Analysis of DFT-based channel estimators for OFDM

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

1996

Analysis of DFT-based channel estimators for OFDM*

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 1995 Vehicular Technology Conference (VTC'96) in Chicago, Illinois, Jul 25–28, 1995, pp 815-819.

Abstract

In this paper we analyze the performance of three low-complexity channel estimators, based on the discrete Fourier-transform (DFT), for orthogonal frequency-division multiplexing (OFDM) systems. Estimators of this type have been analyzed for discrete-time channels, and we extend this analysis to continuous-time channels. We present analytical expressions for their mean-squared error (MSE) and evaluate their complexity versus symbol-error rate (SER) for 16-QAM. The analysis shows that this type of estimators may experience an irreducible error floor at high SNRs. However, in one of the three estimators the error floor can be eliminated while the complexity stays low and the performance is maximized.

Contents

Chapter 1

Introduction

Wireless digital communication systems using coherent signaling schemes, such as a 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 [2, 3], DPSK is appropriate for relatively low data rates, such as in the European digital-audio broadcast (DAB) system [4]. However, for more spectrally-efficient OFDM systems, coherent modulation is more appropriate.

We address linear estimators for OFDM where all channel attenuations in a received symbol are estimated simultaneously. Using the linear minimum mean-squared error (LMMSE) estimator, which takes advantage of the correlation between all N subcarriers, requires an $N \times N$ matrix multiplication. This complexity can be large depending on the number of subcarriers in the system. This paper presents and analyzes three low-complexity, suboptimal, approximations of the LMMSE channel estimator. These estimators all share the property that they use the discrete Fourier transform (DFT) to estimate the channel in the time domain. Estimators of this type have been proposed [5, 6, 7, 8], but only analyzed for discrete-time channels [6].

The addressed estimators take the N noisy frequency-domain observations and transform them to the time domain by an inverse DFT (IDFT). The linear estimation is then performed in the time-domain, and the result transformed back to the frequency domain by a DFT. The transforms can be implemented with fast algorithms requiring only a few multiplications per estimated attenuation, but there are still N coefficients to estimate simultaneously in the time-domain. However, an OFDM symbol time is, by design, much larger than the length of the channel. The time-domain estimation takes advantage of the fact that this concentrates the channel power to a relatively small number of time-domain samples. Three of the strategies for doing time-domain approximations are; approximating time-domain samples with low channel power as zero, ignoring cross correlations and ignoring differences in variance. The three estimators analyzed here use these three strategies cumulatively.

After presenting the OFDM system model in Section 2, we introduce the three DFT-based estimators in Section 3. In Section 4 we present an analysis of the average mean-squared error (MSE) and show that there is an irreducible MSE-floor inherent in DFT-based low-complexity estimators. We also illustrate their performance by presenting the uncoded 16-QAM symbol error rate for a 64 tone OFDM system. A complexity versus performance comparison is done, which singles out the second of the three estimators as a good trade-off. A summary and concluding remarks appear in Section 5.

Chapter 2

System model

Figure 2.1 displays the OFDM base-band system used in this paper. We assume that the use of a cyclic prefix (CP) both preserves the orthogonality of the subcarriers and eliminates inter-symbol interference (ISI) between consecutive OFDM symbols [9]. Further, the channel $g(\tau;t)$ is assumed to be slowly Rayleigh-fading and considered constant during one OFDM symbol. The number of tones in the system is N, and the length of the cyclic prefix is L samples.

Under these assumptions we can describe the system as a set of parallel Gaussian channels [2], shown in Figure 2.2, with correlated attenuations

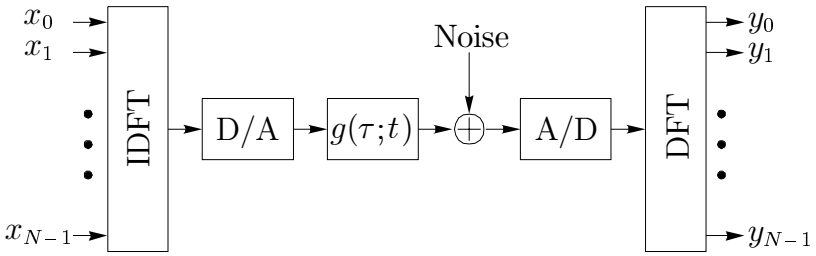

Figure 2.1: Base band OFDM system. The cyclic prefix is not shown in this figure.

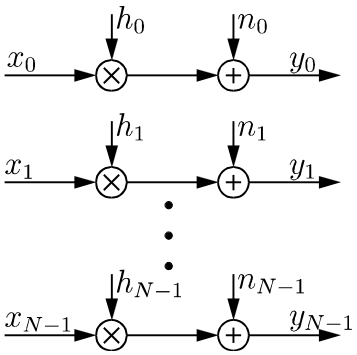

Figure 2.2: The OFDM system, described as a set of parallel Gaussian channels with correlated attenuations.

$$h_k = G\left(\frac{k}{NT_s}\right),\ k = 0 \ldots N-1, \tag{2.1}$$

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

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n}, \tag{2.2}$$

where $\mathbf{y}$ is the received vector, $\mathbf{X}$ is a diagonal matrix containing the transmitted signal points, $\mathbf{h}$ is a channel attenuation vector, and $\mathbf{n}$ is a vector of independent and identically distributed complex, zero-mean, Gaussian noise variables with variance σ_n^2. Without loss of generality, we assume that $E\left\{|h_k|^2\right\} = 1$.

