Semi-blind Joint Maximum Likelihood Channel Estimation and Data Detection for MIMO Systems

M. Abuthinien, S. Chen, and L. Hanzo

***Abstract*—Semi-blind joint maximum likelihood (ML) channel estimation and data detection is proposed for multiple-input multiple-output (MIMO) systems. The joint ML optimization over channel and data is decomposed into an iterative two-level optimization loop. An efficient optimization search algorithm referred to as the repeated weighted boosting search (RWBS) is employed at the upper level to identify the unknown MIMO channel while an enhanced ML sphere detector termed as the optimized hierarchy reduced search algorithm is used at the lower level to perform ML detection of the transmitted data. Only a minimum pilot overhead is required to aid the RWBS channel estimator's initial operation, which not only speeds up convergence but also avoids ambiguities inherent in blind joint estimation of both the channel and data.**

***Index Terms*—Channel estimation, data detection, joint maximum likelihood estimation, multiple-input multiple-output.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) technologies are capable of substantially improving the achievable system's capacity and/or quality of service [1]–[4]. The system's ability to approach the MIMO capacity heavily relies on the channel state information. Accurately estimating a MIMO channel is much more challenging than its single-input single-output (SISO) counterpart. The various MIMO channel estimation methods can be classified into three categories: training-based methods, blind methods and semi-blind methods. Pure training-based schemes are computationally less demanding but a high proportion of training symbols is required in order to obtain a reliable MIMO channel estimate, which considerably reduces the achievable system throughput. The family of blind methods for joint channel estimation and data detection does not require training symbols and hence does not reduce the achievable system throughput, although this is achieved at the expense of high computational complexity. Moreover, blind joint channel estimation and data detection results in a certain grade of estimation and decision ambiguities [5]. Semi-blind schemes do not suffer from this ambiguity problem and are computationally simpler than their blind counterparts, at the cost of requiring a few training symbols.

Many semi-blind methods have been developed for MIMO systems. In the schemes of [6]–[9], a few training symbols are used to provide an initial MIMO channel estimate, and the channel estimator as well as the data detector iteratively exchange their information, where the channel estimator relies on decision-directed adaptation. In [10], the MIMO channel matrix is decomposed into the product of a whitening matrix and a rotational unitary matrix. The first matrix is estimated blindly while the second is estimated with the aid of training symbols. In contrast to these proposals, our novel contribution extends the approach developed for SISO systems in [11] and that advocated for single-input multiple-output systems in [12], where the joint ML channel and data estimation optimization process is decomposed into two levels. At the upper level a global optimization algorithm searches for an optimal channel estimate, while at the lower level a ML data detector recovers the transmitted data. Joint ML channel estimation and data detection is achieved by iteratively exchanging information between the channel estimator and the data detector. Specifically, at the upper level we use the repeated weighted boosting search (RWBS) algorithm [13] as the channel estimator, which searches the MIMO channel space by evolving a population of MIMO channel matrices, while at the lower level we use the optimized hierarchy reduced search algorithm (OHRSA) aided detector [14], which is an advanced extension of the complex sphere decoder [15], to provide ML data estimates for the MIMO channel population. Only a few training symbols are used to provide an initial least squares channel estimate (LSCE) [16] for aiding the RWBS channel estimator to improve its convergence. The employment of a minimum training overhead has an additional benefit in terms of avoiding the ambiguities inherent in pure blind joint channel estimation and data detection.

Throughout our discussions we adopt the following notational conventions. Boldface capitals and lower-case letters stand for matrices and vectors, respectively, while $\mathbf{I}_K$ and $\mathbf{1}_{K\times L}$ denote the $K\times K$ identity matrix and the $K\times L$ matrix of unity elements, respectively. Furthermore, $(\ )^T$ and $(\ )^H$ are the transpose and Hermitian operators, respectively, while $\|\ \|^2$ and $|\ |$ denote the norm and the magnitude operators, respectively. Finally, $E[\ ]$ is the expectation operator.

