenario, the low-rank estimator is more advantageous. Also, the low-rank estimators improve their performance as the cyclic prefix decreases in size.

Appendix A

Optimal rank reduction

The optimal rank reduction is found from the correlation matrices

$$\begin{aligned}
\mathbf{R}_{h\widehat{h}_{ls}} &= E\left\{\mathbf{h}\widehat{\mathbf{h}}_{ls}^H\right\} = \mathbf{R}_{hh} \\
\mathbf{R}_{\widehat{h}_{ls}\widehat{h}_{ls}} &= E\left\{\widehat{\mathbf{h}}_{ls}\widehat{\mathbf{h}}_{ls}^H\right\} = \mathbf{R}_{hh} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}
\end{aligned}$$

and the SVD

$$\mathbf{R}_{h\widehat{h}_{ls}}\mathbf{R}_{\widehat{h}_{ls}\widehat{h}_{ls}}^{-1/2} = \mathbf{Q}_1\mathbf{D}\mathbf{Q}_2^H \tag{A.1}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are unitary matrices and $\mathbf{D}$ is a diagonal matrix with the singular values $d_0 \geq d_1 \geq \cdots \geq d_{N-1}$ on its diagonal. The best low-rank estimator [10] is then

$$\widehat{\mathbf{h}}_p = \mathbf{Q}_1 \left[\begin{array}{cc} \mathbf{D}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right] \mathbf{Q}_2^H \mathbf{R}_{\widehat{h}_{\mathrm{ls}}\widehat{h}_{\mathrm{ls}}}^{-1/2}\widehat{\mathbf{h}}_{ls}, \tag{A.2}$$

where $\mathbf{D}_p$ is the $p \times p$ upper left corner of $\mathbf{D}$, *i.e.* we exclude all but the p largest singular vectors. In this paper we have $\mathbf{R}_{h\widehat{h}_{ls}} = \mathbf{R}_{hh}$ and $\mathbf{R}_{\widehat{h}_{ls}\widehat{h}_{ls}} = \mathbf{R}_{hh} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}$ and we note that they share the same singular vectors, *i.e.* the ones of $\mathbf{R}_{hh} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$. Thus, we may express (A.1) as

$$\begin{aligned}
&\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \left(\mathbf{U}\left(\mathbf{\Lambda} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)\mathbf{U}^H\right)^{-1/2} = \\
&\mathbf{U}\mathbf{\Lambda}\left(\mathbf{\Lambda} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)^{-1/2}\mathbf{U}^H = \mathbf{Q}_1\mathbf{D}\mathbf{Q}_2^H \\
\Rightarrow\ &\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{U} \text{ and } \mathbf{D} = \mathbf{\Lambda}\left(\mathbf{\Lambda} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)^{-1/2}
\end{aligned}$$

The rank-p estimator (A.2) now becomes

$$\begin{aligned}
\widehat{\mathbf{h}}_p &= \mathbf{U}\left[\begin{array}{cc} \mathbf{D}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right]\mathbf{U}^H\left(\mathbf{U}\left(\mathbf{\Lambda} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)\mathbf{U}^H\right)^{-1/2}\widehat{\mathbf{h}}_{ls} = \\
&= \mathbf{U}\left[\begin{array}{cc} \mathbf{D}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right]\left(\mathbf{\Lambda} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)^{-1/2}\mathbf{U}^H\widehat{\mathbf{h}}_{ls} = \mathbf{U}\left[\begin{array}{cc} \mathbf{\Delta}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right]\mathbf{U}^H\widehat{\mathbf{h}}_{ls}
\end{aligned}$$

where $\mathbf{\Delta}_p$ is the $p \times p$ upper left corner of

$$\mathbf{\Delta} = \mathbf{\Lambda} \left(\mathbf{\Lambda} + \frac{\beta}{\mathrm{SNR}} \mathbf{I} \right)^{-1} = \mathrm{diag} \left(\frac{\lambda_0}{\lambda_0 + \frac{\beta}{\mathrm{SNR}}}, \cdots, \frac{\lambda_{N-1}}{\lambda_{N-1} + \frac{\beta}{\mathrm{SNR}}}, \right).$$

