

Semi-blind Channel Estimation Based on Superimposed Pilots for Single-Carrier MIMO Systems

Mohammad-Ali Khalighi and Salah Bourennane

Institut Fresnel, UMR CNRS 6133, École Centrale Marseille, Marseille, France

Email: {Ali.Khalighi, Salah.Bourennane}@fresnel.fr

Abstract— We propose a semi-blind channel estimator based on superimposed pilots for the case of single-carrier single-user MIMO systems. The proposed estimator is based on the expectation maximization (EM) algorithm. Numerical results show that even with this semi-blind estimator, the classical approach of time-multiplexed pilots still outperforms the superimposed approach, except for a large number of receive antennas and not too short frames.

I. INTRODUCTION

Superimposed or overlaid pilots are of special interest as they cause no loss in the data transmission rate [1], [2]. However, only under particular conditions the superimposed pilots (SP) technique could be preferred to the classical time-multiplexed pilots technique. The classical pilot insertion is usually referred to as pilot symbol assisted modulation (PSAM) in single-carrier systems and to pilot tone assisted modulation (PTAM) in multi-carrier systems. Although several works have recently shown the preference of the SP technique to PTAM in multi-carrier systems like orthogonal frequency division multiplexing (OFDM), in terms of channel capacity and performance [3], [4], [5], we can evoke a real *hesitation* for the case of single-carrier systems that we consider in this work. The main problem in this latter case is the non-zero cross-correlation of training and data symbols that results in an error floor at high signal-to-noise ratios (SNR) [6]. To reduce this problem, it is proposed in [7] to modify the pilot sequences according to the actual data sequences to be transmitted, so as to result in a low data-pilot cross-correlation. Such a solution may not be interesting in practice as the receiver has to estimate the *unknown* pilot sequence prior to channel estimation.

Here we consider the channel estimation in multiple-input multiple-output (MIMO) systems. It is shown in [8], [9] that the SP scheme can be preferred to PSAM for a larger number of receive than transmit antennas, and for relatively short channel coherence times. In [9], a simple iterative decision-directed estimator-detector is proposed, and it is shown that, when the power allocated to pilots is large enough, a considerable improvement in the performance can be obtained after few iterations. Yet, the resulting performance is not systematically better than that with PSAM. Notice that, although being a decision-directed approach, the channel estimation in [9] is

based only on pilots. It is recently proposed to use a semi-blind (also called data-aided) channel estimation and it is shown that the capacity of the optimized semi-blind SP scheme is larger than that of pilot-only-based or semi-blind PSAM [10]. Our aim in this work is to implement a semi-blind channel estimator, based on the expectation maximization (EM) algorithm, and to see whether or not by this semi-blind estimation the SP scheme can practically be preferable to PSAM.

The paper is organized as follows. Section II presents our system model. In Section III we consider the channel estimation problem and present briefly the decision-directed solution proposed in [9]. Next, we present our semi-blind estimator based on EM. Numerical results are presented in Section IV to compare the performance of different schemes.

II. SYSTEM MODEL

We consider single-carrier modulation in a single-user context. Rayleigh uncorrelated flat fading is considered for MIMO sub-channels. Also, the quasi-static model is considered for channel time variations. The presented results can easily be applied to the block-fading model as well. Gray-mapped QAM modulation is used and channel coding is performed using a non-recursive non-systematic convolutional (NRNSC) code. We denote by M_T and M_R the number of antennas at transmitter and at receiver, respectively. The block diagram of the transmitter is shown in Fig.1, where random interleaving and the simple spatial multiplexing (known also as V-BLAST [11]) scheme are used. Pilot symbols are added to data symbols prior to transmission. Pilot sequences for M_T antennas are QPSK modulated and are chosen according to the Walsh-Hadamard series that ensures their orthogonality.

