Analysis of Semiblind Channel Estimation for FIR-MIMO Systems

Tianbin Wo[1], Peter Adam Hoeher[2], Ansgar Scherb[3], Karl-Dirk Kammeyer[4]

[1,2] University of Kiel, Germany, {wtb,ph}@tf.uni-kiel.de

[3,4] University of Bremen, Germany, {scherb,kammeyer}@ant.uni-bremen.de

Abstract

This paper deals with joint data detection and channel estimation for frequency-selective multiple-input multiple-output (MIMO) systems with focus on the analysis of the channel estimator. First, we present a scheme alternating between joint Viterbi detection and least squares channel estimation and analyze its performance in terms of unbiasedness. Since in the proposed technique the channel estimator exploits both known pilot symbols (non-blind) as well as unknown information bearing symbols (blind), this channel identification scheme is referred to as semiblind. Second, we derive the Cramer-Rao lower bound (CRLB) for semiblind channel estimation of frequency-selective MIMO channels, which provides a theoretical lower bound of the achievable mean squared error (MSE) of any unbiased estimator. By simulation the MSE performance of the proposed algorithm is evaluated and compared to the CRLB.

1 Introduction

An important obstacle for achieving the potential capacity of MIMO systems is the difficulty in acquiring reliable channel state information. If a purely training-based scheme is used, a large amount of training symbols will be necessary for obtaining reliable estimates of the large number of channel coefficients, which will significantly degrade the system bandwidth efficiency. In comparison, a semiblind channel estimator (SBCE) tries to extract the channel information carried by all observations, and is able to achieve very low MSE with using just a few training symbols. Due to its amazing performance, SBCE has recently attracted increasing attentions. Simulation results of such type of channel estimators can be found in a number of papers [1]-[4], but detailed theoretical analysis is still not sufficient, especially for FIR-MIMO channels.

In this paper, we try to give a thorough analysis of SBCE with focus on the MSE performance. We consider block-fading frequency-selective MIMO channels. The obtained results are universal for systems with an arbitrary number of antennas and an arbitrary channel memory length. As an example, a SBCE algorithm with least squares channel estimator and maximum likelihood data detector will be first introduced and analyzed. It will be shown that the presented semiblind channel estimator is biased at low SNR but tends to be unbiased at high SNR. Please note that this statement is also true if the system adopts a linear data detector or even a successive data detector. Since the estimator is unbiased at high SNR, its MSE performance will be limited by the CRLB. Scherb et al. derived the CRLB for SBCE in the case of single-input single-output (SISO) systems, while in this paper we extend the derivation to the case of frequency-selective MIMO systems. As the obtained analytical expression of CRLB involves high order integration, we provide reasonable approximations for low SNR and high SNR, respectively. Interestingly but reasonably, the minimum mean squared error achievable by any unbiased channel estimator at high SNR will be the same as that all data symbols are a-priori known at the receiver, but only the training symbols are known at low SNR. We thereafter provide simulation results to show that the MSE performance of the presented SBCE coincides with the CRLB at high SNR but exceeds CRLB at low SNR due to biasing. Of particular interest is the SNR value where a semiblind channel estimator begin to approach the CRLB, which means that a SBCE will be able to fully exploit the channel information carried by all observations for SNRs larger than this value. Under this motivation, we derive an analytical expression for determining this SNR value.

The rest of this paper is organized as follows. In Sec. 2, a general MIMO channel model is illustrated. Sec. 3 describes a practical SBCE algorithm, while Sec. 4 concentrates on the derivation of the CRLB for SBCE. In Sec. 5, numerical results are compared with the theoretical bounds, and detailed performance analysis is provided. Finally, Sec. 6 gives the conclusion.

Throughout this paper, we use $(\cdot)^H$ to denote Hermitian conjugate. $\mathrm{tr}\{\cdot\}$ denotes the trace of a matrix. $\otimes$ is the Kronecker product operator, and $\mathrm{vec}\{\cdot\}$ is the column stacking operator, with $\mathrm{vec}\{\mathbf{A}\} = [\mathbf{a}_1^T, \cdots, \mathbf{a}_N^T]^T$, where $A = [\mathbf{a}_1, \ldots, \mathbf{a}_N]$ is an $M \times N$ matrix.

This work has been supported by the German Research Foundation (DFG) under contract no. HO 2226/6-1.

