# ENHANCED DECISION-DIRECTED CHANNEL ESTIMATION OF TIME-VARYING FLAT MIMO CHANNELS

Emna Eitel and Joachim Speidel
Institute of Telecommunications, University of Stuttgart
Stuttgart, Germany


## ABSTRACT

This paper presents a combined pilot-assisted and decision-directed approach to estimate a Rayleigh flat fading MIMO channel. The suggested algorithm is very attractive in comparison to other tracking techniques due to its simplicity and numerical robustness. We introduce a moving-average filter to suppress the noise and show that optimizing the filter parameters such as length and initialization coefficients leads to significant performance improvement. Simulation results for different Doppler frequencies show the effectiveness of the proposed scheme.


## I INTRODUCTION

MIMO systems with coherent detection can deliver high channel capacity provided that an accurate knowledge of the channel is available at the receiver. The performance can be even increased if the channel state information (CSI) is also available at the transmitter. Algorithms to precisely estimate the CSI are therefore crucial. Often periodical pilot-assisted channel estimation (PACE) is employed. In [1], an effective signal to noise ratio (SNR) was derived to take account of the channel estimation (CE) errors as an additive source of noise. The optimal ratio between the training and data phase was also considered. However, in fast-varying channels, PACE does not only decrease the bandwidth efficiency but is also incapable of detecting the fast variations of the channel. Therefore, tracking techniques have to be applied.

A method, that does not require additional pilots is decision-directed channel estimation (DDCE). It considers the recovered data as "new pilots" and can therefore feed the CE module with data that permanently takes account of the actual channel state. The advantage of this method in comparison to others such as interpolation, Wiener or Kalman filtering is that it does nor introduce a delay nor does it require any knowledge about the statistical properties of the channel. Furthermore, it is not subject to numerical problems and shows a small computational complexity. The main drawback of DDCE is its sensitivity to wrongly detected data which causes an error propagation, especially in the low SNR regime. Therefore, in this paper DDCE is combined with a periodical PACE scheme as described in [2] and [3].

We also analyze different filtering techniques in order to suppress the noise inherent to the decision-directed channel estimates. The maximum a posteriori (MAP) filtering in [4] tends to be numerically unstable if the channel varies slowly, as it is the case for many filters of this kind [5]. To tackle this problem, we introduce a moving-average filter. This solution is also suggested in [6]. The optimal filter design is however not considered in detail. We show that due to the recursive nature of the algorithm, filter length and initialization conditions have an important impact on the CE quality.

## II SYSTEM MODEL

We consider an $M \times N$ MIMO system. The $N \times 1$ receive signal vector at time instant $k$ is given by:

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{s}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{s}(k)$ denotes the $M \times 1$ sent signal vector, $\mathbf{H}(k)$ denotes the $N \times M$ MIMO flat fading channel matrix and $\mathbf{n}(k)$ the $N \times 1$ additive white gaussian noise (AWGN) vector whose complex elements are iid and $CN(0, 2\sigma_0^2)$. The channel matrix can be written as $\mathbf{H}(k) = \left[\begin{array}{ccc} \mathbf{h}_1(k) & \cdots & \mathbf{h}_M(k) \end{array}\right]$, with $\mathbf{h}_i(k)$ denoting the *i*th column of $\mathbf{H}(k)$.

The element $h_{ij}(k)$ of $\mathbf{H}(k)$ represents the channel coefficient between the *j*th transmit and the *i*th receive antenna which is assumed to be $CN(0, 1)$ and realizes a discrete-time Rayleigh flat fading process in the base-band. Its temporal autocorrelation function satisfies:

$$E\left\{h_{ij}(k)h_{ij}(k')^*\right\} = J_0(2\pi f_d(k-k')T_s) \tag{2}$$

where $T_s$ stands for the symbol period, $f_d$ is the Doppler frequency and $J_0$ ist the first kind Bessel function of zeroth order. For convenience,we define at this level the normalized Doppler frequency $f_{dnorm} = f_d \cdot T_s$. In order for the receiver to estimate the channel, orthogonal pilot symbol vectors $\mathbf{s}_p$ are periodically sent during the training period that takes $L_p$ symbol periods $T_s$. The training phase is followed by a data transmission phase where $L_d$ symbol vectors are sent. At the end of the training phase, a channel estimate $\widehat{\mathbf{H}}$ is computed by means of the received pilots according to the maximum Likelihood (ML) or the minimum mean squared error (MMSE) principle. In the absence of tracking, the pilot-based channel estimate is used for the coherent detection of data symbols during the subsequent $L_d$ symbol periods. The periodical frame structure is illustrated in Fig. 1. We further introduce the parameter $\tau$ which denotes the time elapsed since the last PACE, i.e. $0 \leq \tau \leq L_d T_s$.