Chapter 3

DFT-based estimators

The task of the channel estimator is to estimate the channel attenuations $\mathbf{h}$ from the observations $\mathbf{y}$, given the transmitted symbols $\mathbf{X}$. For the sake of a tractable analysis we assume the x_ks to be known at the receiver.

Since OFDM systems are designed such that the symbol time is significantly longer than the duration of the channel impulse response, the inverse DFT of the channel attenuation vector $\mathbf{h}$ has most of its power concentrated to relatively few samples. As an illustration of this power concentration, Figure 3.1 shows the channel power in the time domain for two channel types, sample-spaced and non-sample-spaced. Sample-spaced channels are channels that have all fading impulses at integer multiples of the system sampling rate, and for which the DFT gives optimal power concentration [6].

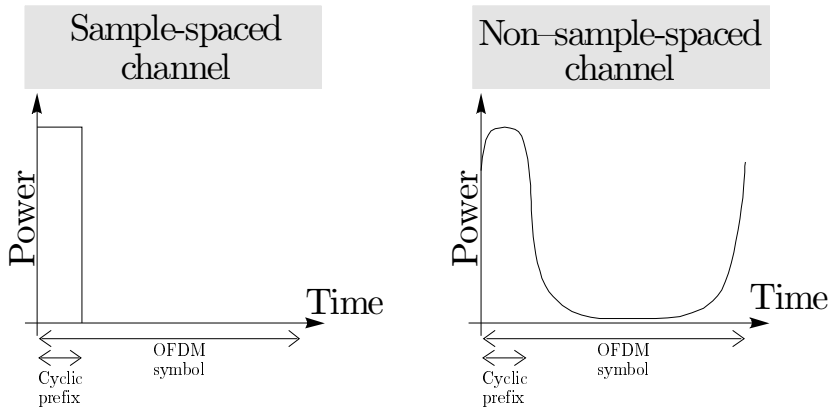

Figure 3.1: Schematic pictures of the time-domain power distribution of the channel, i.e., of IDFT ($\mathbf{h}$).

DFT-based estimation of sample-spaced channels is addressed in [6], and the three estimators we present and analyze are generalizations thereof for non-sample-spaced channels. For clarity, we first calculate the *linear minimum mean-squared error* (LMMSE) estimate of $\mathbf{h}$.

We base our estimates on the LS estimate (the backrotated observations)

$$\widehat{\mathbf{h}}_{ls} = \mathbf{X}^{-1}\mathbf{y} = \mathbf{h} + \widetilde{\mathbf{n}}, \tag{3.1}$$

where $\widetilde{\mathbf{n}} = \mathbf{X}^{-1}\mathbf{n}$ is a vector of independent Gaussian noise variables with covariance matrix $\mathbf{R}_{\widetilde{n}\widetilde{n}} = \sigma_n^2 \left(\mathbf{X}\mathbf{X}^H\right)^{-1}$. The LS estimate $\widehat{\mathbf{h}}_{ls}$ constitutes a sufficient statistic since $\mathbf{X}$ is non-singular. The LS estimate is a noisy observation of the channel attenuations and can be smoothed using correlation properties of the channel. The optimal linear estimator in terms of mean-squared error (MSE) is [8, 10]

$$\widehat{\mathbf{h}} = \mathbf{W}_{\mathbf{X}}\widehat{\mathbf{h}}_{ls}, \tag{3.2}$$

where

$$\mathbf{W}_{\mathbf{X}} \triangleq \mathbf{R}_{hh}\left(\mathbf{R}_{hh} + \sigma_n^2\left(\mathbf{X}\mathbf{X}^H\right)^{-1}\right)^{-1}, \tag{3.3}$$

and $\mathbf{R}_{hh} = E\left\{\mathbf{h}\mathbf{h}^H\right\}$ is the auto-covariance matrix of the channel vector $\mathbf{h}$.

At this point we recognize that the weighting matrix $\mathbf{W}_{\mathbf{X}}$ is of size $N \times N$ and depends on the transmitted data $\mathbf{X}$. As a first step towards low-complexity estimators we want to find a weighting matrix that is independent of the transmitted data. This can be obtained by considering $\widehat{\mathbf{h}}_{ls}$ to be our observation and derive an LMMSE estimator that considers $\mathbf{X}$ to be stochastic with independent and uniformly distributed constellation points. In that case, the auto-covariance matrix of the noise becomes $\mathbf{R}_{\widetilde{n}\widetilde{n}} = \frac{\beta}{\mathrm{SNR}}\mathbf{I}$, where $\beta \triangleq E\left\{|x_k|^2\right\} E\left\{|x_k|^{-2}\right\}$ is a constellation factor ($\beta = 17/9$ for 16-QAM) and $\mathrm{SNR} \triangleq E\left\{|x_k|^2\right\}/\sigma_n^2$ is the per-symbol signal-to-noise ratio.

The LMMSE estimate of $\widehat{\mathbf{h}}_{ls}$ from $\widehat{\mathbf{h}}_{ls}$ now becomes

$$\widehat{\mathbf{h}}_{lmmse} = \mathbf{W}\widehat{\mathbf{h}}_{ls}, \tag{3.4}$$

where the fixed weighting matrix is given by

$$\mathbf{W} \triangleq \mathbf{R}_{hh}\left(\mathbf{R}_{hh} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)^{-1}. \tag{3.5}$$

This LMMSE estimator still requires N multiplications per estimated attenuation and we use it both as a reference and as a starting point in the derivation of the DFT-based low-complexity estimators.