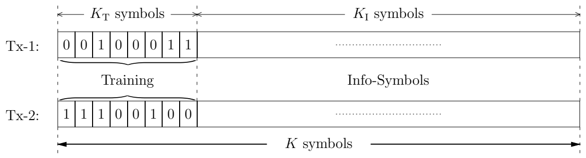

Fig. 1. Burst structure, $N_\mathrm{T} = 2$, $L = 2$

2 Channel Model

The equivalent discrete-time model of a frequency-selective $(N_\mathrm{R} \times N_\mathrm{T})$-MIMO channel (including transmit and receive filter, physical channel and baud-rate sampling) is given by

$$\mathbf{r}(k) = \sum_{l=0}^{L} \mathbf{H}(l)\mathbf{s}(k-l) + \mathbf{n}(k), \tag{1}$$

where k is the discrete time index, and $\mathbf{s}(k) \in \mathbb{C}^{N_\mathrm{T}\times 1}$ denotes the symbol vector. The channel output vector is given by $\mathbf{r}(k) \in \mathbb{C}^{N_\mathrm{R}\times 1}$ and $\mathbf{n}(k) \in \mathbb{C}^{N_\mathrm{R}\times 1}$ is a complex zero-mean white Gaussian noise vector with covariance $\mathrm{E}\{\mathbf{n}(k)\mathbf{n}^H(k)\} = \sigma_n^2 \mathbf{I}_{N_\mathrm{R}}$. The frequency-selective channel $\mathbf{H}(l) \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{T}}$ has order L and is assumed to be constant over K symbol periods (block fading). Hence, the transmission of K consecutive symbol vectors can be compactly rewritten as

$$\mathbf{R} = \mathbf{H}\mathbf{S} + \mathbf{N}, \tag{2}$$

where $\mathbf{R} = [\mathbf{r}(0), \cdots, \mathbf{r}(K+L-1)]$, $\mathbf{H} = [\mathbf{H}(0), \cdots, \mathbf{H}(L)]$ and $\mathbf{N} = [\mathbf{n}(0), \cdots, \mathbf{n}(K+L-1)]$. The matrix $\mathbf{S}$ is of block Toeplitz form, i.e.

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}(0) & \cdots & \mathbf{s}(K-1) & & \mathbf{0} \\ & \ddots & & & \ddots \\ \mathbf{0} & & \mathbf{s}(0) & \cdots & \mathbf{s}(K-1) \end{bmatrix}. \tag{3}$$

Given the burst structure shown in Fig. 1, the symbol matrix may be written as a horizontal concatenation of two sub-blocks

$$\mathbf{S} = [\mathbf{S}_\mathrm{T}, \mathbf{S}_\mathrm{I}], \tag{4}$$

where $\mathbf{S}_\mathrm{T}$ contains training symbols only, while $\mathbf{S}_\mathrm{I}$ contains data symbols and several training symbols due to the Toeplitz structure of $\mathbf{S}$. Correspondingly, the channel output can also be written as a horizontal concatenation

$$\mathbf{R} = [\mathbf{R}_\mathrm{T}, \mathbf{R}_\mathrm{I}] = [\mathbf{H}\mathbf{S}_\mathrm{T} + \mathbf{N}_\mathrm{T}, \mathbf{H}\mathbf{S}_\mathrm{I} + \mathbf{N}_\mathrm{I}]. \tag{5}$$

This notation is given to ease the tasks of algorithm description and performance analysis.

3 Semiblind Channel Estimation

In comparison with traditional algorithms, a SBCE utilizes not only the training but also the data symbols to perform channel estimation. Since the data symbols are unknown at the receiver, data detection becomes a necessary step. Borrowing ideas from Turbo decoding, the procedure of SBCE can be done iteratively:

1) Initial training-based channel estimation;
2) Given channel knowledge, detect data symbols;
3) Given data knowledge, perform channel estimation by taking the whole block as virtual training;
4) Repeat step 2 and step 3 until a certain stopping criterion is reached.

In the following, the issue of data detection and channel estimation for MIMO systems will be tackled respectively, and then an analysis on the biasing is provided.

3.1 Joint Viterbi Detection

Given perfect channel knowledge, the optimal data detector for a MIMO system in the sense of the ML criterion is so called joint Viterbi detector (JVD) [5][6], which searches for the most likely data sequence according to the following formula:

$$\widehat{\mathbf{S}} = \arg\min_{\widetilde{\mathbf{S}}\in\mathcal{S}} \left\{ \|\mathbf{R} - \mathbf{H}\widetilde{\mathbf{S}}\|_\mathrm{F}^2 \right\}. \tag{6}$$

JVD is an extension of the well-known trellis-based data detector for SISO systems to the case of MIMO systems, and delivers better performance compared to linear algorithms, such as zero-forcing and minimum mean squared error. Due to its non-linear property, there is no requirements on the relationship between N_R and N_T, that is the number of receive antennas is allowed to be less than the number of transmit antennas. However, JVD suffers from exponentially growing complexity w.r.t. $N_\mathrm{T}(L+1)$ and hence often becomes computationally prohibitive in practice. Since in this paper we focus on the analysis of theoretical MSE bounds, JVD will be used as the data detector for the sake of easy analysis and clearity.