### A Computing the raw channel estimates

The combined PACE and DDCE scheme is illustrated in Fig. 2. A first channel estimate $\widehat{\mathbf{H}}$ is computed at the end of the training phase. Then, in the tracking phase, new channel coefficients can be estimated based on the detected symbols as follows [4]:

$$\tilde{h}_{ij}(k) = \frac{y_i(k) - \sum_{q=1, q\neq j}^{M} \hat{h}_{iq}(k)\hat{s}_q(k)}{\hat{s}_j(k)} \tag{3}$$

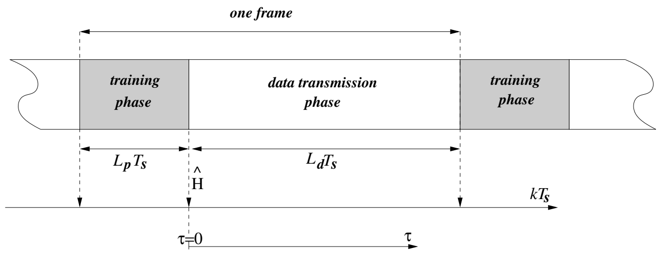


Figure 1: Frame structure with periodically alternating training and data transmission phases

$\tilde{h}_{ij}(k)$ are denoted as raw channel estimate, i.e. without filtering[1]. Executing (3) for each component $i = 1, \ldots, N$ of a column vector $\tilde{\mathbf{h}}_j(k)$ yields:

$$\tilde{\mathbf{h}}_j(k) = \frac{1}{\hat{s}_j(k)} \left[ \mathbf{y}(k) - \sum_{q=1, q\neq j}^{M} \hat{\mathbf{h}}_q(k)\hat{s}_q(k) \right], j = 1, \ldots, M \quad (4)$$

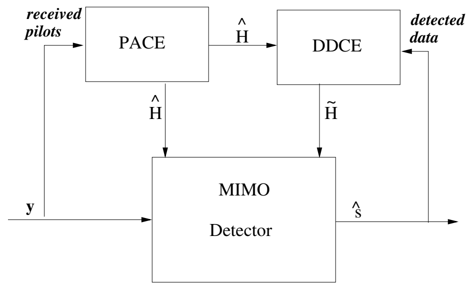


Figure 2: MIMO receiver with pilot-assisted and decision-directed channel estimation

### B Processing order of the raw channel estimate columns

The processing order of the raw channel estimate columns in (4) can be differently chosen according to whether the computation is done at a single time instant or upon an interval of e.g. $B$ time instants.

At a given time instant $k$, applying (4) for all $j = 1, \ldots, M$ simultaneously[2] is time-efficient. However, one has to make the assumption that some old channel estimates are still valid since it is otherwise not possible to compute $M \times N$ unkown variables upon one channel observation (1).

An alternative method would be to wait until a block of $B$ data vectors is detected, such that $B$ observations are available. To be able to compute all $M \times N$ unkown variables, $B$ must satisfy $B \geq M$. Assuming that the channel matrix did not change in the interval of $B$ symbol periods and the detected symbols are identical to the sent ones , we can write the following:

$$\mathbf{Y} = \mathbf{H}\hat{\mathbf{S}} + \mathbf{N} \quad (5)$$

where $\hat{\mathbf{S}} = \begin{bmatrix} \hat{\mathbf{s}}_0 & \hat{\mathbf{s}}_1 & \cdots & \hat{\mathbf{s}}_{B-1} \end{bmatrix}$ and $\mathbf{Y} = \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{B-1} \end{bmatrix}$ are the matrices of the last $B$ detected and receive vectors respectively. $\mathbf{N} = \begin{bmatrix} \mathbf{n}_0 & \mathbf{n}_1 & \cdots & \mathbf{n}_{B-1} \end{bmatrix}$ is the corresponding AWGN matrix.