We now use the property of OFDM systems identified above and in [6, 7, 8]: $\mathbf{h}$ is the sampled frequency response of a channel with short time duration compared to the OFDM symbol length and, hence, its associated cyclic impulse response $\mathbf{g} = \mathrm{IDFT}(\mathbf{h})$ has only a few taps with significant power. If we perform the estimation in the time-domain, we can reduce the complexity of the estimation by using this power concentration.

This prompts the estimator structure in Figure 3.2, where the LS estimate is transformed into its time-domain equivalent $\widehat{\mathbf{g}}_{ls} = \mathrm{IDFT}\left(\widehat{\mathbf{h}}_{ls}\right)$. The smoothing is then performed by a linear transformation

$$\widehat{\mathbf{g}} = \mathbf{Q}\widehat{\mathbf{g}}_{ls} \tag{3.6}$$

and the result is transformed back to the frequency-domain, $\widehat{\mathbf{h}} = \mathrm{DFT}(\widehat{\mathbf{g}}_{ls})$. The important benefit of this estimator structure in terms of complexity is the low complexity of the DFT/IDFT (implemented as fast transforms) and the time-domain power concentration. This offers a simplification of (3.6) without sacrificing too much in performance.

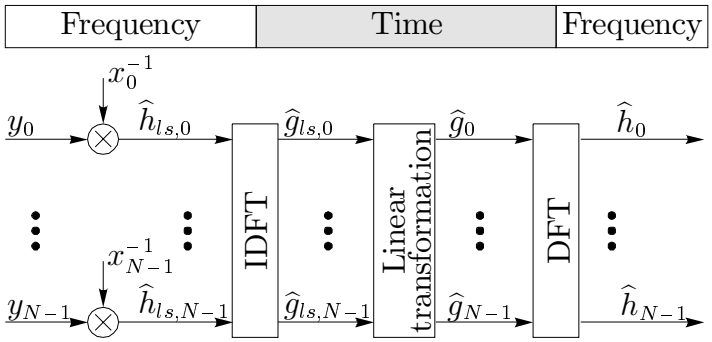

Figure 3.2: Structure of DFT-based channel estimators, where the linear transformation consists of a matrix multiplication $\widehat{\mathbf{g}} = \mathbf{Q}\widehat{\mathbf{g}}_{ls}$.

Our approach is to find sparse approximations of the LMMSE estimator's equivalent time-domain smoothing matrix

$$\mathbf{Q} = \mathbf{F}^H \mathbf{W} \mathbf{F}, \tag{3.7}$$

where $\mathbf{F}$ is the $N \times N$ unitary DFT matrix and $\mathbf{W}$ is defined in (3.5). This will reduce the number of required multiplications, and thus the estimator complexity. A straightforward way is to simply ignore the coefficients in $\widehat{\mathbf{g}}_{ls}$ that contain more noise than channel power and only transform the remaining elements back to the frequency domain. For sample-spaced channels, this is a fruitful approach [5, 6] since the major part of the coefficients only contain noise and no channel power. If the channel is not sample-spaced, however, the channel power is still concentrated but distributed over all coefficients. Due to the lost channel power in the ignored coefficients the simplification causes an irreducible error floor [8].

We now move through three simplification steps and obtain three different low-complexity estimators, of which the last is the straightforward approach described above. The general concept is based on reducing the number of non-zero elements in the time-domain matrix multiplication (3.6), with the aim of reducing the computational complexity and preserving the performance. The three estimators are selected as follows: (See Appendix A for a detailed derivation.)

- **Estimator A**

  By choosing the M coefficients in $\widehat{\mathbf{g}}_{ls}$ that have the highest channel power, we restrict the linear transform in the time-domain to a fixed matrix of size $M \times M$. If M is chosen much smaller than N, the complexity reduction compared to the LMMSE is considerable. The complexity of the time-domain processing in this case is M^2/N multiplications per estimated attenuation. This estimator converges to the LMMSE when $M \to N$. We have presented a related estimator previously in [8].

- **Estimator B**

  Further reductions in complexity can be done by ignoring cross-correlation between the M chosen taps in $\widehat{\mathbf{g}}_{ls}$ and only weighting them individually. This essentially means that we restrict the time-domain processing to be a diagonal $M \times M$ matrix multiplication. Since this matrix only has M non-zero elements, the complexity of this processing is M/N

multiplications per estimated attenuation. To the authors' knowledge, this estimator has not been presented before.

- **Estimator C**

  In this last estimator, we further restrict the time-domain processing to only use the M chosen coefficients directly as input to the DFT. This means restricting the matrix to an $M \times M$ identity matrix, which does not require any multiplications at all. When $M \to N$ this estimator converges to the LS estimator. This estimator is similar to the estimator designed for sample-spaced channels in [5, 6].

Estimator	**Restriction** (cumulative)	**Linear transformation**	**Effective matrix size**	**Required mult./attenuation**
LMMSE	N.A.	N.A.	$N \times N$	$N+1$
A	Use M coefficients.	$\mathbf{Q}_A = \begin{bmatrix} \mathbf{Q}_{M\times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$	$M \times M$	$\log_2 N + \frac{M^2}{N} + 1$
B	Diagonal matrix.	$\mathbf{Q}_B = \begin{bmatrix} \mathbf{D}_{M\times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$	$M \times M$ (diag.)	$\log_2 N + \frac{M}{N} + 1$
C	Identity matrix.	$\mathbf{Q}_C = \begin{bmatrix} \mathbf{I}_{M\times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$	$M \times M$ (identity)	$\log_2 N + 1$
LS	N.A.	N.A.	N.A.	1

Table 3.1: Analysed DFT-based estimators. See Appendix A for details.