3.2 Least Squares Channel Estimation

A least squares (LS) channel estimator minimizes the distance between the channel output and its noiseless hypotheses, written as

$$\begin{aligned} \widehat{\mathbf{H}} &= \arg\min_{\widetilde{\mathbf{H}}\in\mathcal{H}} \left\{ \|\mathbf{R} - \widetilde{\mathbf{H}}\mathbf{S}\|_\mathrm{F}^2 \right\} \\ &= \mathbf{R}\mathbf{S}^H \left(\mathbf{S}\mathbf{S}^H\right)^{-1}, \end{aligned} \tag{7}$$

which in turn maximizes the likelihood function $p(\mathbf{R}|\widetilde{\mathbf{H}}, \mathbf{S})$, and is often also called ML channel estimator. For SBCE, the initial channel estimation is performed over the training only:

$$\widehat{\mathbf{H}} = \mathbf{R}_\mathrm{T}\mathbf{S}_\mathrm{T}^H \left(\mathbf{S}_\mathrm{T}\mathbf{S}_\mathrm{T}^H\right)^{-1}, \tag{8}$$

while in later iterations, the knowledge of data symbols will also be utilized:

$$\widehat{\mathbf{H}} = \mathbf{R}\widehat{\mathbf{S}}^H \left(\widehat{\mathbf{S}}\widehat{\mathbf{S}}^H\right)^{-1}, \tag{9}$$

with $\widehat{\mathbf{S}} = [\mathbf{S}_\mathrm{T}, \widehat{\mathbf{S}}_\mathrm{I}]$. If the number of symbol errors in $\widehat{\mathbf{S}}_\mathrm{I}$ is small enough, (9) will hopefully produce a better channel estimate than (8).

3.3 On the Bias of SBCE

It is easy to find that (8) always delivers unbiased channel estimates, while it remains to be an interesting question whether $\widehat{\mathbf{H}}$ in (9) is biased or not. Considering sufficient large block lengths, the following assumption

$$\widehat{\mathbf{S}}\widehat{\mathbf{S}}^H \approx K\mathbf{I}_{N_\mathrm{T}(L+1)} \tag{10}$$

should be valid. Then we may approximate (9) by

$$\begin{aligned}\widehat{\mathbf{H}} &\approx \frac{1}{K}\mathbf{R}\widehat{\mathbf{S}}^H \\ &= \frac{1}{K}[\mathbf{H}\mathbf{S}_\mathrm{T}+\mathbf{N}_\mathrm{T},\mathbf{H}\mathbf{S}_\mathrm{I}+\mathbf{N}_\mathrm{I}][\mathbf{S}_\mathrm{T},\widehat{\mathbf{S}}_\mathrm{I}]^H \qquad (11)\\ &= \frac{1}{K}\Big(\mathbf{H}(\mathbf{S}_\mathrm{T}\mathbf{S}_\mathrm{T}^H+\mathbf{S}_\mathrm{I}\widehat{\mathbf{S}}_\mathrm{I}^H)+\mathbf{N}_\mathrm{T}\mathbf{S}_\mathrm{T}^H+\mathbf{N}_\mathrm{I}\widehat{\mathbf{S}}_\mathrm{I}^H\Big).\end{aligned}$$

Given optimal training [7], and letting P_s denote the symbol error rate at the output of the data detector, we have

$$\mathbf{S_T}\mathbf{S_T}^H = K_\mathrm{T}\mathbf{I}_{N_\mathrm{T}(L+1)} \tag{12}$$

$$\mathrm{E}\big\{\mathbf{S}_\mathrm{I}\widehat{\mathbf{S}}_\mathrm{I}^H\big\} = (1-2P_s)K_\mathrm{I}\mathbf{I}_{N_\mathrm{T}(L+1)}. \tag{13}$$

Let $\mathbf{E_I} = \widehat{\mathbf{S}}_\mathbf{I} - \mathbf{S_I}$ denote the symbol estimation errors. The mean value of $\widehat{\mathbf{H}}$ can now be written as

$$\mathrm{E}\big\{\widehat{\mathbf{H}}\big\} \approx \frac{K_\mathrm{T}+(1-2P_s)K_\mathrm{I}}{K}\mathbf{H} + \frac{\mathrm{E}\{\mathbf{N}_\mathrm{I}\mathbf{E}_\mathrm{I}^H\}}{K}, \tag{14}$$

where the right-most term indicates the correlation between the noise and the symbol estimation errors. It is clear that $\widehat{\mathbf{H}}$ tends to be unbiased for very high SNR, because both P_s and $\mathbf{E}_\mathrm{I}$ will become zero, while at low SNR it is biased by a positive real scaling factor and by an additive noise-error cross correlation matrix. Contrary to the statements in [8], $\mathrm{E}\{\mathbf{N}_\mathrm{I}\mathbf{E}_\mathrm{I}^H\}$ should not be assumed to be zero, since the noise is exactly the cause of data detection errors. This statement will be attested by the numerical results provided in Sec. 5.