II. System Model

We consider a MIMO system consisting of n_T transmitters and n_R receivers, which communicates over flat fading channels. The system is described by the MIMO model

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

where k is the symbol index, $\mathbf{H}$ denotes the $n_R\times n_T$ MIMO channel matrix, $\mathbf{s}(k)=[s_1(k)\quad s_2(k)\ldots s_{n_T}(k)]^T$ is the

Manuscript received June 18, 2007; revised October 5, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gerald Matz.

The authors are with School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: msaa05r@ecs.soton.ac.uk; sqc@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/LSP.2007.911758

transmitted symbols vector of the n_T transmitters with the symbol energy given by $E\left[|s_m(k)|^2\right]=\sigma_s^2$ for $1\leq m\leq n_T$, $\mathbf{y}(k) = [y_1(k) \quad y_2(k)\ldots y_{n_R}(k)]^T$ denotes the received signal vector, and $\mathbf{n}(k)=[n_1(k) \quad n_2(k)\ldots n_{n_R}(k)]^T$ is the complex-valued Gaussian white noise vector associated with the MIMO channels with $E\left[\mathbf{n}(k)\mathbf{n}^H(k)\right]=2\sigma_n^2\mathbf{I}_{n_R}$.

Specifically, the narrowband MIMO channel matrix is defined by $\mathbf{H}=[h_{l,m}]$, for $1\leq l\leq n_R$ and $1\leq m\leq n_T$, where $h_{l,m}$ denotes the nondispersive channel coefficient linking the mth transmitter to the l-th receiver. Furthermore, the fading is assumed to be sufficiently slow, so that during the time period of a short block of N symbols, all the entries in the MIMO channel matrix $\mathbf{H}$ may be deemed unchanged. From frame to frame, the channel impulse response taps $h_{l,m}$ are independently and identically distributed complex-valued Gaussian processes with zero mean and $E\left[|h_{l,m}|^2\right]=1$. The signal-to-noise ratio (SNR) is defined by $E_b/N_o=\sigma_s^2/2\sigma_n^2$.

III. Proposed Semi-Blind Scheme

Let us consider joint channel estimation and data detection based on the observation vector $\mathbf{y}(k)$ over a relatively short length of N symbols. Define the $n_R\times N$ matrix of received data as $\mathbf{Y}=[\mathbf{y}(1) \quad \mathbf{y}(2)\ldots\mathbf{y}(N)]$ and the corresponding $n_T\times N$ matrix of transmitted symbols as $\mathbf{S}=[\mathbf{s}(1) \quad \mathbf{s}(2)\ldots\mathbf{s}(N)]$. Then the probability density function of the received data matrix $\mathbf{Y}$ conditioned on the MIMO channel matrix $\mathbf{H}$ and the transmitted symbol matrix $\mathbf{S}$ can be written as

$$p(\mathbf{Y}|\mathbf{H},\mathbf{S})=\frac{1}{(2\pi\sigma_n^2)^{n_R\times N}}e^{-(1/2\sigma_n^2)\sum_{k=1}^{N}\|\mathbf{y}(k)-\mathbf{H}\ \mathbf{s}(k)\|^2}. \tag{2}$$

The ML estimation of $\mathbf{S}$ and $\mathbf{H}$ can be obtained by jointly maximizing $p(\mathbf{Y}|\mathbf{H},\mathbf{S})$ over $\mathbf{S}$ and $\mathbf{H}$. Equivalently, the joint ML estimation can be obtained by minimzing the cost function

$$J_{\mathrm{ML}}(\check{\mathbf{S}},\check{\mathbf{H}})=\frac{1}{n_R\times N}\sum_{k=1}^{N}\left\|\mathbf{y}(k)-\check{\mathbf{H}}\ \check{\mathbf{s}}(k)\right\|^2. \tag{3}$$