A. Signal detection at receiver

At the receiver, iterative soft MIMO detection and soft channel decoding are performed together with channel estimation, as shown in Fig.2. MIMO soft detection is based either on the maximum a posteriori (MAP) criterion or on parallel interference cancellation (PIC). For the latter case, we use a simplified version of PIC, described in [12]. Soft-PIC detector general block diagram is shown in Fig.3. Soft-input

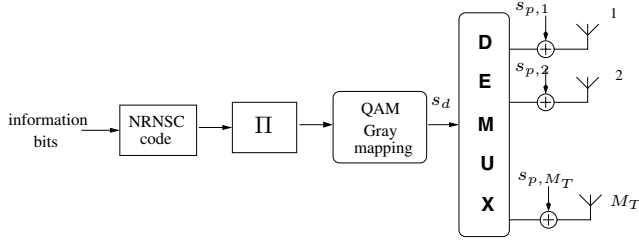


Fig. 1. Block diagram of the transmitter

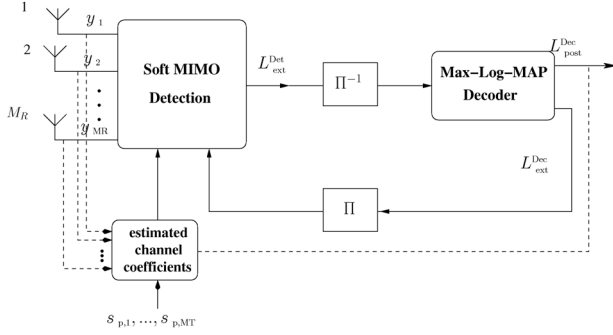


Fig. 2. Block diagram of the receiver

soft-output (SISO) channel decoding is based on the Max-Log-MAP algorithm. Details on the detectors can be found in [12], [13].

III. ITERATIVE CHANNEL ESTIMATION

Let us denote by \mathbf{s} the vector of transmitted symbols (of size M_T) at a given time sample, and by \mathbf{H} the MIMO channel matrix of dimension $(M_R \times M_T)$. The vector of received symbols on the receiver array, \mathbf{y} , is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \quad (1)$$

where \mathbf{n} is the vector of AWGN of zero mean and variance σ_n^2 , assumed to be known at receiver. For the SP scheme, $\mathbf{s} = \mathbf{s}_d + \mathbf{s}_p$, where \mathbf{s}_d and \mathbf{s}_p denote respectively data and pilot symbols with the average power of σ_d^2 and σ_p^2 .

To estimate the channel coefficient H_{ij} , we calculate the cross-correlation between the sequences \mathbf{y}_i and $\mathbf{s}_{p,j}$ whose lengths equal the frame length N_s in our quasi-static channel model:

$$\Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i) = \frac{1}{N_s} \sum_{n=1}^{N_s} s_{p,j}^*[n] \cdot y_i[n]. \quad (2)$$

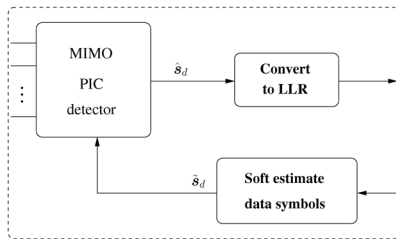


Fig. 3. Soft-PIC detector

Here n is the time reference in the frame and \cdot^* denotes complex conjugate. We have:

$$y_i[n] = \sum_{t=1}^{M_T} (s_{p,t}[n] + s_{d,t}[n])H_{it} + \mathbf{n}_i[n]. \quad (3)$$

Assuming mutually orthogonal pilot sequences, we obtain:

$$\begin{aligned} \Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i) &= \sigma_p^2 H_{ij} + \frac{1}{N_s} \sum_{n=1}^{N_s} \sum_{t=1}^{M_T} s_{d,t}[n] s_{p,j}^*[n] H_{it} \\ &+ \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{n}_i[n] s_{p,j}^*[n]. \end{aligned} \quad (4)$$

At the first iteration, we use $\Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i)/\sigma_p^2$ as the estimate of H_{ij} . Notice that there are two interference terms on this estimate $\hat{H}_{ij}^{(1)}$: one term from the non-zero cross-correlation of data and pilot sequences, and the other from the cross-correlation of noise samples with pilots. The more important one is obviously the former. Specially, at high SNR, this interference term becomes important and results in an error floor in the detector performance. To reduce this interference and to improve the estimation quality, we consider two solutions. The first one is the simple decision-directed (DD) scheme, already proposed in [9], and the second one is a semi-blind (SB) approach that we study in this work.

