Blind MIMO Eigenmode Transmission Based on the Algebraic Power Method

Tobias Dahl, Nils Christophersen, and David Gesbert, *Member, IEEE*

***Abstract*—Identification of the channel matrix is of main concern in wireless multiple input multiple output (MIMO) systems. To maximize the SNR, the best way to utilize a MIMO system is to communicate on the top singular vectors of the channel matrix. Here, we present a new approach for direct blind identification of the main independent singular modes, without estimating the channel matrix itself. The right and left singular vectors with maximum corresponding singular values are determined using payload data and are continuously updated while at the same time being used for communication. The feasibility of the approach is demonstrated by simulating the performance over a noisy, fading time-varying channel. Mathematically, the technique is related to the iterative numerical Power method for finding eigenvalues of a matrix as well as the "time reversal mirror" technique developed within acoustics.**

***Index Terms*—Channel identification, eigen-modes, MIMO systems, singular modes, singular value decomposition (SVD).**

I. Introduction

Wireless multiple input multiple output (MIMO) systems are capable of delivering large increases in capacity through utilization of parallel communication channels [9], [10], [17]. Appearing first in a series of information theory articles published by members of Bell Labs, MIMO systems now constitute a major research area in telecommunications. It is also considered to be one of the technologies that have a chance to resolve the bottlenecks of traffic capacity in the forthcoming broadband wireless Internet access networks [Universal Mobile Telephone Services (UMTS)—and beyond].

Multiple antennas, both at the transmitter and the receiver, create a *matrix* channel. The key advantage is the possibility of transmitting over several spatial *modes* of the channel matrix within the same time-frequency slot at no additional power expenditure. In addition, if the channel matrix is known both at the transmitter (TX) and the receiver (RX), certain spatial modes (*singular modes*) of the matrix channel can be used to maximize the SNR for every realization of the channel. The singular modes can be used to transport independent data streams (to increase data rate), or one may choose to exploit the top mode (associated with the largest singular value) in order to maximize the spatial diversity advantage.

Manuscript received September 3, 2002; revised June 23, 2003. Part of this work was supported by the Norwegian operator Telenor under MIMO contract 15010010. Part of this work was presented at ICASSP 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alex B. Gershman.

The authors are at the Department of Informatics, University of Oslo, N-0316 Blindern, Norway (e-mail: tobias@ifi.uio.no; nilsch@ifi.uio.no; gesbert@ifi.uio.no).

Digital Object Identifier 10.1109/TSP.2004.832000

For an N(receive) $\times$ M(transmit) channel matrix $\mathbf{H}$ of rank $K_0 \leq \min(N, M)$, these modes are realized through the singular value decomposition (SVD) $\mathbf{H} = \mathbf{USV}^*$. Here, $(.)^*$ denotes the complex conjugate transpose. $\mathbf{S}$ is the diagonal matrix of singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{K_0} > 0$, and

$$\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_{K_0}] \in C^{N \times K_0} \quad (1)$$

$$\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_{K_0}] \in C^{M \times K_0} \quad (2)$$

are unitary matrices whose columns can be used as receive and transmit vectors $\{\mathbf{u}_i\}$ and $\{\mathbf{v}_i\}$, respectively. One can select a number K ($K \leq K_0$) of transmit/receive vectors for communication. Using the top K SVD elements, K independent data streams can be carried. With $K = 1$, the MIMO channel is used to provide the maximum diversity benefit. In a scattering-rich environment, MIMO channels will have full rank, but to preserve a diversity advantage on each substream, it makes sense to choose K strictly below K_0.

In practice, however, the singular modes are seldom utilized for communication. The reason is that with present methods, one would have to know the complete channel matrix at both sides and perform computationally demanding SVD operations. The cost, both in terms of reduced bandwidth and processor capacity, has hitherto been considered too high to make this approach practical. Current techniques typically require knowledge of the channel matrix only at the receiver, either through training data (e.g., V-BLAST, [9], [11]) or blind estimation (e.g., [2], [7], [13], [15], [18]–[20]). Compared with the singular vector approach, this implies that the transmit array gain is not realized and that one is unable to transmit on selected top singular vectors, those giving maximum performance/complexity tradeoff.

In this paper, we further extend our method outlined in [3], which treats a time division duplex (TDD) system exhibiting reciprocity, i.e., if the uplink channel is $\mathbf{H}$ at some point in time, the downlink channel is $\mathbf{H}^T$. The technique allows estimation and tracking of the singular vectors at both sides while, at the same time, utilizing these singular vector estimates for two-way communication. No explicit knowledge of $\mathbf{H}$ is needed, and no training data are required. The technique relies on a key "need to know" observation: In order to both transmit and receive in this way, one party needs only one set of the top K singular vectors, say, $\{\mathbf{u}_i\}$, and the other party only the corresponding set $\{\mathbf{v}_i\}$; no party needs both sets, and other aspects of $\mathbf{H}$ are irrelevant to both parties. The method estimates the necessary eigenstructure directly in a distributed manner so that no party is actually estimating the channel matrix *per se*, unlike other classes of algorithms [7], [13], [15], [18], [19]. There is no need