Thus, the joint ML channel and data estimation is obtained as

$$(\hat{\mathbf{S}},\hat{\mathbf{H}})=\arg\left\{\min_{\check{\mathbf{S}},\check{\mathbf{H}}}J_{\mathrm{ML}}(\check{\mathbf{S}},\check{\mathbf{H}})\right\}. \tag{4}$$

The joint ML optimization search defined in (4) is computationally prohibitive. The complexity of this optimization process may be reduced to a tractable level, if it is decomposed into an iterative search carried out over all the possible data symbols first and then over the channel matrices as

$$(\hat{\mathbf{S}},\hat{\mathbf{H}})=\arg\left\{\min_{\check{\mathbf{H}}}\left[\min_{\check{\mathbf{S}}}J_{\mathrm{ML}}(\check{\mathbf{S}},\check{\mathbf{H}})\right]\right\}. \tag{5}$$

At the inner or lower-level optimization we use the OHRSA-aided ML detector [14] to find the ML data estimate for the given channel. The detailed implementation of the OHRSA-aided ML detector can be found in [14] and will not be repeated here. In order to guarantee a joint ML estimate, the search algorithm used at the outer or upper-level optimization should be capable of finding a global optimal channel estimate efficiently, and we employ the RWBS algorithm [13] to perform the upper-level optimization. Motivations and analysis of the RWBS algorithm as a global search algorithm are detailed in [13] and will not be repeated here. Conceptually, a joint ML channel estimation and data detection can be carried out using the following iterative loop.

Outer-level Optimization: The RWBS algorithm searches the MIMO channel parameter space via evolving a population of channel matrices to find a global optimal estimate $\hat{\mathbf{H}}$ by minimising the mean square error (mse)

$$J_{\mathrm{mse}}(\check{\mathbf{H}})=J_{\mathrm{ML}}(\hat{\mathbf{S}}(\check{\mathbf{H}}),\check{\mathbf{H}}) \tag{6}$$

where $\hat{\mathbf{S}}(\check{\mathbf{H}})$ denotes the ML estimate of the transmitted data for the given channel $\check{\mathbf{H}}$.

Inner-level Optimization: Given the MIMO channel matrix $\check{\mathbf{H}}$ the OHRSA-aided ML detector finds the ML estimate of the transmitted data and feeds back the corresponding ML metric $J_{\mathrm{mse}}(\check{\mathbf{H}})$ to the upper level.

Pure blind joint data and channel estimation for MIMO systems has an inherent permutation and scaling ambiguity problem. Scaling ambiguity refers to the fact that the detected data and the estimated channel matrix columns can only be resolved within a complex-valued factor. In the permutation ambiguity, the detected data and the estimated channel matrix columns are reordered. The reason for this is clear from the cost function defined in (3), which is invariant with respect to a re-ordering and scaling of the channel matrix and the data matrix. In fact, let a pair of the MIMO channel and data estimates be $\hat{\mathbf{S}}$ and $\hat{\mathbf{H}}$. Define $\hat{\mathbf{H}}^*=\hat{\mathbf{H}}\ \mathbf{T}$ and $\hat{\mathbf{S}}^*=\mathbf{T}^H\hat{\mathbf{S}}$, where $\mathbf{T}$ is any $n_T\times n_T$ unitary matrix [5]. Then $J_{\mathrm{ML}}(\hat{\mathbf{H}},\hat{\mathbf{S}})=J_{\mathrm{ML}}(\hat{\mathbf{H}}^*,\hat{\mathbf{S}}^*)$. To resolve the permutation and scaling ambiguities, a few training symbols can be employed to identify the correct unitary matrix $\mathbf{T}$. Let the number of training symbols be t, and denote the available training data as $\mathbf{Y}_t=[\mathbf{y}(1) \quad \mathbf{y}(2)\ldots\mathbf{y}(t)]$ and $\mathbf{S}_t=[\mathbf{s}(1) \quad \mathbf{s}(2)\ldots\mathbf{s}(t)]$. The correct unitary matrix $\mathbf{T}$ is determined from all the possible realizations $\check{\mathbf{T}}$ by solving the following optimization:

$$\mathbf{T}=\arg\min_{\check{\mathbf{T}}}\left\{\left\|\mathbf{Y}_t-\hat{\mathbf{H}}\ \check{\mathbf{T}}\,\mathbf{S}_t\right\|^2\right\}. \tag{7}$$