A. Decision-Directed approach

From the second iteration, we calculate soft-estimates of the transmitted data symbols \tilde{s}_d by using a posteriori LLRs at the SISO decoder output and cancel their effect in $\Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i)$. At iteration $m > 1$, we first calculate the data-removed vectors $\tilde{\mathbf{y}}_i^{(m)}$, and then the new estimate of H_{ij} , as shown below.

$$\tilde{y}_i^{(m)}[n] = y_i[n] - \sum_{t=1}^{M_T} \tilde{s}_{d,t}[n] \hat{H}_{it}^{(m-1)} \quad (5)$$

$$\hat{H}_{ij}^{(m)} = \frac{1}{\sigma_p^2} \Gamma(\mathbf{s}_{p,j}, \tilde{\mathbf{y}}_i^{(m)}) = \frac{1}{N_s \sigma_p^2} \sum_{n=1}^{N_s} s_{p,j}^*[n] \cdot \tilde{y}_i^{(m)}[n] \quad (6)$$

B. Semi-blind approach

The semi-blind estimator that we propose is based on the EM algorithm. Let Θ be the vector of the parameters to be estimated (containing the entries of \mathbf{H} , stacked). For a given frame of symbols, let us stack the entries of $\mathbf{y}[n]$, $n = 1, \dots, N_s$ in a vector \mathbf{Y} . We constitute the vector \mathbf{S}_d in the same way. The maximum likelihood estimate of Θ , given the vector of observations \mathbf{Y} , is:

$$\hat{\Theta} = \arg \max_{\Theta} \log p(\mathbf{Y}|\Theta) \quad (7)$$

Here $p(\cdot|\cdot)$ denotes the conditional pdf. The EM algorithm consists of two update steps; the expectation (E) and the maximization (M) step. These are described below for the m^{th} update.

$$Q(\Theta, \hat{\Theta}^{(m-1)}) = \mathbb{E} \left\{ \log p(\mathbf{Y}, \mathbf{S}_d|\Theta) \mid \mathbf{Y}, \hat{\Theta}^{(m-1)} \right\} \quad (8)$$

$$\hat{\Theta}^{(m)} = \arg \max_{\Theta} Q(\Theta, \hat{\Theta}^{(m-1)}) \quad (9)$$

Here $E\{\cdot\}$ denotes expectation and is calculated on data symbols. We define \mathcal{S} as the set of constellation points corresponding to data symbol sequences with the cardinality $|\mathcal{S}| = 2^{BM_T N_s}$, B being the number of bits per symbol. The auxiliary function Q can be described as follows:

$$Q(\Theta|\hat{\Theta}^{(m-1)}) = \sum_{\mathbf{S}_{d_u} \in \mathcal{S}} \log p(\mathbf{Y}|\mathbf{S}_d = \mathbf{S}_{d_u}, \Theta) p(\mathbf{S}_{d_u}|\mathbf{Y}, \hat{\Theta}^{(m-1)}) \quad (10)$$

where \mathbf{S}_{d_u} corresponds to a possibly-transmitted data sequence. Based on the assumptions of flat channel and independent noise samples, conditioned to channel state, $y[n]$ and $y[n+1]$ are independent. We can hence write:

$$Q(\Theta|\hat{\Theta}) = \sum_{n=1}^{N_s} \sum_{u=1}^{2^{BM_T}} \left(cte - \frac{\|\mathbf{y}[n] - \mathbf{H}(\mathbf{s}_{d_u} + \mathbf{s}_p[n])\|^2}{\sigma_n^2} \right) APP_n(\mathbf{s}_{d_u}, \hat{\Theta}) \quad (11)$$

where cte is a constant term and \mathbf{s}_{d_u} is a possibly-transmitted data symbol vector. To simplify our notations, we did not specify the iteration number. Also, $APP_n(\mathbf{s}_{d_u}, \hat{\Theta})$ is the probability of the transmission of \mathbf{s}_{d_u} , i.e., the probability of $\mathbf{s}_{d_u} = \mathbf{s}_d[n]$, and is calculated using the a posteriori probabilities (APP) on its constituting bits at the channel decoder output. For the sake of simplicity we denote it hereafter by $APP_{n,u}$. Also, we denote the summations in (11) simply by their indices, i.e., n and u . Differentiating $Q(\Theta|\hat{\Theta})$ with respect to \mathbf{H} and setting it to zero to find its local maximum, we obtain:

$$\sum_n \sum_u (\mathbf{s}_{d_u} + \mathbf{s}_p[n]) [\mathbf{y}^\dagger[n] - (\mathbf{s}_{d_u}^\dagger + \mathbf{s}_p^\dagger[n]) \mathbf{H}^\dagger] APP_{n,u} = 0 \quad (12)$$

where \dagger stands for complex conjugate transpose. After some simple manipulations, we obtain:

$$\hat{\mathbf{H}} = \mathbf{R}_{YS} \mathbf{R}_{SS}^{-1} \quad (13)$$

where,

$$\mathbf{R}_{YS} = \sum_n \sum_u \mathbf{y}[n] (\mathbf{s}_{d_u}^\dagger + \mathbf{s}_p^\dagger[n]) APP_{n,u} \quad (14)$$

$$\mathbf{R}_{SS} = \sum_n \sum_u (\mathbf{s}_{d_u} + \mathbf{s}_p[n]) (\mathbf{s}_{d_u}^\dagger + \mathbf{s}_p^\dagger[n]) APP_{n,u} \quad (15)$$

The evaluation of \mathbf{R}_{YS} and \mathbf{R}_{SS} as presented above is computationally complex. We propose to reformulate these matrices by considering soft-estimates of the data transmitted symbols which can be calculated using the APPs at the channel decoder output. Actually, these soft-estimates, that we denote by \tilde{s}_d , are already available when Soft-PIC detection is performed. For a vector of M_T data symbols, we have:

$$\tilde{s}_d[n] = \sum_u \mathbf{s}_{d_u} APP_{n,u} \quad (16)$$

Now, using these soft-estimates, it can be shown that:

$$\mathbf{R}_{YS} = \sum_n \mathbf{y}[n] (\tilde{\mathbf{s}}_d^\dagger[n] + \mathbf{s}_p^\dagger[n]) \quad (17)$$

$$\mathbf{R}_{SS} = \bar{\mathbf{R}}_S + \sum_n \left[\tilde{s}_d[n] \mathbf{s}_p^\dagger[n] + \mathbf{s}_p[n] \tilde{s}_d^\dagger[n] \right] + N_s \sigma_p^2 \mathbf{I}_{M_T} \quad (18)$$

where \mathbf{I}_{M_T} is the Identity matrix of dimension $(M_T \times M_T)$ and

$$\bar{\mathbf{R}}_S = \sum_n \sum_u \mathbf{s}_{d_u} \mathbf{s}_{d_u}^\dagger APP_{n,u} \quad (19)$$

The $(i, j)^{th}$ entry of $\bar{\mathbf{R}}_S$ can be calculated as follows.

$$\bar{R}_{S,ij} = \begin{cases} N_s \sigma_d^2 & ; i = j \\ \sum_n \tilde{s}_{d_i}[n] \tilde{s}_{d_j}^*[n] & ; i \neq j \end{cases} \quad (20)$$

1) *Initializing EM*: To permit the EM algorithm to bootstrap, we should provide it with a primary estimate, i.e., we have to initialize it. This primary estimate can be obtained via ML or least-squares (LS) estimation. For the former, we obtain the following expression for $\hat{\mathbf{H}}^{(1)}$, the estimate at the first iteration.

$$\hat{\mathbf{H}}_{ML}^{(1)} = \frac{1}{N_s(\sigma_d^2 + \sigma_p^2)} \sum_n \mathbf{y}[n] \mathbf{s}_p^\dagger[n] \quad (21)$$

We do not provide here the details on the ML estimator expression because of lack of space; it may otherwise be obtained via the expressions provided for EM estimation by setting $\tilde{s}_d = 0$. The LS estimate, on the other hand, comes back to the expression (2) at the beginning of this section:

$$\hat{\mathbf{H}}_{LS}^{(1)} = \frac{1}{N_s \sigma_p^2} \sum_n \mathbf{y}[n] \mathbf{s}_p^\dagger[n]. \quad (22)$$

We verified that the LS estimate provides a better convergence of the EM algorithm than the ML estimate. A similar statement is reported in [14]. Therefore, in what follows, we use the LS estimate for initializing EM.