The outlined estimators (A-C) are summarized in Table 3.1, where we have also included the LMMSE and LS estimators as references. The table shows the total computational complexity, including the IDFT and DFT[1] for estimators A–C. In general, unless M is close to N, the complexity decreases in Table 3.1 from the LMMSE to the LS.

[1]To obtain a complexity measure we have assumed that N is a power of two and that the DFT and the IDFT requires $\frac{1}{2}N\log_2 N$ complex multiplications each [11].

Chapter 4

Performance analysis

The estimators presented in the previous section all have different computational complexities, and the design variations give them different performances as well. To illustrate the performance of these estimators, we calculate the MSE and use formulae from [12] to obtain the uncoded 16-QAM symbol-error rate (SER).

The parameters we have chosen for the OFDM system are $N = 64$ subcarriers and a cyclic prefix of length $L = 8$ samples. The impulse response of the physical channel

$$g(\tau; t) = \sum_m \alpha_m(t)\delta(\tau - \tau_m)$$

consists of independent Rayleigh-fading (α_m) impulses, uniformly distributed (τ_m) over the length of the cyclic prefix and with a constant power delay profile. From this channel model we calculate the auto-covariance matrix $\mathbf{R}_{hh}$ of the channel attenuations $\mathbf{h}$ [13]. The cyclic impulse response $\mathbf{g} = \mathbf{F}^H\mathbf{h}$ has a corresponding auto-covariance matrix $\mathbf{R}_{gg} = \mathbf{F}^H\mathbf{R}_{hh}\mathbf{F}$. We assume that the variances $\gamma_k = E\left\{|g_k|^2\right\}$ are in decreasing order, *i.e.*, $\gamma_0 \geq \gamma_1 \ldots \geq \gamma_{N-1}$. The auto-covariance matrix of the M used transform coefficients, $\mathbf{R}_{gg,M}$, is thus the upper left $M \times M$ corner of $\mathbf{R}_{gg}$.

Through direct calculation of the auto-covariance matrix of the estimation error

$$\mathbf{R}_{e_Q e_Q} = E\left\{\left(\widehat{\mathbf{h}} - \mathbf{h}\right)\left(\widehat{\mathbf{h}} - \mathbf{h}\right)^H\right\} \tag{4.1}$$

for all estimators, we obtain their respective average MSEs

$$\text{MSE} = \frac{1}{N}\,\text{Trace}\left(\mathbf{R}_{e_Q e_Q}\right). \tag{4.2}$$

Note that the diagonal elements of $\mathbf{R}_{e_Q e_Q}$ are the individual error variances for each channel attenuation. The calculations are derived in Appendix A. The final MSE expressions for the LMMSE estimator, the low-complexity estimators (A–C) and the LS estimator are displayed in Table 4.1. Appendix A also contains a note on sample-spaced channels that is of interest when comparing this analysis with the analysis in [6]. As displayed in the table, estimators A, B, and C experience an error floor (MSE) due to the channel power in the $N - M$ excluded channel taps. This error floor is the same for all DFT-based low-complexity estimators. The individual ranking of the low-complexity estimators in terms of MSE, for a fixed M, is $\text{MSE}_A \leq \text{MSE}_B \leq \text{MSE}_C$.

Estimator	**Average MSE**
LMMSE	$\frac{1}{N}\frac{\beta}{\mathrm{SNR}}\sum_{k=0}^{N-1}\frac{\lambda_{k,N}}{\lambda_{k,N}+\frac{\beta}{\mathrm{SNR}}}$
A	$\frac{1}{N}\frac{\beta}{\mathrm{SNR}}\sum_{k=0}^{M-1}\frac{\lambda_{k,M}}{\lambda_{k,M}+\frac{\beta}{\mathrm{SNR}}}+\underline{\mathrm{MSE}}$
B	$\frac{1}{N}\frac{\beta}{\mathrm{SNR}}\sum_{k=0}^{M-1}\frac{\gamma_{k}}{\gamma_{k}+\frac{\beta}{\mathrm{SNR}}}+\underline{\mathrm{MSE}}$
C	$\frac{M}{N}\frac{\beta}{\mathrm{SNR}}+\underline{\mathrm{MSE}}$
LS	$\frac{\beta}{\mathrm{SNR}}$

$\lambda_{k,M}$ – eigenvalues of $\mathbf{R}_{gg,M}$.
γ_k – diagonal elements of $\mathbf{R}_{gg}$ in decreasing order.
$\underline{\mathrm{MSE}}=\frac{1}{N}\sum_{k=M}^{N-1}\gamma_k$ (MSE floor)

Table 4.1: Average MSE for the investigated estimators. $\mathbf{R}_{gg,M}$ denotes the covariance matrix of the M dominating taps in $\mathbf{g}$. Note that the $\lambda_{k,N}$s are eigenvalues of $\mathbf{R}_{gg}$.

