Impact of channel estimation errors on the performance of linear FIR equalizers for frequency selective MIMO channels.

Dieter Van Welden, Frederik Simoens, Heidi Steendam and Marc Moeneclaey
DIGCOM research group, TELIN Department,
Ghent University, Sint-Pietersnieuwstraat 41, 9000 GENT, Belgium
E-mail: {dmvwelde,fsimoens,hs,mm}@telin.ugent.be

Receivers require channel state information (CSI) to be able to detect the transmitted information. Therefore the channel has to be estimated. This imperfect channel knowledge results in an extra noise term at the receiver side. In this paper, the influence of a Gaussian channel estimation error on the performance of a linear Minimum Mean Squared Error (MMSE) Finite Impulse Response (FIR) equalizer is investigated. The BER degradation is calculated as a function of the variance of this Gaussian error. The performance of a pilot aided Maximum Likelihood (ML) estimator is also investigated based on the results for Gaussian errors.

1 Introduction

Frequency selective Multiple-Input Multiple-Output (MIMO) fading channels are of great interest since they enable a high data-rate wireless communication [1]. Furthermore, these frequency selective MIMO channels offer a high space- and frequency-diversity. To mitigate the interference arising from the multi-antenna set-up and the frequency selectivity of the channel, several receiver structures have recently been proposed [2, 3, 4]. All these receivers require knowledge of the CSI. In practice, this information is unavailable at the receiver and has to be estimated before detection of the transmitted symbols can take place. Because of noise and other disturbing effects, the channel estimators cannot provide perfect CSI and the resulting channel estimation errors will cause a degradation of the receiver's performance.

In this paper, we investigate the sensitivity of a linear MMSE FIR equalizer to channel estimation errors. We consider a channel estimate that consists of the correct channel tap disturbed by a Gaussian distributed error and compute the BER degradation. Based on this analysis, we investigate the performance of a pilot aided ML estimator.

2 System Model

Consider a MIMO frequency-selective fading channel with N_T transmit antennas, N_R receive antennas and a channel memory of length L. The received signal, sampled at symbol-rate, can be modeled as

$$\mathbf{y}(k) = \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{a}(k-l) + \mathbf{w}(k) \tag{1}$$

where $\mathbf{H}(l)$ denotes the $N_R \times N_T$ channel matrix corresponding to the l-th tap of the channel impulse response, $\mathbf{a}(k)$ is the vector of the transmitted symbols and $\mathbf{y}(k)$ the vector of the received signals at instant k, and $\mathbf{w}(k)$ denotes the additive white Gaussian noise (AWGN) vector with independent real and imaginary components and variance N_0. We assume the different channel taps are i.i.d. Rayleigh distributed with variance 1.

At the receiver, we use a filterbank consisting of N_T MMSE FIR filters [3], each with $L'N_R$ coefficients, to detect the symbols transmitted by the different transmit antennas. We choose the filter coefficients symmetric around the L time instants during which the symbol vector to be estimated contributes to the received signals. So we define L' as $L' = 2q + L$. The estimate of the k-th transmitted symbol vector $\mathbf{a}(k)$ is given by

$$\hat{\mathbf{a}}(k) = \mathbf{G}^{\mathbf{H}} \left(\mathbf{H}\mathbf{H}^{\mathbf{H}} + \frac{N_0}{E_s}\mathbf{I} \right)^{-1} \mathbf{y}_{\mathbf{EXT}} \tag{2}$$

where $\mathbf{y}_{\mathbf{EXT}} = [\mathbf{y}(k-q)^{\mathsf{T}}, ..., \mathbf{y}(k+q+L-1)^{\mathsf{T}}]^{\mathsf{T}}$ has dimension $L'N_R \times 1$, $\mathbf{G} = [\mathbf{0}_{N_T \times qN_R}, \mathbf{H}(0)^{\mathsf{T}}, ..., \mathbf{H}(L-1)^{\mathsf{T}}, \mathbf{0}_{N_T \times qN_R}]^{\mathsf{T}}$ with $\mathbf{0}_{i \times j}$ denoting the $i \times j$ all zeros matrix and $\mathbf{H}$ a $L'N_R \times (L' + L - 1)N_T$ matrix given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}(L-1) & \dots & \mathbf{H}(0) & \dots & \mathbf{0} \\ \vdots & \ddots & & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{H}(L-1) & \dots & \mathbf{H}(0) \end{bmatrix} \tag{3}$$