TABLE I
SUMMARY OF SYMBOLS

Symbol	Meaning
$\mathbf{H} \in C^{N \times M}$	channel matrix
$K_0 \leq min(N, M)$	rank of $\mathbf{H}$
$K \leq K_0$	number of channels chosen for communication
$\{\mathbf{u}_i\}$,$\{\mathbf{v}_i\}$	singular vectors
$\{\sigma_i\}$	singular values
$\mathbf{U}, \mathbf{S}, \mathbf{V}$	singular vectors/values matrices
$\hat{\mathbf{U}}, \hat{\mathbf{S}}, \hat{\mathbf{V}}$	estimates of singular vector/values
n	symbol block length
$c, c_x, c_y \in C$	symbols
$\mathbf{C}, \mathbf{C}_x, \mathbf{C}_y \in C^{K \times n}$	symbol blocks
$\hat{\mathbf{C}}, \hat{\mathbf{C}}_x, \hat{\mathbf{C}}_y$	symbol block estimates
$\mathbf{x}, \mathbf{y}$	transmit/receive vectors
$\mathbf{X}, \mathbf{Y}$	transmit/receive blocks

for SVD calculations or higher order statistics that would usually be needed if one wanted to estimate the complete channel blindly.

The basic principle utilized follows from a simple result in matrix theory: If an arbitrary initial data vector is transmitted over the same reciprocal and noiseless TDD MIMO channel multiple times, the result will converge toward a top singular vector of the matrix—the left top vector on one side and the right top vector on the other side. In numerical linear algebra, this result forms the basis for the so-called *Power method*, which is an iterative method for finding eigenvectors and eigenvalues of a matrix [12]. ("Power" here refers to the fact that the method raises the matrix to higher and higher powers.) This result from matrix theory also underpins the time reversal mirror technique developed by Fink [8] in acoustics. By repeatedly sending a sound pulse into a medium, recording the reflected signal and resending it after normalization and time-reversal (i.e., complex conjugation in the frequency domain), convergence toward a top eigenvector is reached. This physical process is nothing but a *Power method*.

Computationally, our approach basically comes at the expense of a QR decomposition (essentially a Gram–Schmidt orthogonalitation) for every singular vector update. This is far less than for an SVD, which involves the estimation of eigenvectors/eigenvalues and extraction of square roots etc. In fact, the QR is commonly used as a starting point for the SVD [12].

We believe that our approach has the potential for making the singular vector approach feasable for TDD MIMO systems. It does not reduce the effective bandwidth since no training data are required, and the computational burden is modest. The critical test for the method will be if the quality of the singular vector estimates are sufficiently good to produce acceptable BER values for fading, time-varying channels at common SNRs. In [3], we gave a first presentation of basic principles for the BPSK case; here, we provide a full description in a more general setting and further simulation results. To keep matters simpler, we consider the number of singular modes to use K as a fixed design parameter and load all K singular modes equally (no water filling).

Other authors have also commented on sending-resending schemes. Andersen [1] has observed independently that such a procedure leads to convergence toward the top singular mode of the channel. Kilfoyle [14], on the other hand, uses training data to find singular modes in a nonflat fading (underwater) channel but also comments that there is important information in the data vectors sent up and down. However, none of these estimate the singular modes while using the channel for communication.

The paper is organized as follows. In Section II, we present the algorithm, and in Section III, we demonstrate performance and convergence by simulations. In Section IV, we conclude on the findings and discuss lines of further research.

The symbols used in the subsequent elaborations are summarized in Table I.

II. Algorithm

A. Preliminaries: Algebraic Power Method

We briefly recapture the basic version of the Power method, which is an iterative numerical method for finding the top eigenvector of real symmetric matrices. For a full reference, see [12]. Assume a symmetric matrix $\mathbf{A} \in R^{P \times P}$ of rank $r \leq P$. It has an eigen-decomposition

$$\mathbf{A} = \sum_{j=1}^{r} \lambda_j \mathbf{v}_j \mathbf{v}_j^T \tag{3}$$

where $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ are the orthonormal eigenvectors corresponding to the strictly positive and ordered eigenvalues

$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0$. Now, decompose an arbitrary vector $\mathbf{x}^{(0)} \in R^P$:

$$\mathbf{x}^{(0)} = \mathbf{x}_R^{(0)} + \mathbf{x}_0^{(0)} = \sum_{k=1}^{r} \alpha_k \mathbf{v}_k + \mathbf{x}_0^{(0)} \tag{4}$$

for some α_k, where $\mathbf{x}_R^{(0)}$ and $\mathbf{x}_0^{(0)}$ are in the range and null space of $\mathbf{A}$, respectively. Assuming $\mathbf{x}_R^{(0)} \neq 0$ without loss of generality (numerically this is always the case), repeated premultiplication of $\mathbf{x}^{(0)}$ by $\mathbf{A}$ leads to the *Power* term $\mathbf{A}^i$:

$$\mathbf{x}^{(i)} = \underbrace{\mathbf{A}\mathbf{A}\cdots\mathbf{A}}_{i \text{ times}} \mathbf{x}^{(0)} = \mathbf{A}^i \mathbf{x}^{(0)}.$$