Using the formulae from [12], we display 16-QAM SER curves for two different numbers of included transform coefficients ($M = 10$ and $M = 40$) in Figure 4.1. The 16-QAM SER of the full LMMSE estimator and the LS estimator are also included in the figure, as references. Note that for $M = 10$ there is no visible difference in SER between the three low-complexity estimators and that the SER levels off at high SNRs, due to the error floor. It is only for higher values of M that a difference in SER is noticeable between the low-complexity estimators. It is also noteworthy that even at $M = 40$ (out of a possible $N = 64$) the error floor is visible, and the low-complexity estimators perform worse than the LS estimator above a certain SNR. For $M = 40$, the figure also shows that the performance of the estimators is decreasing from A to C.

The performance of the low-complexity estimators, especially for high SNRs, depends strongly on the number of included taps. An analysis of this behavior shows that the three low-complexity estimators are quite different in terms of complexity versus performance. This is illustrated in Figure 4.2, which shows that the complexity needs to be high to make estimator A the best in terms of SER. In general, estimators B and C are more efficient per required multiplication. For a 64 subcarrier system, the complexity of estimator C is always 7 multiplications per attenuation while estimator B has a complexity of between 7 and 8 multiplications per attenuation, depending on the number of taps used.

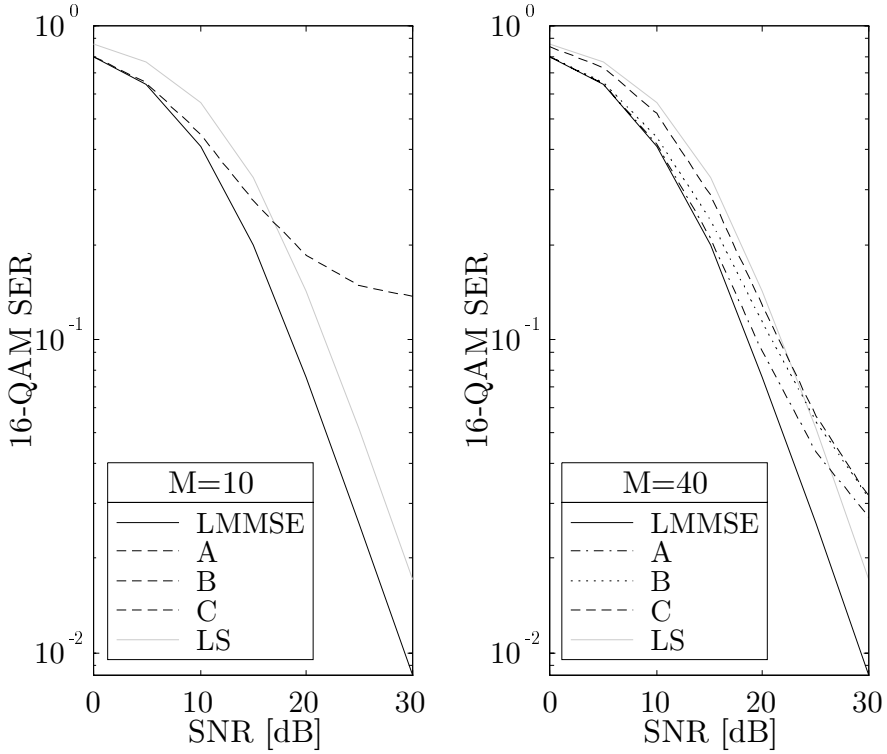

Figure 4.1: Uncoded 16-QAM SER for three systems using the low-complexity estimators A, B, and C, respectively. (Note: Estimators A, B and C have the same SER for $M = 10$.)

If we want to eliminate the error floor entirely, we have to use all taps ($M = N$) in estimators A–C. Effectively, this turns estimator A into the high-complexity LMMSE estimator and estimator C into the low-performance LS estimator. However, estimator B which has lower complexity than estimator A and better performance than estimator C is a good compromise between the two. Estimator A does not have a complexity low enough to compete with the other two and the approximations in estimator C seem to be too crude to provide a good estimate. This enables us to single out estimator B as the best trade-off between complexity and performance among the three estimators analyzed.

As a final comparison, we present the 16-QAM SER performance of estimator B (with all taps included), the LMMSE estimator, and the LS estimator in Figure 4.3. The SER performance of estimator B is good for low SNRs where it resembles that of the LMMSE estimator, and this at less than 1/8 of the complexity (number of multiplications) of the LMMSE. At high SNRs it converges to the SER of the LS estimator, but there is no error floor and the performance never becomes worse than that of the LS estimator. Estimator B has more than a 2 dB gain over the LS estimate for SNRs less than 15 dB and a 3 dB gain for SNRs below 5 dB.