To compute the decision-directed estimate $\tilde{\mathbf{H}}$, we can apply the MAP method, i.e search for the channel estimate that maximizes the probability density function $p(\tilde{\mathbf{H}}|\mathbf{Y}, \hat{\mathbf{S}})$. Therefore, we rewrite (6) in vector form to apply standard results from estimation theory:

$$\underbrace{vec(\mathbf{Y})}_{\tilde{\mathbf{y}}} = \underbrace{\left(\hat{\mathbf{S}} \otimes \mathbf{I}\right)}_{\mathbf{X}} \cdot \underbrace{vec(\mathbf{H})}_{\mathbf{h}} + \underbrace{vec(\mathbf{N})}_{\tilde{\mathbf{n}}} \quad (6)$$

where $\otimes$ is the Kronecker product. The MAP estimate is given by the well-known equation [7]:

$$\tilde{\mathbf{h}}_{MAP} = vec\left\{\tilde{\mathbf{H}}_{MAP}\right\} = \mathbf{R}_{\mathbf{hh}}\mathbf{X}^H \left(\mathbf{X}\mathbf{R}_{\mathbf{hh}}\mathbf{X}^H + \mathbf{R}_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}\right)^{-1} \tilde{\mathbf{y}} \quad (7)$$

where $(.)^H$ refers to the Hermitian of a matrix, $\mathbf{R}_{\mathbf{hh}} = E\left\{\mathbf{h}.\mathbf{h}^H\right\}$ and $\mathbf{R}_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} = E\left\{\tilde{\mathbf{n}}.\tilde{\mathbf{n}}^H\right\}$. Alternatively, we can apply the least minimum square (LMS) method and minimize the mean square error $\left|\tilde{\mathbf{H}} - \mathbf{H}\right|^2$. The solution is given by:

$$\tilde{\mathbf{H}}_{LMS} = \mathbf{Y} \cdot \hat{\mathbf{S}}^H \cdot \left(\hat{\mathbf{S}}\hat{\mathbf{S}}^H\right)^{-1} \quad (8)$$

As it can be seen from (7) and (8), both solutions require a matrix inversion, where the matrix and consequently its invertibility highly depends on the data, which is very undesirable. In [8], an iterative method was suggested to avoid the matrix inversion. The convergence of the method depends however on the SNR range. Besides the iterative algorithm generates a delay until a new channel estimate is available.

In the following, we adopt the simultaneous processing of all columns at a single time instant, i.e. once a symbol vector is detected, (4) is applied for all $j = 1, \ldots, M$ . By this means, we make use of the actual information available about the CSI at each time instant. Furthermore, we avoid numerical problems and complexities induced by matrix inversion and do not introduce any further delay that would impair the real-time functionality of the algorithm.

### C Noise suppression from the raw estimates

In (4) and under the assumption that the symbols $\hat{s}_j(k)$ are correctly detected, the new computed channel estimates are subject to two independent sources of noise. The first one is AWGN $\mathbf{n}(k)$ inherent to the received signal $\mathbf{y}(k)$. The second one is caused by the previous noisy channel estimates. Just after the PACE, the channel estimates are affected by a zero-mean gaussian noise with variance $\sigma_p^2$. In case of uncorrelated MIMO channel, we have $\sigma_p^2(\tau) = \frac{E_s 2\sigma_0^2}{L_t E_p} + 2(1 - J_0(2\pi f_d \tau))$, where $E_p$ and $E_s$ denote the energy of the pilot and data symbols respectively [1, 2]. Making the further assumption that

[1]Please note that in case of phase shift keying (PSK) with normalized symbol energy , i.e $|s_j(k)|^2 = 1$, (3) becomes $\tilde{h}_{ij}(k) = \hat{s}_j^* \left(y_i(k) - \sum_{q=1,q\neq j}^{M} \hat{h}_{iq}(i)\hat{s}_q(k)\right)$. This measure helps to get rid of the division operation needed in (3) but does not restrict the CE method to PSK-constellations.

[2]A sequential computation, i.e after the computation of a new column insert the new value in (4) for the computation of the next one for the same time instant does only change the value of the first processed column and is therefore not relevant.

PSK is used, the mean square error (MSE) of the raw channel estimates is given by:

$$e(k) = E\left[\left|\tilde{h}_{ij}(k) - h_{ij}(k)\right|^2\right] = (M-1)e(k-1) + \frac{2\sigma_0^2}{E_s} \quad (9)$$

which says that at each time instant *k*, $1 \leq k \leq L_d$, the MSE is composed of one part which depends on the CE error of the previous time instant *k-1* and a second part which is directly related to the AWGN of the receive signal and equals the inverse of the SNR per symbol [3].