4 Cramer-Rao Lower Bound

As the presented SBCE is unbiased at high SNR, its MSE performance will be limited by the well-known CRLB. In the following we give a brief overview on the main results of [9] dealing with the CRLB for semiblind channel estimation. Due to the lack of space we do not repeat any derivation within this paper. The interested reader may find details in the above mentioned reference. Although [9] deals with SISO systems, a generalization to MIMO systems is straight forward by employing the vectorized channel model:

$$\underline{\mathbf{r}}^H = \underline{\mathbf{h}}^H\underline{\mathbf{S}} + \underline{\mathbf{n}}^H, \tag{15}$$

where $\underline{\mathbf{r}} = \mathrm{vec}(\mathbf{R}^H)$, $\underline{\mathbf{h}} = \mathrm{vec}(\mathbf{H}^H)$, $\underline{\mathbf{S}} = \mathbf{I}_{N_\mathrm{R}} \otimes \mathbf{S}$ and $\underline{\mathbf{n}} = \mathrm{vec}(\mathbf{N}^H)$.

4.1 Preliminaries

The set $\mathcal{S}$ is defined as a finite set over all possible realizations of $\underline{\mathbf{S}}$, where $\underline{\mathbf{S}}_i$ is the i-th element of $\mathcal{S}$. Let M be the amount of binary information carried by $\underline{\mathbf{S}}$, which depends on modulation index, block length and number of transmit antennas. Then the cardinality of $\mathcal{S}$ is given by $|\mathcal{S}| = 2^M$. Each element $\underline{\mathbf{S}}_i \in \mathcal{S}$ is assumed to be equally probable with

$$p(\underline{\mathbf{S}}) = \frac{1}{|\mathcal{S}|}. \tag{16}$$

In order to obtain a validated statement by the CRLB, it has to be guaranteed that the estimation problem has an unique solution for all $\underline{\mathbf{S}}_i \in \mathcal{S}$. For this reason a necessary condition is that

$$\underline{\mathbf{h}}^H\underline{\mathbf{S}}_i \neq \underline{\mathbf{h}}^H\underline{\mathbf{S}}_j \quad \forall\ \underline{\mathbf{S}}_i \neq \underline{\mathbf{S}}_j, \tag{17}$$

which in turn means that $\underline{\mathbf{h}} \neq \mathbf{0}$ and $\underline{\mathbf{S}}_i$ has to be full row rank. Considering a proper choice of pilots, the full rank requirement can be always satisfied.

4.2 Derivation of the CRLB

The CRLB can be stated as follows: Let $\hat{\underline{\mathbf{h}}}$ be an unbiased estimator of the channel parameter $\underline{\mathbf{h}}$ and

$$\mathbf{C}_{\hat{\underline{\mathbf{h}}}} = \mathrm{E}\{(\underline{\mathbf{h}} - \hat{\underline{\mathbf{h}}})(\underline{\mathbf{h}} - \hat{\underline{\mathbf{h}}})^H\} \tag{18}$$

be the covariance of the estimator. Then the covariance matrix of any unbiased estimator satisfies

$$\mathbf{C}_{\hat{\underline{\mathbf{h}}}} - \mathbf{F}^{-1} \geq \mathbf{0}, \tag{19}$$

where "$\geq \mathbf{0}$" stands for positive-semidefinite, and $\mathbf{F}$ is the Fisher information matrix (FIM) [10]. Since the variance of each entry in $\hat{\underline{\mathbf{h}}}$ is given by the main diagonal of $\mathbf{C}_{\hat{\underline{\mathbf{h}}}}$, the minimum MSE obtained by any unbiased estimator can be directly derived from the inverse of FIM by $\mathrm{MSE}_{\min} = \mathrm{tr}\{\mathbf{F}^{-1}\}$. The FIM for the semiblind channel estimation problem is given by

$$\mathbf{F} = \frac{1}{|\mathcal{S}|\sigma_n^2}\cdot \int_{\mathbb{C}^{(L+K)N_\mathrm{R}}} \frac{\sum\limits_{\underline{\mathbf{S}}_i,\underline{\mathbf{S}}_j\in\mathcal{S}} \underline{\mathbf{S}}_i\mathbf{d}_i\mathbf{d}_j^H\underline{\mathbf{S}}_j^H p(\underline{\mathbf{r}}|\underline{\mathbf{S}}_i,\underline{\mathbf{h}})p(\underline{\mathbf{r}}|\underline{\mathbf{S}}_j,\underline{\mathbf{h}})}{\sum\limits_{\underline{\mathbf{S}}_i\in\mathcal{S}} p(\underline{\mathbf{r}}|\underline{\mathbf{S}}_i,\underline{\mathbf{h}})} d\underline{\mathbf{r}}, \tag{20}$$

where

$$\mathbf{d}_i^H = \frac{\underline{\mathbf{r}}^H - \underline{\mathbf{h}}^H\underline{\mathbf{S}}_i}{\sigma_n} \tag{21}$$

and $p(\underline{\mathbf{r}}|\underline{\mathbf{S}}_i,\underline{\mathbf{h}})$ is the conditional PDF to observe $\underline{\mathbf{r}}$ under the assumption that $\underline{\mathbf{S}}_i$ was transmitted over the channel $\underline{\mathbf{h}}$.