IV. SIMULATION RESULTS

Unless otherwise mentioned, two transmit and two receive antenna MIMO system is considered; performance curves correspond to the fifth iteration of the receiver where almost full convergence is attained, and to a frame of $N_s = 100$ channel-uses. Rayleigh flat fading is considered and channel coefficients are assumed to be uncorrelated and normalized. QPSK modulation is used, the channel code is the NRNSC code $(133, 171)_8$, and the interleaver is random. As in the SP scheme a part of the transmit power is dedicated to pilots, we prefer to present the performance curves in terms of the actual average SNR, i.e., $M_R(\sigma_d^2 + \sigma_p^2)/\sigma_n^2$, instead of E_b/N_0 as done in [9]. In the results to be presented, SNR stands for this actual value and includes the antenna array gain at receiver, i.e., M_R . We denote by α the percentage of the power dedicated to pilots, i.e., $\alpha = \sigma_p^2/(\sigma_p^2 + \sigma_d^2)$. In this way, we can directly see the compromise between the channel estimation quality and the data detection performance, e.g. by increasing α .

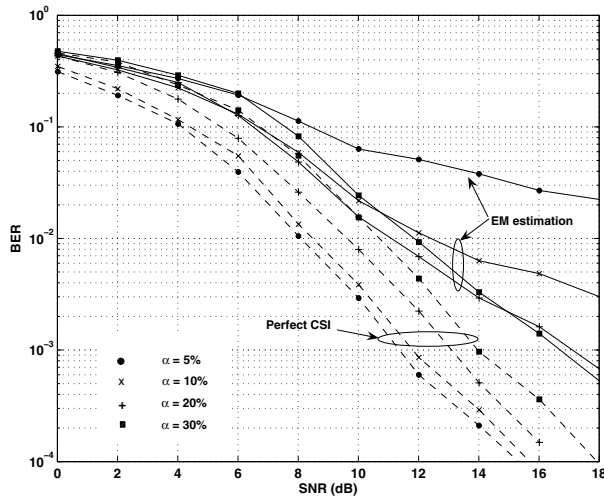


Fig. 4. PIC detection, SB versus perfect estimation, $M_T = M_R = 2$, $N_s = 100$.

A. Performance with semi-blind estimation

Fig.4 compares bit-error-rate (BER) curves versus SNR for the cases of perfect channel knowledge and semi-blind estimation based on EM. PIC detection is performed and several α values are considered. It is seen that the problem of error floor is still present for $\alpha = 5\%$ and 10% . In fact, by increasing SNR, the interference terms in channel estimation become more important and result in an error floor. This is especially the case for small values of α for which the channel estimate is not good enough to permit good soft-estimates of data symbols. Results are much better for increased frame length (not presented here). Also, notice that by increasing α , better channel estimates are obtained, but at the same time, less power is dedicated to data symbols, and hence, increasing α too high will result in an overall performance degradation.

B. Decision-directed versus semi-blind estimation

BER curves versus SNR are shown in Fig.5 for DD and SB estimation methods and PIC detection. The performance improvement by SB estimation is considerable for small α values but, it is less significant for α as large as 20% . We conclude that the already-presented DD estimation approach makes a good trade-off between complexity and performance.

C. MAP versus PIC detection

It is interesting to compare the performance of the receiver for two cases of MAP and PIC detection. The corresponding BER curves are contrasted in Fig.6 for several values of α . We see that for relatively small (and *reasonable*) values of α , the PIC detector (with much less computational complexity) outperforms MAP. Only for $\alpha = 30\%$ the MAP detector shows a better performance. As a matter of fact, for small α , the channel estimates are not precise enough and result in a degradation of the detector performance, as compared to