In the following, the raw channel estimates have to be filtered to reduce the MSE in (9). Let $\tilde{\mathbf{h}}_{ij}(k, L)$ denote the vector consisting of the last $L$ raw estimates anterior to time instant $kT_s$, i.e. $\tilde{\mathbf{h}}_{ij}(k, L) = \begin{bmatrix} \tilde{h}_{ij}(k-L) & \tilde{h}_{ij}(k-L+1) & \cdots & \tilde{h}_{ij}(k-1) \end{bmatrix}^T$. We now determine the filter $\mathbf{w} = \begin{bmatrix} w_0 & w_1 & \cdots & w_{L-1} \end{bmatrix}^T$ such that the filtered decision-directed channel estimate

$$\hat{\tilde{h}}_{ij}(k) = \mathbf{w}^{\mathbf{T}} \tilde{\mathbf{h}}_{ij}(k, L) \quad (10)$$

satisfies an optimization criterion which we will define in the next paragraph.

### D Optimal Combining of the raw estimates

Considering $\hat{\tilde{h}}_{ij}(k)$ according to (10), the optimization problem can be described as follows:

$$\min_{\mathbf{w}} \left\{ E\left[\left|\hat{\tilde{h}}_{ij}(k) - h_{ij}(k)\right|^2\right]\right\} \quad (11)$$

The solution, also called Wiener solution [7, 4], is given by:

$$\mathbf{w} = \mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}^{-1}\mathbf{r} \quad (12)$$

where $\mathbf{r} = E\left\{\tilde{\mathbf{h}}_{ij}(k, L) h_{ij}^*(k)\right\}$ and $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = E\left\{\tilde{\mathbf{h}}_{ij}(k, L)\tilde{\mathbf{h}}_{ij}(k, L)^H\right\}$.

For the computation of $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}^{-1}$, the condition number of $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ is very important to determine the accuracy of the solution[4]. For a normalized Doppler frequency of about $10^{-2}$ or more, the condition number is still large enough such that the inverse matrix can be accurately determined. For smaller values of $f_{dnorm}$ and the channel model given in (2), the matrix is very bad-conditioned. The solution of (12) is therefore confronted with severe numerical problems. In [8], a similar problem was solved by a simple heuristic approach. To improve the conditioning of $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$, the values along the principal diagonal are increased by a very small positive amount $\epsilon$. For instance, choosing $\epsilon = 10^{-8}$ reduces the condition number from $10^{19}$ to $10^{12}$. This solution, although optimal in the MMSE sense, requires a further optimization with regard to the parameter $\epsilon$, which we do not follow in this paper.

[3]From (9), we can state the MSE dependency on the PACE MSE $e(0) = \sigma_p^2(T_s)$ as $e(k) = (M-1)^k e(0) + \frac{(M-1)^k}{M-2} \cdot \frac{2\sigma_0^2}{E_s}$.

[4]Note that in the high SNR range and for the channel model in (2) $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ is given by $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = Toeplitz\begin{bmatrix} 1 & J_0(2\pi f_d T_s) & \cdots & J_0(2\pi f_d (L-1)T_s) \end{bmatrix}$ and $\mathbf{r} = \begin{bmatrix} J_0(2\pi f_d L T_s) & \cdots & J_0(2\pi f_d T_s) \end{bmatrix}^T$.

### E Sub-optimal Combining of the raw estimates

In order to tackle the problem of numerical instabilities when applying the Wiener solution, we adopt a moving-average FIR filter, whose coefficients do not take account of $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$. In doing so, we notice that the length $L$ of the filter and its initialization are a key factor for the performance of the CE since we are dealing with a recursive processing. We further introduce the parameter $L_1$ as one index of the filter coefficients, e.i. $1 \leq L_1 \leq L$. We propose the following four different initialization procedures:

- *Procedure A (ProcA)*: Initialization with zero until $L$. Wait untill $L$ raw estimates are available for beginning with the DDCE. Meanwile use the pilot-assisted channel estimates. The filter coefficients are constant for all time instants and equal $\frac{1}{L}$.

- *Procedure B (ProcB)*: Initialization with pilot-assisted channel estimates until $L$ . Begin immediatly with the DDCE. The filter coefficients are constant for all time instants and equal $\frac{1}{L}$.