Note that $\mathbf{Q}_1 = \mathbf{Q}_2$ since we are estimating the same tones as we are observing (*i.e.* smoothing) and an eigenvalue decomposition could be used to achieve optimal rank reduction. In the general case when *e.g.* pilot-symbol assisted modulation [15] is used and there are known symbols (pilots) on only a part of the subchannels, we have $\mathbf{Q}_1 \neq \mathbf{Q}_2$ since $\mathbf{R}_{h\widehat{h}_{ls}}$ and $\mathbf{R}_{\widehat{h}_{ls}\widehat{h}_{ls}}$ don't share the same singular vectors (the matrices are not even of the same size). Hence, the more general SVD must be used which motivates the nomenclature in this article.

Appendix B

Channel-correlation matrices

Using the channel model in (2.2), the attenuation on tone k becomes

$$h_k = \sum_{i=0}^{M-1} \alpha_i e^{-j2\pi \frac{k}{N} \tau_i},$$

and the correlation matrix for the attenuation vector, $\mathbf{h}$,

$$\mathbf{R_{hh}} = E\left\{\mathbf{hh}^H\right\} = [r_{m,n}]$$

can be expressed as (τ_k's independent)

$$\begin{aligned} r_{m,n} &= \int \cdots \int \prod_{k=0}^{M-1} f_{\tau_k}(\tau_k) \left[\sum_{i=0}^{M-1} \theta(\tau_i) e^{-j2\pi\tau_i \frac{m-n}{N}}\right] d\tau_0 \ldots d\tau_{M-1} \\ &= \sum_{i=0}^{M-1} \int f_{\tau_i}(\tau_i)\, \theta(\tau_i)\, e^{-j2\pi\tau_i \frac{m-n}{N}} d\tau_i, \end{aligned} \tag{B.1}$$

where $\theta(\tau)$ is the multi-path intensity profile and $f_{\tau_k}(\tau_k)$ is the probability density function of τ_k.

The correlation matrices of the three channels used in this paper are calculated below.

- **Synchronized channel**.

  The probability distributions for the delays are

$$\begin{aligned} f_{\tau_0}(\tau_0) &= \delta(\tau_0), \\ f_{\tau_i}(\tau_i) &= \begin{cases} 1/L & \text{if } \tau_i \in [0, L] \\ 0 & \text{otherwise} \end{cases}, i = 1, 2, \ldots, M, \end{aligned}$$

  and the power-delay profile is $\theta(\tau) = C_1 e^{-\tau/\tau_{rms}}$. Substituting in (B.1), and normalizing $r_{k,k}$ to unity, gives us

$$r_{m,n} = \frac{L + \frac{M-1}{\left(\frac{1}{\tau_{rms}} + j2\pi\frac{m-n}{N}\right)}\left(1 - e^{-L\left(\frac{1}{\tau_{rms}} + 2\pi j\frac{m-n}{N}\right)}\right)}{L + (M-1)\,\tau_{rms}\left(1 - e^{-\frac{L}{\tau_{rms}}}\right)}.$$

- **Uniform channel**.

  The probability distributions for the delays are

$$f_{\tau_i}\left(\tau_i\right) = \begin{cases} 1/L & \text{if } \tau_i \in [0, L] \\ 0 & \text{otherwise} \end{cases}, \; i = 1, 2, \ldots, M,$$

  and the power-delay profile is constant $\theta\left(\tau\right) = C_2$. Substituted in (B.1), and normalizing $r_{k,k}$ to unity, gives us

$$r_{m,n} = \begin{cases} 1 & \text{if} \quad m = n \\ \frac{1-e^{-j2\pi L \frac{m-n}{N}}}{j2\pi L \frac{m-n}{N}} & \text{if} \quad m \neq n \end{cases}.$$