The training data $\mathbf{S}_t$ and $\mathbf{Y}_t$ can also be utilised to provide an initial LSCE, defined by

$$\check{\mathbf{H}}_{\mathrm{LSCE}}=\mathbf{Y}_t\mathbf{S}_t^H\left(\mathbf{S}_t\mathbf{S}_t^H\right)^{-1} \tag{8}$$

for aiding the RWBS channel estimator. This not only improves the achievable convergence speed but also helps to resolve the above-mentioned ambiguities associated with pure blind schemes. Therefore, it is no longer required to solve the optimization problem (7). We can now summarise the proposed semi-blind joint ML estimation scheme based on the RWBS channel estimator and the OHRSA data detector.

Specify the three RWBS algorithmic parameters, namely the population size P_S, the number of generations N_G and the number of iterations in weighted boosting search N_I, as well as the control parameter γ for channel population initialization.

○ **Algorithm initialization:** $\check{\mathbf{H}}_{\mathrm{best}}^{(0)}=\check{\mathbf{H}}_{\mathrm{LSCE}}$

○ **Generation loop:** for $(g=1; g\leq N_G; g++)$ {

Generation initialization: $\check{\mathbf{H}}_1^{(g)} = \check{\mathbf{H}}_{\text{best}}^{(g-1)}$ and

$$\check{\mathbf{H}}_i^{(g)} = \check{\mathbf{H}}_1^{(g)} + \Delta(\mathbf{1}_{n_R \times n_T} + j\mathbf{1}_{n_R \times n_T}), \quad 2 \le i \le P_S$$

where Δ is a uniformly distributed random variable taking values from $[-\gamma, \gamma]$

Weighted boosting search initialization: Assign the initial distribution weightings $\delta_i(0) = 1/P_S$, $1 \le i \le P_S$, and compute the cost function of each channel

$$J_{\text{mse}}^{(g)}(i) = J_{\text{mse}}\left(\check{\mathbf{H}}_i^{(g)}\right) = J_{\text{ML}}\left(\hat{\mathbf{S}}\left(\check{\mathbf{H}}_i^{(g)}\right), \check{\mathbf{H}}_i^{(g)}\right)$$

where $\hat{\mathbf{S}}\left(\check{\mathbf{H}}_i^{(g)}\right)$ is the ML data estimate for channel $\check{\mathbf{H}}_i^{(g)}$, provided by the OHRSA-aided ML detector

○ **Weighted boosting search**: for $(l = 1; l \le N_I; l++)$ {

Step I. Boosting:

1) Find $i_{\text{best}} = \arg\min_{1 \le i \le P_S} J_{\text{mse}}^{(g)}(i)$ and $i_{\text{worst}} = \arg\max_{1 \le i \le P_S} J_{\text{mse}}^{(g)}(i)$. Denote $\check{\mathbf{H}}_{i_{\text{best}}}^{(g)} = \check{\mathbf{H}}_{\text{best}}^{(g)}$ and $\check{\mathbf{H}}_{i_{\text{worst}}}^{(g)} = \check{\mathbf{H}}_{\text{worst}}^{(g)}$.