3 Impact of an imperfect channel

From (2), it follows that the receiver needs the knowledge of the channel matrix $\mathbf{H}$ to estimate the transmitted data symbols. In practice, the channel matrix $\mathbf{H}$ is not known at the receiver and an estimate $\hat{\mathbf{H}}$ is made. This estimate $\hat{\mathbf{H}}$ is then substituted in (2) to obtain the estimates of $\mathbf{a}(k)$.

$$\hat{\mathbf{a}}(k) = \hat{\mathbf{G}} \left(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathbf{H}} + \frac{N_0}{E_s}\mathbf{I} \right)^{-1} \mathbf{y}_{\mathbf{EXT}} \tag{4}$$

First, we investigate the effect of a Gausian channel estimation error. We model the l-th tap channel estimate as follows:

$$\hat{\mathbf{H}}(l) = \mathbf{H}(l) + \mathbf{E}(l) \tag{5}$$

where $\mathbf{H}(l)$ denotes the true channel tap, $\hat{\mathbf{H}}(l)$ the estimate and $\mathbf{E}(l)$ a complex Gaussian disturbance matrix with independent components with zero mean and independent real and imaginary parts each with a variance $\frac{\sigma_\epsilon^2}{2}$. It is easily shown that this channel estimation error results in an additional Gaussian noise term at the output of the MMSE FIR filterbank. To investigate the influence of $\mathbf{E}(l)$ we write $\mathbf{H}(l) = \hat{\mathbf{H}}(l) - \mathbf{E}(l)$. As $\mathbf{E}(l)$ and $\hat{\mathbf{H}}(l)$ are both Gaussian distributed with variances σ_ϵ^2 and $1+\sigma_\epsilon^2$ respectively, and are not statistically independent, $\mathbf{E}(l)$ can be decomposed into $\mathbf{E}(l) = \alpha\hat{\mathbf{H}}(l) + \beta$, where α and β do not depend on $\hat{\mathbf{H}}(l)$. This yields [5, chapter 12]

$$\begin{aligned}\mathbf{E}(l) &= E_{\mathbf{E}}\left[\mathbf{E}(l) \mid \hat{\mathbf{H}}(l)\right] + \Delta(l) \\ &= \frac{\sigma_\epsilon^2}{1+\sigma_\epsilon^2}\hat{\mathbf{H}}(l) + \Delta(l) \end{aligned} \tag{6}$$

where the elements of $\Delta(l)$ are i.i.d. Gaussian random variables with independent real and imaginary parts, with zero mean and variance $\frac{\sigma_\epsilon^2}{1+\sigma_\epsilon^2}$. Substituting (6) in (5) yields

$$\mathbf{H}(l) = \frac{1}{1+\sigma_\epsilon^2}\hat{\mathbf{H}}(l) - \Delta(l) \tag{7}$$

Taking into account (1) $\mathbf{y_{EXT}}$ can be rewritten as

$$\mathbf{y_{EXT}} = \mathbf{H}\mathbf{a_{EXT}} + \mathbf{w_{EXT}} \tag{8}$$

where $\mathbf{a_{EXT}} = [\mathbf{a}(k-q-L+1)^{\mathsf{T}}, ..., \mathbf{a}(k+q+L-1)^{\mathsf{T}}]^{\mathsf{T}}$and $\mathbf{w_{EXT}} = [\mathbf{w}(k-q)^{\mathsf{T}}, ..., \mathbf{w}(k+q+L-1)^{\mathsf{T}}]^{\mathsf{T}}$. Further defining $\mathbf{H} = \frac{1}{1+\sigma_\epsilon^2}\hat{\mathbf{H}} - \Delta$, we can rewrite (4) as

$$\hat{\mathbf{a}}(k) = \hat{\mathbf{G}}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathbf{H}} + \frac{N_0}{E_s}\mathbf{I}\right)^{-1}\left[\frac{1}{1+\sigma_\epsilon^2}\hat{\mathbf{H}}\mathbf{a_{EXT}} - \Delta\mathbf{a_{EXT}} + \mathbf{w_{EXT}}\right] \tag{9}$$