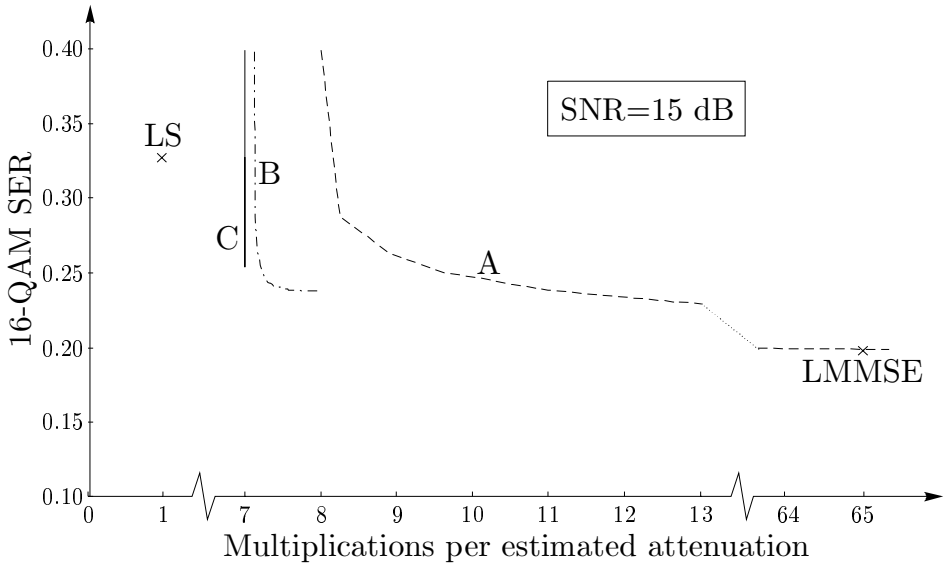

Figure 4.2: SER versus estimator complexity at SNR = 15 dB. Note the discontinuities of the abscissa.

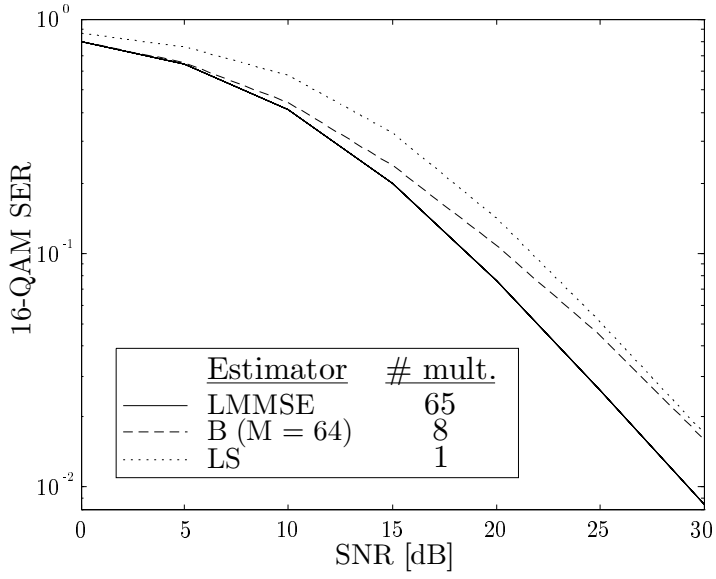

Figure 4.3: Uncoded 16-QAM SER for estimator B where all 64 taps are used ($M = N$). Curves for the LMMSE and LS estimators are included as references.

Chapter 5

Conclusions

In this paper we have presented the MSE and SER performances of three low-complexity DFT-based estimators (A–C). Using M of the N time-domain coefficients yields estimators with complexity (excluding the DFT/IDFT) of M^2/N, M/N and 0 multiplications per estimated channel attenuation, respectively. Only the first of these estimators is potentially of high complexity since $M \leq N$. We have provided analytical expressions for the mean-squared error of all three estimators and shown that, if the number of included taps is less than N, they suffer from an irreducible error floor.

The error-floor can be removed if all taps are used in the linear transform, but only estimator B maintains both its performance and its low complexity in this case. The other two designs either experience a drastically increased complexity (estimator A) or converge to the poor performance of the LS estimator (estimator C). Estimator B maintains good performance with low complexity by ignoring the relatively small cross-correlation between the time-domain channel coefficients. So, we consider estimator B using all N coefficients to be the most suitable of the DFT-based channel estimators. In the investigated system it has almost the same performance as the LMMSE estimator for low SNRs, and this at less than 1/8 of the complexity. Further, in terms of symbol-error rate, this estimator has more than a 2 dB gain over the LS estimator for SNRs less than 15 dB.

Appendix A

Estimator expressions

In this appendix we provide expressions for the linear transformations of the three low-complexity estimators and the MSE for each estimator. At the end we have also included a note on sample-spaced channels, which is of interest when comparing our analysis to the one in [6].

To simplify the matrix notation we assume that the coefficients in the cyclic impulse response $\widehat{\mathbf{g}}_{ls}$ are ordered according to decreasing channel power. This is justified since permutations of the DFT/IDFT coefficients do not change the estimators – it only changes the order in which the coefficients are indexed. The channel power in coefficient $\widehat{g}_{ls,k}$ is denoted $\gamma_k = E\left\{|g_k|^2\right\}$, which are the diagonal elements of $\mathbf{R}_{gg}$.

Before we start the derivation, we write the estimator structure, Figure 3.2, in matrix notation

$$\widehat{\mathbf{h}}_Q = \mathbf{F}\mathbf{Q}\mathbf{F}^H\mathbf{X}^{-1}\mathbf{y} = \mathbf{F}\mathbf{Q}\mathbf{F}^H\widehat{\mathbf{h}}_{ls}, \tag{A.1}$$

where $\mathbf{Q}$ is the matrix representing the linear transformation in the time domain and $\mathbf{F}$ is the $N \times N$ unitary DFT matrix (Note: $\mathbf{F}^{-1} = \mathbf{F}^H$). Further, we need the auto-covariance matrix of $\mathbf{g}$ which is

$$\mathbf{R}_{gg} = E\left(\mathbf{g}\mathbf{g}^H\right) = \mathbf{F}^H\mathbf{R}_{hh}\mathbf{F}. \tag{A.2}$$