2) Normalize the cost function values

$$\bar{J}_{\text{mse}}^{(g)}(i) = \frac{J_{\text{mse}}^{(g)}(i)}{\sum_{q=1}^{P_S} J_{\text{mse}}^{(g)}(q)}, \quad 1 \le i \le P_S.$$

3) Compute a weighting factor β_l according to

$$\eta_l = \sum_{i=1}^{P_S} \delta_i(l-1)\bar{J}_{\text{mse}}^{(g)}(i), \quad \beta_l = \frac{\eta_l}{1-\eta_l}.$$

4) Update the distribution weightings for $1 \le i \le P_S$

$$\tilde{\delta}_i(l) = \begin{cases} \delta_i(l-1)\beta_l^{\bar{J}_{\text{mse}}^{(g)}(i)}, & \text{for } \beta_l \le 1 \\ \delta_i(l-1)\beta_l^{1-\bar{J}_{\text{mse}}^{(g)}(i)}, & \text{for } \beta_l > 1 \end{cases}$$

and normalize them

$$\delta_i(l) = \frac{\tilde{\delta}_i(l)}{\sum_{q=1}^{P_S} \tilde{\delta}_q(l)}, \quad 1 \le i \le P_S.$$

Step II. Population updating:

1) Construct the $(P_S + 1)$th point using the formula

$$\check{\mathbf{H}}_{P_S+1} = \sum_{i=1}^{P_S} \delta_i(l)\check{\mathbf{H}}_i^{(g)}$$

and construct the $(P_S + 2)$th point using the formula

$$\check{\mathbf{H}}_{P_S+2} = \check{\mathbf{H}}_{\text{best}}^{(g)} + \left(\check{\mathbf{H}}_{\text{best}}^{(g)} - \check{\mathbf{H}}_{P_S+1}\right).$$

2) Compute the cost function values $J_{\text{mse}}(\check{\mathbf{H}}_{P_S+1})$ and $J_{\text{mse}}(\check{\mathbf{H}}_{P_S+2})$. Then find

$$i_* = \arg \min_{i=P_S+1, P_S+2} J_{\text{mse}}(\check{\mathbf{H}}_i).$$

The pair $\left(\check{\mathbf{H}}_{i_*}, J_{\text{mse}}(\check{\mathbf{H}}_{i_*})\right)$ then replaces $\left(\check{\mathbf{H}}_{\text{worst}}^{(g)}, J_{\text{mse}}(\check{\mathbf{H}}_{\text{worst}}^{(g)})\right)$ in the population

○} **End of weighted boosting search**

○ **Solution**: $\check{\mathbf{H}}_{\text{best}}^{(g)}$

○} **End of generation loop**

○ **Solution**: $\left(\hat{\mathbf{H}} = \check{\mathbf{H}}_{\text{best}}^{(N_G)}, \hat{\mathbf{S}} = \hat{\mathbf{S}}\left(\check{\mathbf{H}}_{\text{best}}^{(N_G)}\right)\right)$

Let $C_{\text{OHRSA}}(N)$ be the complexity of the OHRSA algorithm to decode the N-sample data matrix $\mathbf{S}$ and let N_{OHRSA} be the number of calls for the OHRSA algorithm required by the RWBS algorithm to converge. Then the complexity of the proposed semi-blind method is expressed as

$$C = N_{\text{OHRSA}} \times C_{\text{OHRSA}}(N). \tag{9}$$

The complexity of the OHRSA detector depends on the SNR [14]. The RWBS algorithm is a simple yet efficient global search algorithm. It can be shown that $N_{\text{OHRSA}} = N_G\left((P_S - 1) + 2N_I\right)$. Appropriate values for the RWBS algorithmic parameters, P_S, N_G and N_I, depends on how hard the objective function to be optimized. Generally, these algorithmic parameters have to be found empirically but some rules are discussed in [13]. The control parameter γ in the channel population initialization also influences the performance.

IV. Simulation Study

A simulation study was carried out to investigate the proposed semi-blind joint ML channel estimation and data detection scheme. We considered a MIMO system with $n_T = 4$ and $n_R = 4$. The modulation scheme was quadrature phase shift keying. The number of pilot symbols used was $t = 4$. The achievable performance was assessed in the simulation using three metrics, and these were the mse defined in (6), the mean channel error (MCE) defined as

$$J_{\text{MCE}}(\check{\mathbf{H}}) = \|\mathbf{H} - \check{\mathbf{H}}\|^2 \tag{10}$$

with $\mathbf{H}$ denoting the true MIMO channel matrix, and the bit error ratio (BER). The three algorithmic parameters of the RWBS were found empirically and the values used were $P_S = 5$, $N_G > 200$ and $N_I = 30$.