Appendix C

Estimator mean-squared error

In this appendix we derive the MSE of the rank$-p$ estimator in (3.5). We also present the MSE-floor, which bounds the achievable MSE from below in low-rank approximations of the LMMSE estimator. To get a general expression for the mean-squared error for the rank$-p$ approximation of the LMMSE estimator, we assume that the estimator has been designed for channel correlation $\mathbf{R}_{hh}$ and signal-to-noise ratio SNR, but the real channel $\widetilde{\mathbf{h}}$ has the correlation $\widetilde{\mathbf{R}}_{hh}$ and the real signal-to-noise ratio is $\widetilde{\mathrm{SNR}}$. From (2.1) and (3.2), we have $\widehat{\mathbf{h}}_{ls} = \widetilde{\mathbf{h}} + \widetilde{\mathbf{n}}$, where the noise term $\widetilde{\mathbf{n}} = \mathbf{X}^{-1}\mathbf{n}$ has the autocovariance matrix $\mathbf{R}_{\widetilde{n}\widetilde{n}} = \frac{\beta}{\widetilde{\mathrm{SNR}}}\mathbf{I}$. The estimation error $\mathbf{e}_p = \widetilde{\mathbf{h}} - \widehat{\mathbf{h}}_p$ of the rank$-p$ estimator (3.5) is

$$\mathbf{e}_p = \mathbf{U}\left(\mathbf{I} - \begin{bmatrix} \mathbf{\Delta}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)\mathbf{U^H}\widetilde{\mathbf{h}} - \mathbf{U}\begin{bmatrix} \mathbf{\Delta}_p & 0 \\ 0 & 0 \end{bmatrix}\mathbf{U^H}\widetilde{\mathbf{n}}, \tag{C.1}$$

and the mean-squared error is

$$\mathrm{mse}\,(p) = \frac{1}{N}\,\mathrm{Trace}\,E\left\{\mathbf{e}_p\mathbf{e}_p^H\right\}, \tag{C.2}$$

To simplify the expression we use that:

- $\widetilde{\mathbf{h}}$ and $\widetilde{\mathbf{n}}$ are uncorrelated, hence the cross terms are cancelled in the expectation.
- $\mathrm{Trace}\left(\mathbf{UAU}^H\right) = \mathrm{Trace}\,\mathbf{A}$ if $\mathbf{U}$ is a unitary matrix, and $\mathrm{Trace}\,(\mathbf{A}+\mathbf{B}) = \mathrm{Trace}\,\mathbf{A} + \mathrm{Trace}\,\mathbf{B}$ [17].
- $\mathrm{Trace}\,(\mathbf{DAD}) = \sum_k a_{k,k} d_k^2$ when $\mathbf{D}$ is a diagonal matrix with the elements d_k on its diagonal and $\mathbf{A}$ (not necessarily a diagonal matrix) has diagonal elements $a_{k,k}$.

Using (C.1) in (C.2), the mean-squared error becomes

$$\begin{aligned} \mathrm{mse}\,(p) \;=\; & \frac{1}{N}\,\mathrm{Trace}\left[\mathbf{U}\left(\mathbf{I} - \begin{bmatrix} \mathbf{\Delta}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)\mathbf{U^H}\mathbf{R}_{\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}}\mathbf{U}\left(\mathbf{I} - \begin{bmatrix} \mathbf{\Delta}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)^H\mathbf{U}^H + \right. \\ & \left. \mathbf{U}\begin{bmatrix} \mathbf{\Delta}_p & 0 \\ 0 & 0 \end{bmatrix}\mathbf{U^H}\mathbf{R}_{\widetilde{\mathbf{n}}\widetilde{\mathbf{n}}}\mathbf{U}\begin{bmatrix} \mathbf{\Delta}_p & 0 \\ 0 & 0 \end{bmatrix}^H\mathbf{U^H}\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{N}\left(\sum_{k=0}^{p-1}\mu_k\left(1-\delta_k\right)^2+\sum_{k=p}^{N-1}\mu_k\right)+\frac{1}{N}\sum_{k=0}^{p-1}\frac{\beta}{\widetilde{\mathrm{SNR}}}\delta_k^2 \\
&= \frac{1}{N}\sum_{k=0}^{p-1}\left(\mu_k\left(1-\delta_k\right)^2+\frac{\beta}{\widetilde{\mathrm{SNR}}}\delta_k^2\right)+\frac{1}{N}\sum_{k=p}^{N-1}\mu_k, \qquad \text{(C.3)}
\end{aligned}
$$