- *Procedure C (ProcC)*: Initialization with zero until $L$ . Begin with the DDCE when $L_1$raw estimates are available. Adapt the filter coefficients correspondingly.

- *Procedure D (ProcD)*: Initialization with pilot-assisted channel estimates untill $L_1$ and with zero from $L_1$ until $L$. Begin immediatly with the DDCE. Adapt the filter coefficients correspondingly.

The four different initialization procedures result in different mean CE errors. Consequently, the bit error ratio (BER) performances of the corresponding CE schemes also differ, which we will thoroughly discuss in Section III.

## III Parameter Optimization

We optimize the values for the parameters $L$ and $L_1$ by numerical evaluations. If not otherwise stated, we adopt the parameter set of Table 1. We generally use an ordered successive interference cancellation (OSIC) MMSE detector according to [9].

Table 1: parameter set for simulation

| $M \times N$ | $T_s$ | modulation | $(L_t, L_d)$ |
|---|---|---|---|
| $2 \times 2$ | $20\mu s$ | BPSK | $(2, 300)$ |

Denoting the modulation order as $Q$, the mean SNR per information bit is defined as $SNR_{dB} = 10 \cdot lg\left(\frac{M \cdot E_S}{Q \cdot 2\sigma_0^2}\right)$.

### A Filter length

The longer the filter, the better the suppression of noise will be. However, in fast varying channels the CSI becomes out of date. Therefore, the filter length $L$ has to take account of the channel coherence time. We evaluate the BER for different values of $f_{dnorm}$ by computer simulation. Fig. 3 shows the results for $f_{dnorm} = 2 \cdot 10^{-4}$ and $2 \cdot 10^{-2}$ , where ProcA is applied. Fig. 3 left shows that a lower error floor can be achieved by increasing $L$. However, exceeding a certain value, the CSI becomes out of date and the BER degrades. This effect occurs even more clearly for higher $f_{dnorm}$, as can be seen in Fig. 3 right. We define the optimal $L$ such that the temporal

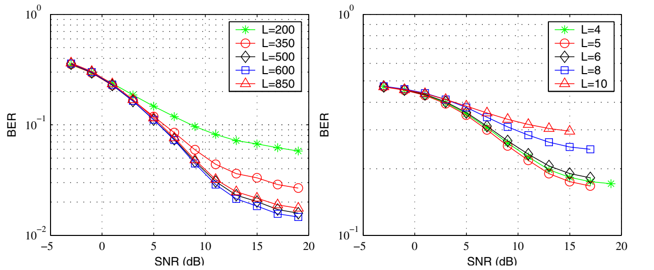


Figure 3: BER as a function of the SNR for various values of the filter length $L$ and for $f_{dnorm} = 2 \cdot 10^{-4}$ (left) and $f_{dnorm} = 2 \cdot 10^{-2}$ (right)

correlation function falls to its $q$ percentile. Using (2) and the approximation $J_0(x) \approx 1 - \frac{x^2}{4}$ for $x \ll 1$, $L$ satisfies:

$$L = \left\lfloor \frac{\sqrt{(1-q)}}{\pi f_d T_s} \right\rfloor \tag{13}$$

Intensive simulations for a wide range of $f_{dnorm}$ yield to $q = 90\%$.

### B Initialization parameter $L_1$

This parameter has to be optimized in the case of initialization procedures C and D. In Fig. 4, we plot the MSE as a function of $\tau$. All DDCE procedures lead to an MSE, which is much lower than the PACE MSE result. The curve shows furthermore that at the end of the frame all procedures yield approximately the same MSE, whereas differences can be noticed at the beginning of the frame for an SNR of -3 dB, see Fig. 4 top right. For a higher SNR of 12 dB for instance, the differences at the beginning of the frame become significantly smaller as illustrated in Fig. 4 bottom right. The same effect is noticed for ProcD. We can therefore conclude that the choice of $L_1$ is crucial in the low SNR range and at the beginning of a frame, whereas in the high SNR range and for long data frames, $L_1$ does not impact the MSE significantly. By means of intensive simulations, we determined the optimal $L_1$ for different $f_{dnorm}$. For instance, using ProcD at $f_{dnorm} = 4 \cdot 10^{-3}$, the optimal value is $L_1 = 15$.