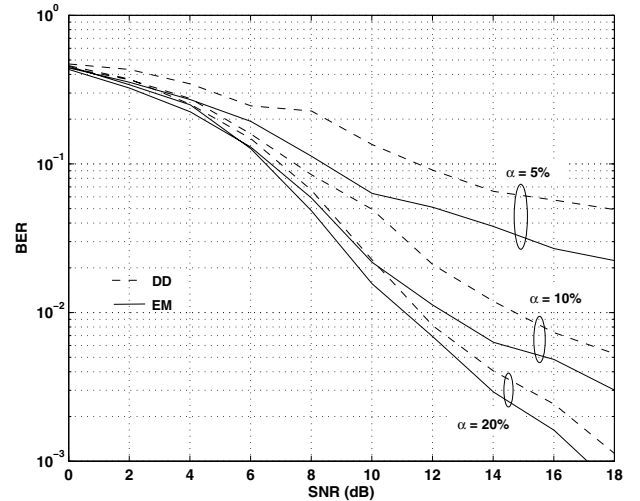


Fig. 5. PIC detection, SB estimation versus DD, $M_T = M_R = 2$, $N_s = 100$.

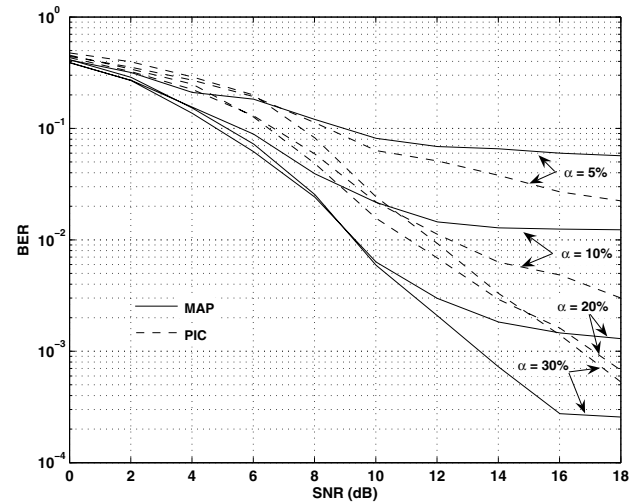


Fig. 6. MAP versus PIC detection, SB estimation, $M_T = M_R = 2$, $N_s = 100$.

the perfect channel knowledge case. This degradation is more considerable for the MAP detector than PIC. In other words, the MAP detector is more sensitive to channel estimation errors. Similar results are obtained for the case of DD estimation (results are not presented).

D. Comparison with PSAM

Here we consider the crucial question: which scheme should be preferred, SP or PSAM? For this purpose, we contrast BER curves for the two methods for the case of iterative PIC detection. For the PSAM method, we perform the simple pilot-only-based estimation that is done once at the first iteration. For the SP method, the receiver corresponds to what considered previously, i.e., to Fig.2. For PSAM, we use N_p

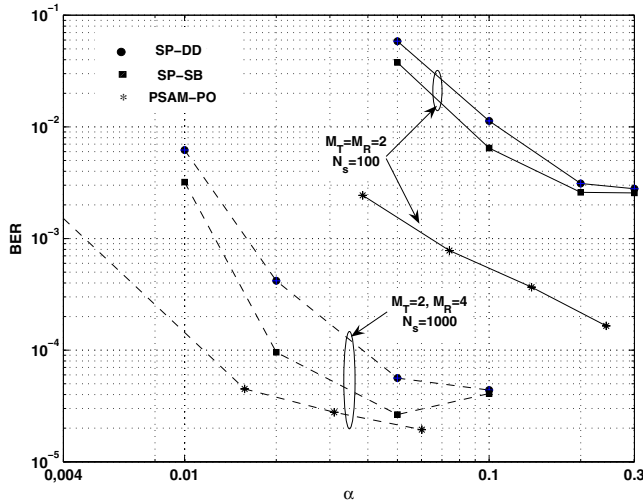


Fig. 7. SP versus PSAM, SNR = 15 dB for $M_R = 2$ and $N_s = 100$, SNR = 10 dB for $M_R = 4$ and $N_s = 1000$