where μ_k is the channel power in the kth transform coefficient, *i.e.*, the kth diagonal element of the matrix $\mathbf{U}^H\widetilde{\mathbf{R}}_{hh}\mathbf{U}$. The MSE can be lower-bounded, $\mathrm{mse}\,(p) \geq \underline{\mathrm{mse}}\,(p)$, by what we call the MSE-floor

$$
\underline{\mathrm{mse}}\,(p) = \frac{1}{N}\sum_{k=p}^{N-1}\mu_k.
$$

If there is no mismatch in SNR or channel correlation, we have $\mu_k = \mathrm{diag}\left(\mathbf{U}^H\mathbf{R}_{hh}\mathbf{U}\right) = \lambda_k$ and $\widetilde{\mathrm{SNR}} = \mathrm{SNR}$, and the MSE becomes

$$
\mathrm{mse}\,(p) = \frac{1}{N}\sum_{k=0}^{p-1}\left(\lambda_k\left(1-\delta_k\right)^2+\frac{\beta}{\mathrm{SNR}}\delta_k^2\right)+\frac{1}{N}\sum_{k=p}^{N-1}\lambda_k.
$$

Bibliography

[1] J.G. Proakis. *Digital communications*. Prentice-Hall, 3rd edition, 1995.

[2] Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers. ETS 300 401, ETSI – European Telecommunications Standards Institute, Valbonne, France, February 1995.

[3] Peter Höher. TCM on frequency-selective land-mobile fading channels. In *Proc. Tirrenia Int. Workshop Digital Commun.*, Tirrenia, Italy, September 1991.

[4] Sarah Kate Wilson, R. Ellen Khayata, and John M. Cioffi. 16-QAM modulation with orthogonal frequency-division multiplexing in a Rayleigh-fading environment. In *Proc. IEEE Vehic. Technol. Conf.*, volume 3, pages 1660–1664, Stockholm, Sweden, June 1994.

[5] Ahmad Chini. *Multicarrier modulation in frequency selective fading channels*. PhD thesis, Carleton University, Ottawa, Canada, 1994.

[6] John M. Cioffi. Personal communication, 1994.

[7] Jan-Jaap van de Beek, Ove Edfors, Magnus Sandell, Sarah Kate Wilson, and Per Ola Börjesson. On channel estimation in OFDM systems. In *Proc. IEEE Vehic. Technol. Conf.*, volume 2, pages 815–819, Chicago, IL, July 1995.

[8] A. Peled and A. Ruiz. Frequency domain data transmission using reduced computational complexity algorithms. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 964–967, Denver, CO, 1980.

[9] Peter Höher. A statistical discrete-time model for the WSSUS multipath channel. *IEEE Trans. Commun.*, 41(4):461–468, November 1992.

[10] Louis L. Scharf. *Statistical signal processing: Detection, estimation, and time series analysis*. Addison-Wesley, 1991.

[11] Ove Edfors, Magnus Sandell, Jan-Jaap van de Beek, Sarah Kate Wilson, and Per Ola Börjesson. Analysis of DFT-based channel estimators for OFDM. Research Report TULEA 1996:17, Div. of Signal Processing, Luleå University of Technology, September 1996.

[12] H. J. Landau and H. O. Pollak. Prolate spheriodal wave functions, Fourier analysis and uncertainty – III: The dimension of the space of essentially time- and band-limited signals. *Bell System Tech. J.*, 41:1295, 1962.

[13] Sarah Kate Wilson. *Digital audio broadcasting in a fading and dispersive channel.* PhD thesis, Stanford University, CA, August 1994.

[14] Magnus Sandell and Ove Edfors. A comparative study of pilot-based channel estimators for wireless OFDM. Research Report TULEA 1996:19, Div. of Signal Processing, Luleå University of Technology, September 1996.

[15] James K. Cavers. An analysis of pilot-symbol assisted modulation for Rayleigh-fading channels. *IEEE Trans. Vehic. Technol.*, 40(4):686–693, November 1991.

[16] Henry Li and James K. Cavers. An adaptive filtering technique for pilot-aided transmission systems. *IEEE Trans. Vehic. Technol.*, 40(3):532–545, August 1991.

[17] Gilbert Strang. *Linear Algebra and Its Applications.* Academic Press, 2nd edition, 1980.