We are minimizing the MSE, and thus need an explicit expression for the auto-covariance matrix of the estimation error (4.1)

$$\begin{aligned} \mathbf{R}_{e_Q e_Q} &= E\left\{\left(\widehat{\mathbf{h}}_\mathbf{Q} - \mathbf{h}\right)\left(\widehat{\mathbf{h}}_\mathbf{Q} - \mathbf{h}\right)^H\right\} \\ &= \mathbf{F}\mathbf{Q}\mathbf{F}^H\left(\mathbf{R}_{hh} + \frac{\beta}{\mathrm{SNR}}\mathbf{I}\right)\mathbf{F}\mathbf{Q}^H\mathbf{F}^H - \\ &\quad \mathbf{R}_{hh}\mathbf{F}\mathbf{Q}^H\mathbf{F}^H - \mathbf{F}\mathbf{Q}\mathbf{F}^H\mathbf{R}_{hh}^H + \mathbf{R}_{hh}, \end{aligned} \tag{A.3}$$

which gives the MSE according to (4.2).

To save space, the calculations below are not presented in detail. However, the respective $\mathbf{Q}$ matrices of the estimators are substituted in (A.3) and the following equalities are applied:

- $\mathrm{Trace}\left(\mathbf{U}\mathbf{R}\mathbf{U}^H\right) = \mathrm{Trace}\left(\mathbf{R}\right)$, when $\mathbf{U}$ is a unitary matrix.
- $\mathrm{Trace}\left(\mathbf{D}_1\mathbf{R}\mathbf{D}_2\right) = \mathrm{Trace}\left(\mathbf{D}_1\mathbf{D}_2\mathbf{R}\right)$, when $\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal matrices.

- $\mathbf{R}_{gg,M} = \mathbf{U}_M^H \mathbf{R}_{hh} \mathbf{U}_M$, where $\mathbf{F} = \left[\begin{array}{cc} \mathbf{U}_M & \mathbf{V}_{N-M} \end{array} \right]$ and $\mathbf{U}_M$ contains the first M columns of $\mathbf{F}$.
- $\mathbf{U}_M^H \mathbf{V}_M = \mathbf{0}$, where $\mathbf{U}_M$ and $\mathbf{V}_{N-M}$ are defined above.

To simplify the MSE expressions we also use that the trace of a matrix is equivalent to the sum of its eigenvalues [14]. Further, relation (A.2) implies that $\mathbf{R}_{hh}$ and $\mathbf{R}_{gg}$ share the same eigenvalues [14], which is used to avoid separate notations.

- **LS estimator**

  The LS estimate (3.1) is used as input to the IDFT in Figure 3.2, and its MSE is

  $$\text{MSE} = \frac{\beta}{\text{SNR}}.$$

  This estimate only requires one multiplication per estimated attenuation. Since it is the input to the rest of the estimators, this one multiplication will show up in the following complexity expressions.

- **LMMSE estimator**

  The LMMSE estimator is given in (3.4) and (3.5) and its MSE becomes

  $$\begin{aligned} \text{MSE} &= \frac{1}{N} \text{Trace} \left(\mathbf{R}_{hh} \left(\mathbf{I} - \left(\mathbf{R}_{hh} + \frac{\beta}{\text{SNR}} \mathbf{I} \right)^{-1} \mathbf{R}_{hh} \right) \right) \\ &= \frac{1}{N} \frac{\beta}{\text{SNR}} \sum_{k=0}^{N-1} \frac{\lambda_{k,N}}{\lambda_{k,N} + \frac{\beta}{\text{SNR}}}, \end{aligned}$$

  where the $\lambda_{k,N}$s are eigenvalues of $\mathbf{R}_{hh}$ (and $\mathbf{R}_{gg}$).

  Implementing this estimator as a matrix multiplication as in (3.4) requires $N+1$ multiplications per estimated attenuation.

- **Estimator A**

  Imposing the first restriction on the linear transformation,

  $$\mathbf{Q}_A = \left[\begin{array}{cc} \mathbf{Q}_{M \times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array} \right],$$

  the minimal MSE is obtained if

  $$\mathbf{Q}_{M \times M} = \mathbf{R}_{gg,M} \left(\mathbf{R}_{gg,M} + \frac{\beta}{\text{SNR}} \mathbf{I} \right)^{-1},$$

  where $\mathbf{R}_{gg,M}$ is the upper left $M \times M$ corner of $\mathbf{R}_{gg}$.

  The MSE of this estimator is

  $$\begin{aligned} \text{MSE} &= \frac{1}{N} \text{Trace} \left(\mathbf{R}_{gg,M} \left(\mathbf{I} - \left(\mathbf{R}_{gg,M} + \frac{\beta}{\text{SNR}} \mathbf{I} \right)^{-1} \mathbf{R}_{gg,M} \right) \right) + \underline{\text{MSE}} \\ &= \frac{1}{N} \frac{\beta}{\text{SNR}} \sum_{k=0}^{M-1} \frac{\lambda_{k,M}}{\lambda_{k,M} + \frac{\beta}{\text{SNR}}} + \underline{\text{MSE}} \end{aligned}$$

where the $\lambda_{k,M}$s are eigenvalues of $\mathbf{R}_{gg,M}$ and

$$\underline{\text{MSE}} = \text{Trace}\left(\mathbf{V}_{N-M}^H \mathbf{R}_{hh} \mathbf{V}_{N-M}\right) = \frac{1}{N} \sum_{k=M}^{N-1} \gamma_k, \tag{A.4}$$

where $\mathbf{V}_{N-M}$ contains the last $N-M$ columns of $\mathbf{F}$. We call $\underline{\text{MSE}}$ the *MSE floor*, since it only depends on the number of excluded taps $(N-M)$ and lower bounds the MSE.