channel-uses for the transmission of mutually orthogonal pilot sequences, to which we attribute the same power as that to data symbols. For a nominal SNR of 15 dB we obtained the BER for several values of α that are shown in Fig.7. In order that the comparison be fair, we modify the noise power in the PSAM case to take into account the loss in the spectral efficiency. The actual E_b/N_0 is adjusted for both cases to result in the same received SNR. The third iteration of the detector is considered; for the SP case with SB estimation the small performance gain may not justify more iterations. Results show that for $N_s = 100$, the PSAM scheme is highly preferable for small α . For large α , the loss in the spectral efficiency may reduce the interest of PSAM. We have also shown the corresponding curves for $N_s = 1000$ and $M_R = 4$ and SNR=10 dB. For this case, the SP scheme seems to be more interesting, taking into account the loss in the spectral efficiency for the PSAM scheme.

V. CONCLUSIONS

We considered MIMO channel estimation using superimposed pilots in a single-carrier single-user context. We proposed a semi-blind estimation method based on the EM algorithm. We showed that a considerable gain can be obtained when low power is allocated to pilots, i.e., for small α values, as compared to an already-proposed decision-directed (DD) method. For relatively large α , however, the DD method seems to be an appropriate solution considering complexity and performance. Comparing to PSAM scheme, however, we showed that except for not too short frames and for more receive than transmit antennas, the PSAM scheme may be preferable to SP even with the proposed SB estimation.

REFERENCES

[1] H. Zhu, B. Farhang-Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication

systems," *IEEE Communications Letters*, vol. 7, no. 1, pp. 30–32, Jan. 2003.

[2] P. Hoher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," *IEEE Global Telecommunications Conference (GlobeCom)*, vol. 4, pp. 2162–2166, 1999, Rio de Janeiro, Brazil.

[3] S. Balasubramanian, B. Farhang-Boroujeny, and V. J. Mathews, "Pilot embedding for channel estimation and tracking in OFDM systems," *IEEE Global Telecommunications Conference (GlobeCom)*, pp. 1244–1248, Nov.-Dec. 2004, Dallas, TX.

[4] A. R. Varma, L. L. H. Andrew, C. R. N. Athaudage, and J. H. Manton, "Iterative algorithms for channel identification using superimposed pilots," *Australian Communications Theory Workshop (AusCTW)*, pp. 195–201, Feb. 2005, Brisbane, Australia.

[5] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami, "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 226–229, Mar. 2005.

[6] V. Jungnickel, T. Haustein, E. Jorswieck, V. Pohl, and C. von Helmolt, "Performance of a MIMO system with overlay pilots," *IEEE Global Telecommunications Conference (GlobeCom)*, vol. 1, pp. 594–598, Nov. 2001, San Antonio, TX.

[7] F. Tsuzuki and T. Ohtsuki, "Channel estimation with selective superimposed pilot sequences under fast fading environments," *IEEE Vehicular Technology Conference*, pp. 62–66, Sept. 2004, Los Angeles, CA.

[8] P. Bohlín and M. Tapio, "Performance evaluation of MIMO communication systems based on superimposed pilots," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. IV, pp. 425–428, May 2004, Montreal, Canada.

[9] M. A. Khalighi, L. Berriche, and J.-F. Hélar, "Overlaid or time-multiplexed pilots for channel estimation in iterative MIMO receivers," *International Symposium on Signal Processing and its Applications (ISSPA)*, pp. 483–486, Aug. 2005, Sydney, Australia.

[10] P. Bohlín and M. Tapio, "Optimized data aided training in MIMO systems," *Vehicular Technology Conference (VTC)*, pp. 679–683, May 2004, Milan, Italy.

[11] G. Golden, G. Foschini, R. Valenzuela, and P. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[12] M. A. Khalighi, J.-F. Hélar, and S. Bourennane, "Choice of appropriate space-time coding scheme for MIMO systems employing channel coding under BICM," *International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, July 2006, Cannes, France.

[13] M. A. Khalighi and J. J. Boutros, "Semi-blind channel estimation using EM algorithm in iterative MIMO APP detectors," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3165–3173, Nov. 2006.

[14] T. Cui and C. Tellambura, "Superimposed pilot symbols for channel estimation in OFDM systems," *IEEE Global Telecommunications Conference (GlobeCom)*, pp. 2229–2233, Nov.-Dec. 2005, St. Louis, MO.