Integration of this expression can only be done numerically, where high accuracy requires high computational effort. However, a more intuitive insight of the CRLB is provided by asymptotical expressions of the FIM at low and high SNR. At high SNR, namely if σ_n^2 tends to zero, the FIM is given by

$$\mathbf{F}_{\text{high}} = \frac{1}{|\mathcal{S}|\sigma^2} \sum_{\underline{\mathbf{S}}_i \in \mathcal{S}} \underline{\mathbf{S}}_i \underline{\mathbf{S}}_i^H, \quad (22)$$

whereas at low SNR, namely if σ_n^2 tends to infinity, the FIM is given by

$$\mathbf{F}_{\text{low}} = \frac{1}{|\mathcal{S}|^2\sigma^2} \sum_{\underline{\mathbf{S}}_i, \underline{\mathbf{S}}_j \in \mathcal{S}} \underline{\mathbf{S}}_i \underline{\mathbf{S}}_j^H. \quad (23)$$

Thus, the minimum MSE of any unbiased estimator is within the range $\mathrm{tr}(\mathbf{F}_{\text{high}}^{-1}) \leq \mathrm{MSE}_{\min} \leq \mathrm{tr}(\mathbf{F}_{\text{low}}^{-1})$. Letting $\underline{\mathbf{S}}_T$ be the training part of $\underline{\mathbf{S}}$, (22) and (23) can be further simplified to

$$\mathbf{F}_{\text{high}} = \frac{1}{\sigma_n^2}\Big(\underline{\mathbf{S}}_{\mathrm{T}}\underline{\mathbf{S}}_{\mathrm{T}}^H + K_{\mathrm{I}}\mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}(L+1)}\Big). \quad (24)$$

and

$$\mathbf{F}_{\text{low}} = \frac{1}{\sigma_n^2}\underline{\mathbf{S}}_{\mathrm{T}}\underline{\mathbf{S}}_{\mathrm{T}}^H, \quad (25)$$

respectively. It is remarkable that $\mathbf{F}_{\text{low}}$ depends only on the pilot part $\underline{\mathbf{S}}_{\mathrm{T}}$. Therefore, in the low SNR regime the semiblind CRLB is identical to the CRLB of a pilot aided channel estimator. In contrast, $\mathbf{F}_{\text{high}}$ takes all available symbols into account. Hence, in the high SNR regime the semiblind CRLB is close to the case of perfectly knowing the values of each symbol at the estimator.

4.3 Saturation Point

Of particular interest is the range of SNR values, at which the asymptotical expressions (22) and (23) provide sufficiently accurate approximations of the true CRLB. Contrary to [9], in this paper we will restrict our focus on the high SNR case, since (as shown in Sec. 3.3) only at high SNR values the SBCE is unbiased and, consequently, bounded by the CRLB. Note that at low SNR the MSE performance of the SBCE may be superior to the CRLB due to its bias.

In the following we present a SNR threshold called "saturation point", after which the true CRLB is sufficiently close to its high SNR approximation. As shown in [9], this threshold depends on the minimum squared distance $\delta_{\min}$[1] between two adjacent transmit hypothesis defined by

$$\delta_{\min} = \min_{i \neq j} \delta_{i,j}, \quad (26)$$

where

$$\delta_{i,j} = \|\underline{\mathbf{h}}^H(\underline{\mathbf{S}}_i - \underline{\mathbf{S}}_j)\|^2 = \|\mathbf{H}^H(\mathbf{S}_i - \mathbf{S}_j)\|_{\mathrm{F}}^2. \quad (27)$$

Let $a \in [0, 1]$ be a parameter quantifying the similarity of the true CRLB to the low and high SNR approximation, where $a \to 1$ means that the true CRLB is close to its low SNR approximation and $a \to 0$ means that the true CRLB is close to its high SNR approximation. The SNR value, at which the true CRLB has weight a, can be approximated by

$$\mathrm{SNR}_a(\delta_{\min}) = -\frac{4\log(a)}{\delta_{\min}}. \quad (28)$$

Experience has shown that a rather good match for the desired saturation point can be found by setting $a = 0.05$ in (28).