### C Initialization procedure

Now that all parameters are optimized, we deal with the question which initialization procedure to choose. Fig. 5 shows the CE MSE as a function of $\tau$ in a frame of length $L_d = 300$. The differences in the four procedures occur at the beginning of the frame in the low SNR range. At the end of the frame, all of them result in the same MSE. In the high SNR range, ProcA leads to the largest MSE on the whole frame whereas ProcC and ProcD perform best. These MSE results are also reflected in the BER performance as can be seen in Fig. 6. In the low SNR range, all precedures perform similarily whereas in the high SNR range ProcC and ProcD result in the lowest BER floor. These results let us expect the discrepancy between the different intialization procedures to be even more accentuated if the frame length is smaller, which will be analyzed in Section IV.

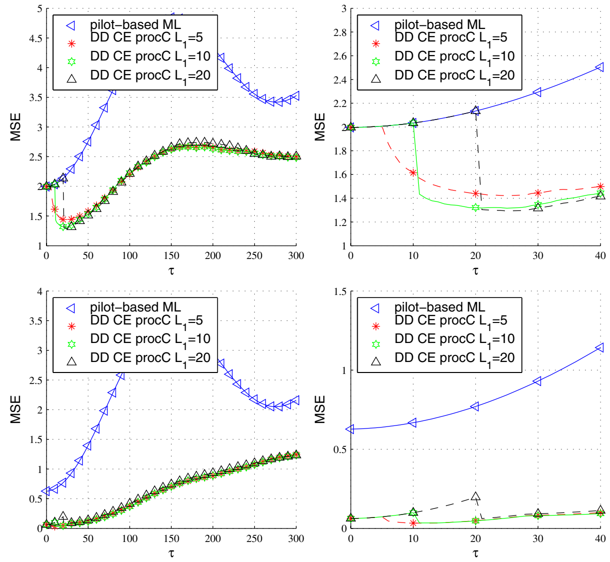


Figure 4: MSE as a function of $\tau$ for ProcC at $f_{dnorm} = 4 \cdot 10^{-3}$ and $SNR = -3dB$ (top left). Zoom for $0 \leq \tau \leq 40$ (top right). MSE as a function of $\tau$ for ProcC at $f_{dnorm} = 4 \cdot 10^{-3}$ and $SNR = 12\ dB$ (bottom left). Zoom for $0 \leq \tau \leq 40$ (bottom right)

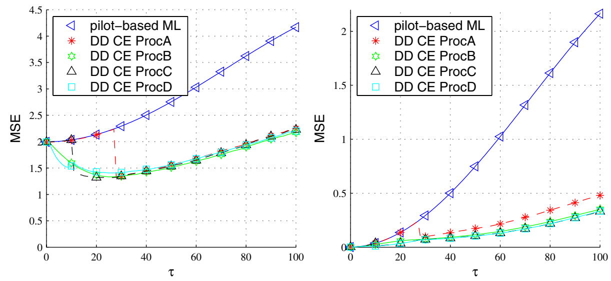


Figure 5: MSE as a function of $\tau$ for the different initialization procedures with optimized parameters at $f_{dnorm} = 4 \cdot 10^{-3}$, SNR = -3 dB (right), SNR = 30 dB (left)

## IV RESULTS

In Section III, we optimized the filter design numerically. In the following, we investigate the impact of some system parameters on the BER.

First, we consider different frame lengths. As expected from the MSE consideration in Section III, the BER difference between ProcC and ProcD is accentuated for smaller data frames. Fig. 7 shows that whereas for $L_d = 300$ no difference between the BER curves can be noticed, the better performance of ProcD becomes evident for $L_d = 50$ and even clearer for $L_d = 35$.

In this paragraph, we investigate the performance of the CE algorithm with four different MIMO detectors, namely, ZF, MMSE, OSIC ZF and OSIC MMSE. Because of the recursive algorithm, the better the detector, the more accurate the detected symbols will be and the better the decision-directed channel estimates are. Therefore, we expect a larger gain for OSIC detectors, which is confirmed by simulations in Fig. 8.