For given $\hat{\mathbf{H}}$, the signal-to-noise ratio (SNR) can be written as

$$SNR(\hat{\mathbf{H}}) = \frac{E_{\mathbf{a}}\left[\left\|E_{\Delta,\mathbf{w}}\left[\hat{\mathbf{a}}(k) \mid \hat{\mathbf{H}}\right]\right\|^2\right]}{Var_{\mathbf{a},\Delta,\mathbf{w}}\left[\hat{\mathbf{a}}(k) \mid \hat{\mathbf{H}}\right]} \tag{10}$$

with $\|\mathbf{A}\|^2 = tr(\mathbf{A}\mathbf{A}^{\mathsf{H}})$. In (10), we have not taken into account the effect of intersymbol interference (ISI). Indeed, the most important cause of detection errors is the AWGN noise and not the ISI. Only at very high $\frac{E_s}{N_0}$, the ISI will become important.

When $\frac{E_s}{N_0}$ is sufficiently large (but not so large that ISI becomes the important detection error cause), $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathbf{H}} + \frac{N_0}{E_s}\mathbf{I}$ can be approximated by $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathbf{H}}$. Hence (10) can be approximated by

$$SNR_{est}(\hat{\mathbf{H}}) = \frac{\frac{E_s}{(1+\sigma_\epsilon^2)^2} tr\left[\hat{\mathbf{G}}^{\mathbf{H}}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathbf{H}}\right)^{-1}\hat{\mathbf{G}}\right]}{\left(N_0 + \frac{E_s L N_T \sigma_\epsilon^2}{1+\sigma_\epsilon^2}\right)\left\|\hat{\mathbf{G}}^{\mathbf{H}}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathbf{H}}\right)^{-1}\right\|^2} \tag{11}$$

where we used $E\left[\mathbf{a_{EXT}}\mathbf{a_{EXT}^H}\right] = E_s\mathbf{I}$, $E\left[\mathbf{\Delta}\mathbf{\Delta}^{\mathbf{H}}\right] = \frac{L N_T \sigma_\epsilon^2}{1+\sigma_\epsilon^2}\mathbf{I}$ and $E\left[\mathbf{w_{EXT}}\mathbf{w_{EXT}^H}\right] = N_0\mathbf{I}$. The SNR for perfect channel knowledge $SNR_{perf}(\mathbf{H})$ is easily obtained by substituting in (11) $\hat{\mathbf{H}}$ by $\mathbf{H}$ and replacing σ_ϵ^2 by zero.

$$SNR_{perf}(\mathbf{H}) = \frac{E_s tr\left[\mathbf{G}^{\mathbf{H}}\left(\mathbf{H}\mathbf{H}^{\mathbf{H}}\right)^{-1}\mathbf{G}\right]}{N_0\left\|\mathbf{G}^{\mathbf{H}}\left(\mathbf{H}\mathbf{H}^{\mathbf{H}}\right)^{-1}\right\|^2} \tag{12}$$

The bit error rate (BER) given $\hat{\mathbf{H}}$ or $\mathbf{H}$ is a function of the SNR: $BER(\mathbf{X}) = f(SNR(\mathbf{X}))$, where $\mathbf{X} = \mathbf{H},\ \hat{\mathbf{H}}$. To obtain the degradation we compare the BER for perfect channel knowledge $E_{\mathbf{H}}[BER_{perf}(\mathbf{H})]$ with the BER in the presence of channel estimation errors $E_{\hat{\mathbf{H}}}[BER_{est}(\hat{\mathbf{H}})]$. Taking into account that $\hat{\mathbf{H}}$ and $\mathbf{H}$ are zero-mean Gaussian distributed with variances $1+\sigma_\epsilon^2$ and 1 respectively, we can define the variable $\tilde{\mathbf{H}}$ as $\hat{\mathbf{H}}(l) = \sqrt{1+\sigma_\epsilon^2}\tilde{\mathbf{H}}(l)$. As the distributions of $\mathbf{H}$ and $\tilde{\mathbf{H}}$ are the same, it can easily be shown that the degradation caused by imperfect channel knowledge is obtained by dividing $SNR_{perf}(\mathbf{H})$ by $SNR_{est}(\hat{\mathbf{H}})$, yielding

$$degradation = \frac{N_0(1+\sigma_\epsilon^2) + N_T L E_s \sigma_\epsilon^2}{N_0} = 1 + \sigma_\epsilon^2 + N_T L \sigma_\epsilon^2 \frac{E_s}{N_0} \tag{13}$$