Implementing this estimator according to Figure 3.2 requires $\log_2 N + \frac{M^2}{N} + 1$ multiplications per estimated attenuation.

- **Estimator B**

  Applying the second restriction on the linear transformation,

$$\mathbf{Q}_C = \begin{bmatrix} \mathbf{D}_{M\times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

  where $\mathbf{D}_M = \text{diag}\left(\delta_0, \delta_1, \ldots, \delta_{M-1}\right)$, we obtain a minimal MSE if

$$\delta_k = \frac{\gamma_k}{\gamma_k + \frac{\beta}{\text{SNR}}}.$$

  The MSE of this estimator is

$$\text{MSE} = \frac{1}{N} \frac{\beta}{\text{SNR}} \sum_{k=0}^{M-1} \frac{\gamma_k}{\gamma_k + \frac{\beta}{\text{SNR}}} + \underline{\text{MSE}},$$

  where $\underline{\text{MSE}}$ is given by (A.4).

  Implementing this estimator according to Figure 3.2 requires $\log_2 N + \frac{M}{N} + 1$ multiplications per estimated attenuation.

- **Estimator C**

  Applying the last restriction on the linear transformation,

$$\mathbf{Q}_D = \begin{bmatrix} \mathbf{I}_{M\times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

  we do not have any choice in the design, except for M, and the MSE becomes

$$\text{MSE} = \frac{M}{N} \frac{\beta}{\text{SNR}} + \underline{\text{MSE}}$$

  where, again, $\underline{\text{MSE}}$ is given by (A.4).

  Implementing this estimator according to Figure 3.2 requires $\log_2 N + 1$ multiplications per estimated attenuation.

A note on sample-spaced channels

The above expressions are derived for a general case, but they have some interesting properties for sample-spaced channels that are worth noting. Consider a channel impulse response

$$g(\tau;t) = \sum_{m=0}^{L} \alpha_m(t)\,\delta(\tau - mT_s),$$

where the fading amplitudes $\alpha_m(t)$, of the sample-spaced impulses, are independent. Then $\mathbf{F}$ diagonalizes $\mathbf{R}_{hh}$, and $\mathbf{R}_{gg}$ becomes diagonal with only L non-zero elements. The channel power in the kth coefficient of $\widehat{\mathbf{g}}_{ls}$ is therefore equivalent to the kth largest eigenvalue $\lambda_{k,N}$ of $\mathbf{R}_{hh}$ (and $\mathbf{R}_{gg}$), and the eigenvalues of $\mathbf{R}_{gg,M}$ becomes $\lambda_{k,M} = \lambda_{k,N} = \gamma_k$, of which only the first L are non-zero. Hence, the MSEs for the estimators become:

Estimator	**Average MSE**
LMMSE	$\frac{1}{N}\frac{\beta}{\mathrm{SNR}}\sum_{k=0}^{N-1}\frac{\gamma_k}{\gamma_k+\frac{\beta}{\mathrm{SNR}}}$
A	$\frac{1}{N}\frac{\beta}{\mathrm{SNR}}\sum_{k=0}^{M-1}\frac{\gamma_k}{\gamma_k+\frac{\beta}{\mathrm{SNR}}} + \underline{\mathrm{MSE}}$
B	$\frac{1}{N}\frac{\beta}{\mathrm{SNR}}\sum_{k=0}^{M-1}\frac{\gamma_k}{\gamma_k+\frac{\beta}{\mathrm{SNR}}} + \underline{\mathrm{MSE}}$
C	$\frac{M}{N}\frac{\beta}{\mathrm{SNR}} + \underline{\mathrm{MSE}}$
LS	$\frac{\beta}{\mathrm{SNR}}$

First of all, estimator A is now equivalent to estimator B and, since $\gamma_k = 0$ for $k \geq L$, both estimator A and B are equivalent to the LMMSE estimator for $M \geq L$. Further, since $\underline{\mathrm{MSE}} = 0$ for $M \geq L$, we can choose $M = L$ in estimator C which reduces the noise, compared to the LS estimator, to a fraction L/N. This last observation was also done in [6].

Bibliography

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

[2] John A. C. Bingham. Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Commun. Mag.*, 28(5):5–14, May 1990.

[3] Leonard J. Cimini. Analysis and simulation of a digital mobile channel using orthogonal frequency-division multiplexing. *IEEE Trans. Commun.*, COM-33(7):665–675, July 1985.

[4] 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.

[5] Ahmad Chini, Mohammed S. El-Tanany, and Samy A. Mahmoud. Transmission of high rate ATM packets over indoor radio channels. *IEEE J. Select. Areas Commun.*, 14(3):469–476, April 1996.

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

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

[8] 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.

[9] 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.

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

[11] A.V. Oppenheim and R.V. Schafer. *Discrete-time signal processing*. Prentice-Hall, 1989.

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

[13] Ove Edfors, Magnus Sandell, Jan-Jaap van de Beek, Sarah Kate Wilson, and Per Ola Börjesson. OFDM channel estimation by singular value decomposition. Research Report TULEA 1996:18, Div. of Signal Processing, Luleå University of Technology, September 1996.

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