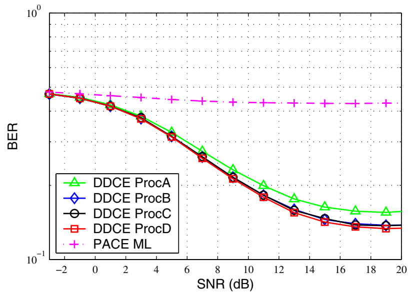


Figure 6: BER as a function of SNR for the different initialization procedures at $f_{dnorm} = 4 \cdot 10^{-3}$. $L = 27$ and $L_1 = 15$

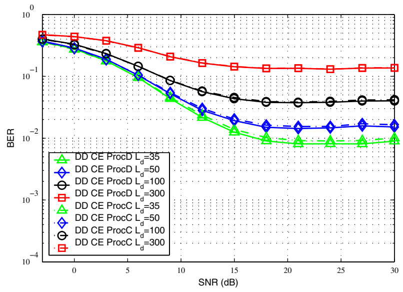


Figure 7: BER as a function of SNR for various frame lengths using ProcC and ProcD with optimized parameters at $f_{dnorm} = 4 \cdot 10^{-3}$

We want to emphasize that our method is applicable indepen-

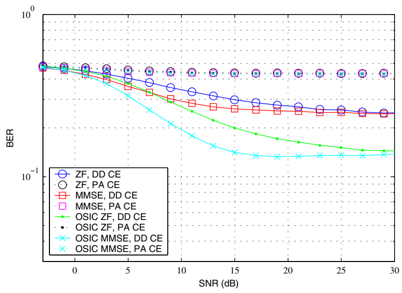


Figure 8: BER as a function of SNR for different receiver types with ProcD at $f_{dnorm} = 4 \cdot 10^{-3}$

dently of the detection scheme, which is a great advantage. We also simulated the ideal case where all received symbols are error free, which gives us an upper bound. In this case, only the noise inherent to the PACE and the AWGN term contribute to the CE error. From Fig. 9, we conclude that the error propagation induced by wrongly detected symbols impairs the BER significantly. The gain achieved by perfect detection is even higher at higher Doppler frequencies.

## V Conclusion

We presented a combined pilot-assisted and decision-directed channel estimation scheme that is capable of improving the channel estimates making use of the minimal required number of pilots. The algorithm is not only bandwidth-efficient, but also numerically robust and significantly reduces the BER floor induced by the channel estimation errors for pilot-assisted

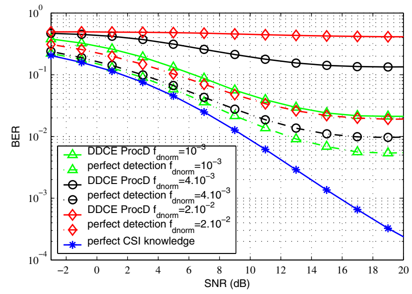


Figure 9: BER as a function of SNR with perfect detection using ProcD with $L_1 = 15$ and $L = 27$

methods only. We also showed that a moving-average filter with optimized parameters, such as length and initialization, can further improve the performance.

## References

[1] V. Pohl, V. Jungnickel, and C. von Helmolt, "Optimal length of the training and data phase in continuous flat fading mimo channels," *IEEE Intern. Conf. on Commun. (ICC)*, vol. 6, pp. 3153–3157, June 2004.

[2] Q. Sun, D. Cox, H. Huang, and A. Lozano, "Estimation of continuous flat fading mimo channels," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 549–553, October 2002.

[3] M. Kiessling and J. Speidel, "Statistical transmit processing for enhanced mimo channel estimation in presence of correlation," *IEEE International Conference on Global Communications (GLOBECOM)*, December 2003.

[4] J. Gao and H. Liu, "Decision-directed estimation of mimo time-varying rayleigh fading channels," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1412–1417, July 2005.

[5] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, Inc., 1996, ISBN 0-13-322760-X.

[6] A. Grant, "Joint decoding and channel estimation for linear mimo channels,"o *Wireless Commun. and Networking Conf.*, vol. 3, pp. 1009–1012, September 2000.

[7] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993, ISBN 0-13-345711-7.

[8] X. Deng, A. M. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for mimo systems," *Proceedings of IEEE International Conference on Communications (ICC 03)*, vol. 4, no. 4, pp. 2326– 2329, May 2003.

[9] A. Boronka, "Analysis and results on optimized mimo detection with theshold-based adaptive cancellation," *IEEE Vehicular Technology Conference (VTC)*, May 2004.

[10] K. E. Baddour and N. C. Beaulieu, "Autoregressive models for fading channel simulation," *Proc. IEEE Global Telecommun. Conf.*, vol. 2, pp. 1187–1192, 2001.