For stochastic channel models $\delta_{\min}$ is a random variable depending on the momentary channel realization. In this case the mean of $\delta_{\min}$ given by

$$\bar{\delta}_{\min} = \mathrm{E}\{\delta_{\min}\} \quad (29)$$

may be taken for further analysis. As derived in Appendix A, an upper limit and also an accurate approximation of $\bar{\delta}_{\min}$ for complex Gaussian frequency-selective MIMO channels of length $Q = L + 1$ (each channel gain is i.i.d. with $h_{m,n}(l) \sim \mathcal{CN}(0, 1/Q)$) is given by

$$\bar{\delta}_{\min} = \frac{wN_{\mathrm{T}}}{Q}\int_0^\infty \frac{e^{-\delta}\delta^{QN_{\mathrm{R}}}[\Gamma(QN_{\mathrm{R}}, \delta)]^{N_{\mathrm{T}}-1}}{[(QN_{\mathrm{R}}-1)!]^{N_{\mathrm{T}}}}d\delta, \quad (30)$$

where w is the minimum squared distance between two admissible data symbols and $\Gamma(a, b)$ is the incomplete gamma function defined in Appendix A. In case of BPSK mapping the minimum squared distance is $w = 4$. Please note that (30) does not hold for channel coding and space-time coding, which may have an enormous effect on $\delta_{\min}$ and, consequently, may shift the saturation point to significantly lower SNR values. Eq. (30) is analytically solvable, but the solution may become rather complex for high numbers of receive and transmit antennas and for large channel lengths. In Tab. I we have compared the values calculated by (30) with simulated means of $\delta_{\min}$ for certain MIMO setups. Apparently, the analytically obtained values are rather close to the simulated values and the accuracy increases for large numbers of receive antennas. Since the degree of diversity becomes higher with increasing channel length, the values of frequency-selective systems are slightly larger than in the correponding non-frequency-selective cases.

TABLE I

$\bar{\delta}_{\min}$ AND THE CORRESPONDING SATURATION POINTS ($a = 0.05$) FOR FREQUENCY-SELECTIVE CHANNELS ($L = 2$)

N_{T}	N_{R}	simul. $\bar{\delta}_{\min}$	anal. $\bar{\delta}_{\min}$	simul. SNR_a [dB]	anal. SNR_a [dB]
2	2	6.1731	6.1953	2.8806	2.865
2	4	13.3815	13.4211	-0.4794	-0.4923
2	8	28.3170	28.3339	-3.7348	-3.7374

5 Numerical Results

In this section, we will provide numerical results concerning the biasing, the mean squared error and the

[1]Please note that the definition of $\delta_{i,j}$ and $\delta_{\min}$ differ from [9] for the sake of readability.

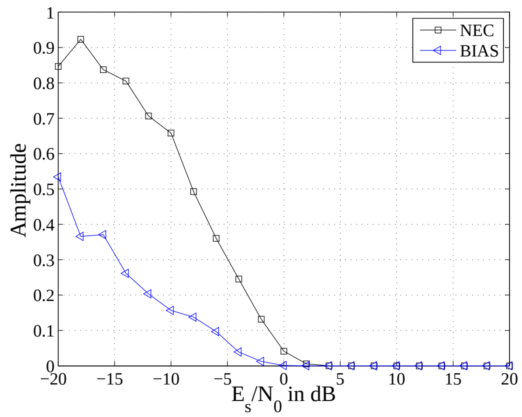

Fig. 2. BIAS vs. E_s/N_0, $N_\mathrm{R}=2$, $K=148$

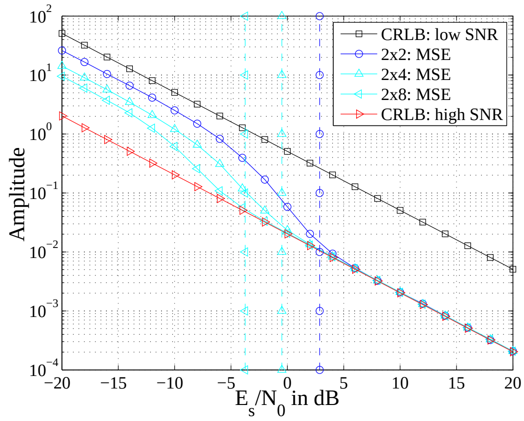

Fig. 3. MSE and CRLB, $K=148$

Cramer Rao lower bound, respectively. Specifically, we choose a system with $N_\mathrm{T}=2$ and $L=2$ as the simulation platform, and use $K_\mathrm{T}=8$ training symbols per burst, which is actually the minimum number of training symbols in order to perform the proposed SBCE algorithm [4]. Since flat-fading channel is just a special case ($L=0$), all the analysis and theoretical conclusions in the following hold for flat channels as well.