Fig. 1 shows the MCE performance after 1000 OHRSA evaluations over a range of γ values. The results were averaged over 50 different channel realisations. It can be seen from Fig. 1 that the optimal value of γ for this case was 0.04. This value of γ was used for all the other simulations. Figs. 2 and 3 depict the convergence performance of the proposed semi-blind scheme averaged over 50 different channel realisations and given $\gamma = 0.04$ in terms of the mse and MCE, respectively, for different SNR values as well as for two frame lengths $N = 50$ and 100. It can be seen from Fig. 2 that the mse converged to the noise floor. Fig. 4 shows the BER improvement of the semi-blind scheme over the training based one with the same number of pilot symbols, in comparison with the

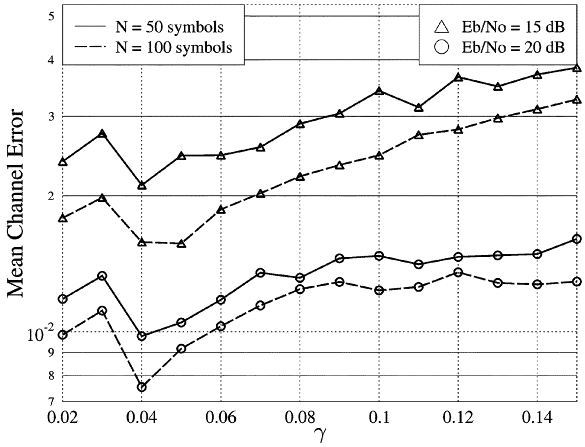

Fig. 1. Mean channel error as a function of γ after 1000 OHRSA evaluations and averaged over 50 different channel realizations, for two different values of E_b/N_o and two different values of N.

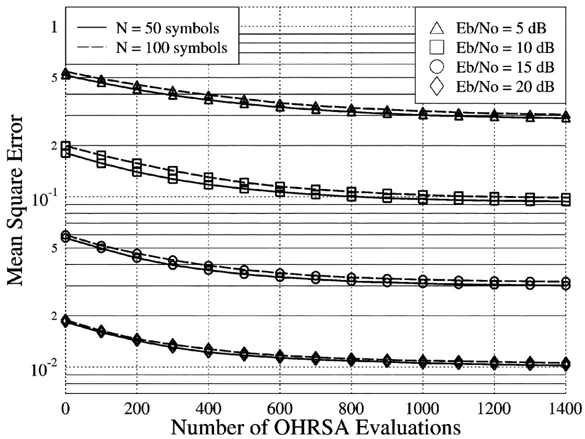

Fig. 2. Convergence of mean square error averaged over 50 different channel realizations and given $\gamma = 0.04$, for different values of E_b/N_o and N.

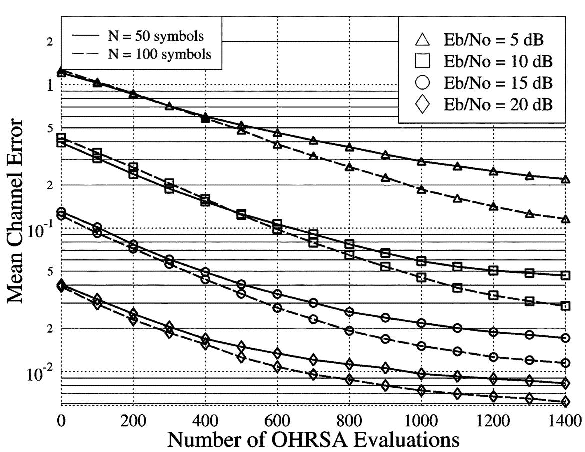

Fig. 3. Convergence of mean channel error averaged over 50 different channel realisations and given $\gamma = 0.04$, for different values of E_b/N_o and N.