When the channel estimates are obtained by pilot-aided ML-channel estimation, the variance σ_ϵ^2 is given by [6]

$$\sigma_\epsilon^2 = \frac{N_0}{LN_T} tr\left[\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right)^{-1}\right] = \frac{N_0}{E_s K} \tag{14}$$

where K is the length of the pilot sequence and $\mathbf{A}$ is defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}(1) & \dots & \mathbf{a}(K) & \dots & \mathbf{0} \\ \vdots & \ddots & & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{a}(1) & \dots & \mathbf{a}(K) \end{bmatrix} \tag{15}$$

where $\mathbf{a}(k)$ denotes the k-th pilot vector. For random pilot sequences it can be shown that (14) is only valid if K is sufficiently large such that $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ is close to a diagonal matrix. However for small K, it is possible to find pilot sequences such that $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ is also close to a diagonal matrix; in which case our analysis is still valid. Substituting (14) in (13) yields the degradation as function of K

$$degradation = 1 + \frac{N_0}{E_s K} + \frac{N_T L}{K} \tag{16}$$

4 Simulation Results

For all simulations the transmitted symbols were BPSK modulated. Fig. 1 *(left)* shows the BER performance for different values of σ_ϵ^2. We see that at low $\frac{E_s}{N_0}$, the noise $w_n(k)$ added by the channel dominates the overall error-rate and the channel estimation error does not significantly degrade the performance. For higher $\frac{E_s}{N_0}$ ($\geq 20dB$) however, the influence of the channel estimation error becomes apparent. The error floor, which is inherent to MMSE FIR filters, grows as the channel estimation error increases. Secondly, we consider the performance when the channel estimate is obtained with a pilot-based Maximum Likelihood (ML) algorithm. Fig. 1 *(right)* illustrates the impact of the pilot sequence length on the BER performance of our receiver. We observe that the BER degradation corresponding to a given pilot sequence length remains approximately constant for all $\frac{E_s}{N_0}$ for sufficiently large K; this was expected from (16).

Fig. 2 shows the degradation as function of the pilot sequence length K, for a target BER of 10^{-3}. The simulation results are given for two cases: for the case of a random pilot sequence and for the case of an optimized pilot sequence. As expected, for short

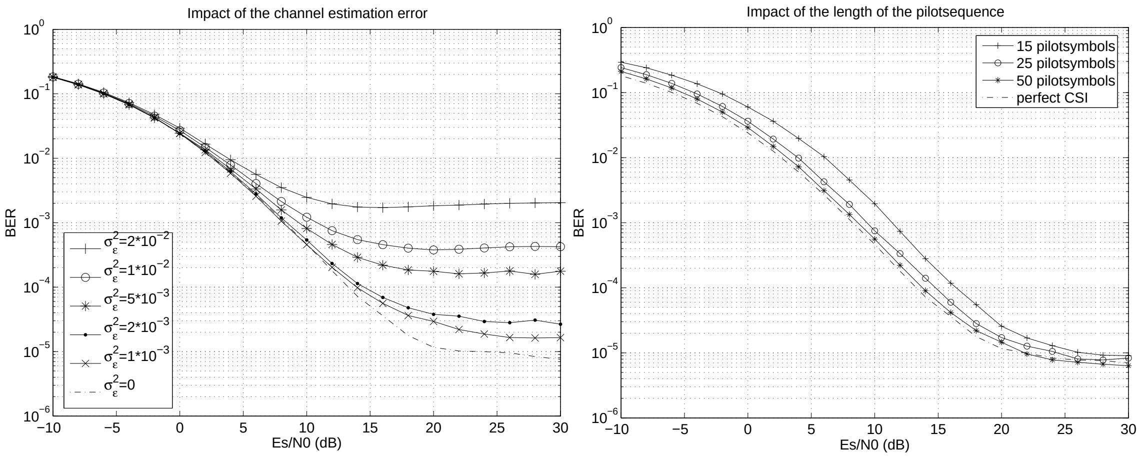

Figure 1: BER of a FIR MMSE receiver ($q = 4$) with Gaussian channel estimation errors ($N_T = 2$ transmitting antennas, $N_R = 2$ receiving antennas and $L = 3$ channel taps).