5.1 Biasing of SBCE

In order to measure the biasing of the semi-blind channel estimator, we define the degree of biasing as

$$\mathbf{BIAS} \doteq \|\mathbf{H}-\mathrm{E}\{\widehat{\mathbf{H}}\}\|_\mathrm{F}^2. \tag{31}$$

Given a randomly chosen channel matrix

$$\mathbf{H}=[\mathbf{H}(0),\mathbf{H}(1),\mathbf{H}(2)], \tag{32}$$

where

$$\begin{aligned}\mathbf{H}(0) &= \begin{bmatrix}0.32-0.39j & 0.40-0.50j\\ 0.23-0.15j & -0.21-0.02j\end{bmatrix}\\ \mathbf{H}(1) &= \begin{bmatrix}-0.34-0.48j & 0.13-0.46j\\ -0.11-0.43j & 0.10-0.55j\end{bmatrix}\\ \mathbf{H}(2) &= \begin{bmatrix}-0.48+0.60j & 0.01-0.11j\\ -0.90+0.02j & -0.41+0.39j\end{bmatrix},\end{aligned} \tag{33}$$

the simulation results of **BIAS** versus SNR is provided in Fig. 2. The statement in Section 3.3 is hence proved, that the presented semiblind channel estimator is biased at low SNR while unbiased at high SNR. Besides, the curve of **NEC** $\doteq \|\mathrm{E}\{\mathbf{N}_\mathrm{I}\mathbf{E}_\mathrm{I}^H\}/K\|_\mathrm{F}^2$ shows that the noise-error correlation is not negligible. Indeed, the value of **NEC** is significant w.r.t. the value of **BIAS**.

5.2 Cramer-Rao Lower Bound

From (22) and (23), we can obtain the CRLB approximation for high SNR as

$$\mathrm{MSE} \geq \mathrm{tr}\{\mathbf{F}_\mathrm{high}^{-1}\} \approx \frac{N_\mathrm{R}N_\mathrm{T}(L+1)}{K}\sigma_n^2, \tag{34}$$

and the CRLB approximation for low SNR as

$$\mathrm{MSE} \geq \mathrm{tr}\{\mathbf{F}_\mathrm{low}^{-1}\} = N_\mathrm{R}\sigma_n^2\mathrm{tr}\{(\mathbf{S}_\mathrm{T}\mathbf{S}_\mathrm{T}^H)^{-1}\}. \tag{35}$$

As we may notice that both approximations are linear w.r.t. the noise power, we can expect two parallel straight lines in a log-plot, which is demonstrated in Fig. 3. At the high SNR range, the MSE curves approach the lower bound. In this situation, the discussed semi-blind channel estimator tends to be unbiased, hence its performance is exactly bounded by the CRLB. However, at low SNR, the MSE curve is even lower than the CRLB approximation. Fortunately, this phenomenon can be explained by the biasing of the estimator at low SNR.

According to Tab. I, the vertical dashed lines in Fig. 3 mark the lowest possible SNR thresholds, where SBCE may go into saturation. We will find that the saturation point shifts to lower SNRs as the number of receive antennas increases. Therefore for SBCEs, the more receive antennas the better, and the better channel estimation quality will in turn yields lower symbol error rate. In comparison, purely training-based channel estimators (TBCE) cannot benefit from reception diversity. The MSE of TBCE is always fixed, as long as the number of transmit antennas and the channel memory length is fixed. It can be observed that the saturation point is considerably close to the point, where the MSE curves of the proposed SBCE approaches the high SNR approximation of CRLB. Furthermore, the gap becomes smaller for an increasing number of receive antennas.

6 Conclusions

In this paper, we analyzed the performance of semiblind channel estimation for frequency-selective MIMO systems. An algorithm alternating between joint Viterbi detection and least squares channel estimation was presented. We also presented asymptotic approximations of the CRLB for SBCE in the high and low SNR regime. It was shown that the proposed algorithm is asymptotically unbiased in the high SNR regime, whereas it suffers from biasing in the low SNR regime. For this reason the proposed algorithm may perform better within the low SNR range as provided by the CRLB, whereas in the high SNR regime the CRLB is an admissible lower bound of the MSE performance. We also examined the SNR threshold, where SBCEs theoretically and practically may approach the asymptotic CRLB at high SNR. As shown by simulations, the presented algorithm is rather close to the theoretical bounds.

APPENDIX

A Derivation of (30)

For the sake of readability, in the following we will drop the subscript of $\delta_{\min}$ and denote the quantity simply by δ. The mean of δ is given by

$$\bar{\delta} = \mathrm{E}\{\delta\} = \int_0^\infty \delta p(\delta) d\delta. \tag{36}$$

Hence, for calculating $\bar{\delta}$ the PDF $p(\delta)$ is required. We will first give an approximation for $p(\delta)$ in the case of flat Rayleigh-fading channels, and then extend the obtained expression to frequency-selective channels.

Recall the definition of the minimum squared distance:

$$\delta = \min_{i \neq j} \delta_{i,j} = \min_{i \neq j} \|\mathbf{H}(\mathbf{S}_i - \mathbf{S}_j)\|_{\mathrm{F}}^2. \tag{37}$$

Since δ is mostly achieved if $\mathbf{S}_i$ and $\mathbf{S}_j$ differ only in one symbol, in the following we neglect those $\delta_{i,j}$ differing in more than one symbols. As shown in Table I this assumption provides a rather good match of the real situation. The remaining distances can be split into N_T ensembles w.r.t. the transmit antennas.