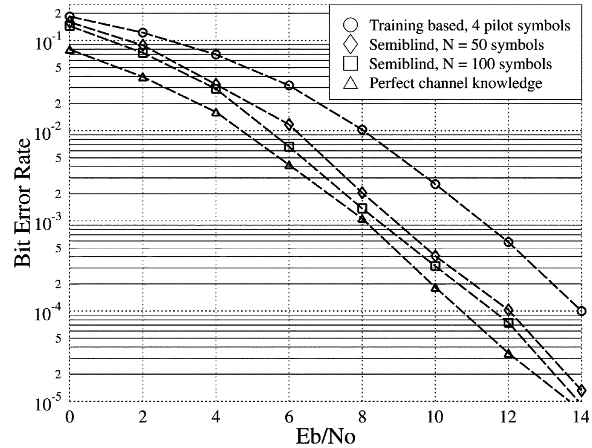

Fig. 4. Bit error rate of the proposed semi-blind scheme with $\gamma = 0.04$ and two different values of frame length N after 1200 OHRSA evaluations, in comparison with the training-based case using only four pilot symbols and the case of perfect channel knowledge.

case of perfect channel knowledge. It was also observed in our simulation study that, for the training-based scheme to achieve the same BER performance of the semi-blind one having a frame length $N = 50$ and with only 4 pilot symbols, the number of training symbols had to increase to at least 16. For the graphic clarification, the BER curve of the training-based scheme with 16 pilot symbols was not shown in Fig. 4.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Person. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[4] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[5] L. Tang, R. W. Liu, V. C. Soon, and Y. F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits Syst.*, vol. 38, no. 5, pp. 499–509, May 1991.

[6] A. Medles and D. T. M. Slock, "Semiblind channel estimation for MIMO spatial multiplexing systems," in *Proc. VTC2001-Fall*, Oct. 7–11, 2001, vol. 2, pp. 1240–1244.

[7] C. Cozzo and B. L. Hughes, "Joint channel estimation and data detection in space-time communications," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1266–1270, Aug. 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. Technol.*, vol. 53, no. 4, pp. 1085–1104, Jul. 2004.

[9] T. Wo, P. A. Hoeher, A. Scherb, and K. D. Kammeyer, "Performance analysis of maximum-likelihood semiblind estimation of MIMO channels," in *Proc. VTC2006-Spring*, Melbourne, Australia, May 7–10, 2006, vol. 4, pp. 1738–1742.

[10] A. K. Jagannatham and B. D. Rao, "Whitening-rotation-based semi-blind MIMO channel estimation," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 861–869, Mar. 2006.

[11] S. Chen and Y. Wu, "Maximum likelihood joint channel and data estimation using genetic algorithms," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1469–1473, May 1998.

[12] S. Chen, X. C. Yang, L. Chen, and L. Hanzo, "Blind joint maximum likelihood channel estimation and data detection for SIMO systems," *Int. J. Autom. Comput.*, vol. 4, no. 1, pp. 47–51, 2007.

[13] S. Chen, X. X. Wang, and C. J. Harris, "Experiments with repeating weighted boosting search for optimization in signal processing applications," *IEEE Trans. Syst., Man, Cybern. B*, vol. 35, no. 4, pp. 682–693, Aug. 2005.

[14] J. Akhtman, A. Wolfgang, S. Chen, and L. Hanzo, "An optimized-hierarchy-aided approximate log-MAP detector for MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1900–1909, May 2007.

[15] D. Pham, K. R. Pattipati, P. K. Willet, and J. Luo, "An improved complex sphere decoder for V-BLAST systems," *IEEE Signal Process. Lett.*, vol. 11, no. 9, pp. 748–751, Sep. 2004.

[16] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.