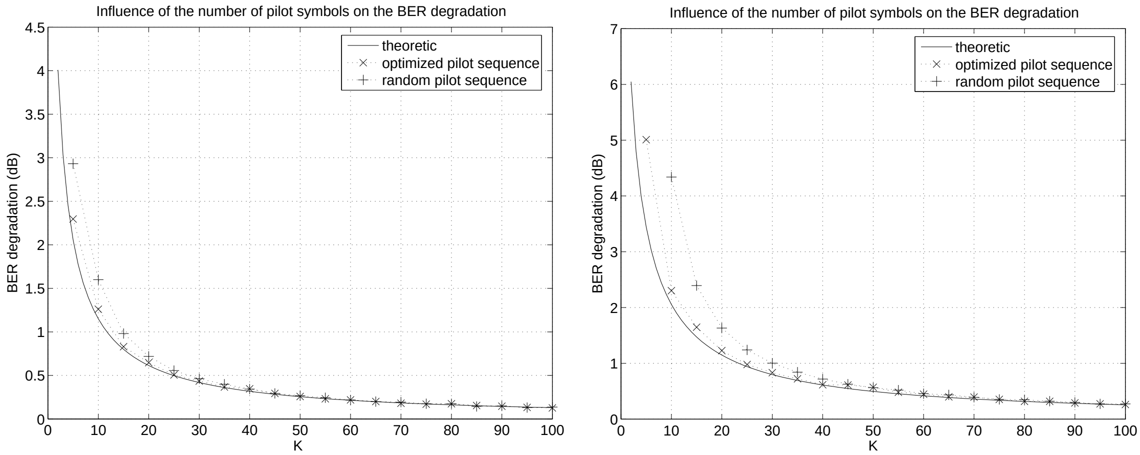

Figure 2: Influence of the length of the number of pilotsymbols on the BER degradation of a FIR MMSE receiver($q = 4$, $N_T = 1$, $N_R = 1$, $L = 3$ (left) and $q = 4$, $N_T = 2$, $N_R = 2$, $L = 3$ (right))

pilot sequences the theoretical curve does not match the simulation results. For small K the approximations in (14) are not valid as $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ is not close to a diagonal matrix, even in the case of optimized pilot sequences. For larger K the theoretical curve matches the simulation results. The curve for the optimized pilot sequences converges faster to the theoretic curve than for the random pilot sequences. This is explained as for optimized pilot sequences, $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ can be better approximated by a diagonal matrix than for the case of random pilot sequences.

5 Conclusions

We have computed the BER degradation for a linear MMSE FIR equalizer as function of a channel estimation error. For a channel estimate obtained with a ML algorithm we have shown that this BER degradation can be written as function of the length of the pilot sequence. This theoretic result was verified with simulation results.

References

[1] G. Foschini and M. Gans. On Limits of Wireless Communication in a Fading Environment when using Multiple Antennas. *Wireless Pers. Commun.*, 6(3):311–335, 1998.

[2] T. Abe and T. Matsumoto. Space-Time turbo equalization in frequency selective MIMO channels. *IEEE Trans. Veh. Technol.*, pages 469–475, May 2003.

[3] X. Wang and H.V. Poor. Iterative (turbo) soft interference cancellation and decoding for coded CDMA. *IEEE Trans. Comm.*, 47:1046–1061, July 1999.

[4] N. Al-Dhahir, A.F. Naguib and A.R. Calderbank. Finite-Length MIMO Decision Feedback Equalization for Space-Time Block-Coded Signals over Multipath-Fading Channels. *IEEE Trans. Vehicular Technology*, 50(4):1176–1182, July 2001.

[5] H. Meyr, M. Moeneclaey, S.A. Fechtel. *Synchronization, Channel Estimation, and Signal Processing*, volume 2 of *Digital Communication Receivers*. John Wiley & Sons, 1997.

[6] E. G. Larsson and P. Stoica. *Space-Time Block coding for Wireless Communications*. Cambridge University Press, 2003.