The n-th ensemble comprises all squared distances, where the position of the differing symbol is associated with the n-th transmit antenna. Note that each ensemble has coinciding distances, e. g. for the n-the ensemble the squared distance is given by

$$\Delta_n = w \sum_m |h_{m,n}|^2, \tag{38}$$

where Δ_n $(1 \leq n \leq N_\mathrm{T})$ are independent with each other and identically distributed. Given i.i.d. channel gains with $h_{m,n} \sim \mathcal{CN}(0,1)$, Δ_n conforms to the χ^2 distribution with $2 \cdot N_\mathrm{R}$ degrees of freedom given by

$$p(\Delta_n) = \frac{\Delta_n^{N_\mathrm{R}-1} e^{-\Delta_n/w}}{w^{N_\mathrm{R}}(N_\mathrm{R}-1)!}. \tag{39}$$

Hence, δ is the minimum of all ensemble distances Δ_n and the PDF of δ is given by

$$\begin{aligned} p(\delta) &\approx \sum_{n=1}^{N_\mathrm{T}} p(\Delta_n = \delta) \prod_{i=1, i\neq n}^{N_\mathrm{T}} \mathrm{P}(\Delta_i \geq \delta) \qquad (40) \\ &= N_\mathrm{T} \cdot p(\Delta_n = \delta) \cdot [\mathrm{P}(\Delta_i \geq \delta)]^{N_\mathrm{T}-1} \\ &= \frac{N_\mathrm{T}\delta^{N_\mathrm{R}-1} e^{-\frac{\delta}{w}}}{w^{N_\mathrm{R}}(N_\mathrm{R}-1)!} \left(\frac{\Gamma(N_\mathrm{R}, \frac{\delta}{w})}{(N_\mathrm{R}-1)!} \right)^{N_\mathrm{T}-1}, \end{aligned}$$

where the probability $\mathrm{P}(\Delta_i \geq \delta)$ is the complementary cumulative distribution function (CCDF) of Δ_i. $\Gamma(a,b)$ is the incomplete Gamma function defined as

$$\Gamma(a,b) = \int_b^\infty x^{a-1} e^{-x} dx. \tag{41}$$

Substituting (40) into (36), we finally obtain

$$\bar{\delta} = w N_\mathrm{T} \cdot \int_0^\infty \frac{\delta^{N_\mathrm{R}} e^{-\delta} [\Gamma(N_\mathrm{R}, \delta)]^{N_\mathrm{T}-1}}{[(N_\mathrm{R}-1)!]^{N_\mathrm{T}}} d\delta. \tag{42}$$

This result can be easily extended to the frequency-selective case. Taking into account the by factor $L+1$ increased degree of freedom in (39), the mean of δ is given by Eq. (30).

References

[1] A. Medles and DTM. Slock, "Semiblind channel estimation for MIMO spatial multiplexing systems," in *Proc. IEEE Veh. Techn. Conf. (VTC)*, Fall, 2001.

[2] M. Loncar *et al.*, "Iterative joint detection, decoding, and channel estimation for dual antenna arrays in frequency selective fading," in *Proc. 5th Int. Symp. Wireless Personal Multimedia Communication*, Honolulu, HI, Oct. 2002, pp. 125-129.

[3] C. Cozzo and B. L. Hughes, "Joint channel estimation and data detection in space-time communications," *IEEE Trans. Commun.*, vol. 51, pp. 1266-1270, Aug. 2003.

[4] T. Wo and P. A. Hoeher, "Semi-Blind Channel Estimation for Frequency-Selective MIMO Systems," in *Proc. 14th IST Mobile & Wireless Communications Summit*, Dresden, Germany, June, 2005.

[5] K. Giridhar, S. Chari, J. Shynk, R. Gooch, and D. Artman, "Joint estimation algorithms for cochannel signal demodulation," in *Proc. IEEE ICC' 93*, Geneva, pp. 1497-1501, May 1993.

[6] K. Giridhar, J. Shynk, A. Mathur, S. Chari, and R. Gooch, "Nonlinear techniques for the joint estimation of cochannel signals," *IEEE Trans. Comm.*, vol. 45, no. 4, pp. 473-484, Apr. 1997.

[7] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 384-391, Mar. 2003.

[8] S. Buzzi, M. Lops, and S. Sardellitti, "Performance of iterative data detection and channel estimation for single-antenna and multiple-antennas wireless communications," *IEEE Trans. Veh. Techn.*, vol. 53, no. 4, pp. 1085-1104, July 2004.

[9] A. Scherb, V. Kuehn and K. D. Kammeyer, "Cramer-Rao lower bound for semiblind channel estimation with respect to coded and uncoded finite-alphabet signals," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Monterey, USA, November, 2004

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, PTR Prentice-Hall, Englewood Cliffs, New Jersey, 1993.