The null space component disappears, and

$$\mathbf{x}^{(i)} = \left(\sum_{i=j}^{r} \lambda_j^i \mathbf{v}_j \mathbf{v}_j^T\right) \left(\sum_{k=1}^{r} \alpha_k \mathbf{v}_k\right) = \sum_{j=1}^{r} \lambda_j^i \alpha_j \mathbf{v}_j \tag{5}$$

will be dominated by the term $\lambda_1^i \alpha_1 \mathbf{v}_1$ when i tends to infinity. If the vector $\mathbf{x}^{(i)}$ is normalized after each iteration, this becomes a method for finding the top eigenvector $\mathbf{v}_1$. An extension of this technique called the *method of orthogonal iterations* computes several top eigenvectors by including a QR othogonalization step (basically a Gram–Schmidt procedure) in each iteration. It also carries over to nonsymmetric and nonreal matrices [12]. A related algorithm (sometimes called NIPALS [21]) can be used for finding singular vectors, which is what we are interested in.

In the present paper, we consider the application of these ideas to two-way transmission over an unknown and time-varying TDD MIMO channel. Transmission and retransmission take the role of the iterations in the Power method. Reciprocity is maintained by assuming that the ping-pong time—the time between the begining of a downlink (DL) frame and the begining of the next uplink (UL) frame—is small compared to the channel variability.

B. Estimating the Top Singular Vector Pair Without Communication Data

We start by introducing the basic feedback scheme underlying the procedure and neglect communication data for the time being. Consider first transmission of single vectors only (block length 1) over a (noiseless) flat-fading complex channel $\mathbf{H}$

$$\mathbf{y}^{(i)} = \mathbf{H}\mathbf{x}^{(i-1)} \quad \text{uplink} \tag{6}$$
$$\mathbf{z}^{(i)} = \mathbf{H}^T\mathbf{w}^{(i)} \quad \text{downlink.} \tag{7}$$

Here, $\mathbf{x}^{(i-1)}$ and $\mathbf{w}^{(i)}$ are the UL and DL vectors, respectively. Feedback is introduced as follows:

$$\mathbf{w}^{(i)} := \bar{\mathbf{y}}^{(i)}$$
$$\mathbf{x}^{(i)} := \bar{\mathbf{z}}^{(i)}.$$

Here, the bar $\bar{(.)}$ denotes the complex conjugate vector (elementwise, without vector/matrix transposition). In effect, this states that the signal received by one party is returned to the other party after it has been complex conjugated. This scheme is turned into an algebraically equivalent form that lends itself more easiliy to analysis. To do this, rewrite (7) as follows:

$$\mathbf{z}^{(i)} = \mathbf{H}^T\mathbf{w}^{(i)} = \mathbf{H}^T\bar{\mathbf{y}}^{(i)}$$
$$\mathbf{z}^{(i)} = \bar{\mathbf{x}}^{(i)} = \mathbf{H}^T\bar{\mathbf{y}}^{(i)}$$

and when conjugating the latter equality, we get

$$\mathbf{x}^{(i)} = \mathbf{H}^*\mathbf{y}^{(i)}. \tag{8}$$

Through this *complex conjugation trick*, (6) and (8) will serve as the basis for our analysis.

Following the Power method, transmission starts with an arbitrary vector $\mathbf{x}^{(0)} \in C^M$. Let $\mathbf{H} = \sum_{i=1}^{K_0} \sigma_i \mathbf{u}_i \mathbf{v}_i^*$ be the SVD of $\mathbf{H}$ including only the singular vectors corresponding to the strictly positive singular values. Decomposing $\mathbf{x}^{(0)}$, we get $\mathbf{x}^{(0)} = \sum_{j=1}^{K_0} \alpha_j \mathbf{v}_j + \mathbf{x}_0^{(0)}$ for some set $\{\alpha_j\}$ of constants. Neglecting $\mathbf{x}_0^{(0)}$, which is the component in the null space of $\mathbf{H}$, we obtain

$$\mathbf{y}^{(1)} = \mathbf{H}\mathbf{x}^{(0)} = \left(\sum_i \sigma_i \mathbf{u}_i \mathbf{v}_i^*\right)\left(\sum_j \alpha_j \mathbf{v}_j\right) = \sum_i \sigma_i \alpha_i \mathbf{u}_i \tag{9}$$

$$\mathbf{x}^{(1)} = \mathbf{H}^*\mathbf{y}^{(1)} = \left(\sum_j \sigma_j \mathbf{u}_j \mathbf{v}_j^*\right)^* \left(\sum_i \sigma_i \alpha_i \mathbf{u}_i\right) = \sum_i \sigma_i^2 \alpha_i \mathbf{v}_i. \tag{10}$$

Continuing this way, one arrives at the following recursion, for $i \geq 1$:

$$\mathbf{y}^{(i)} = \sum_k \sigma_k^{(2i-1)} \alpha_k \mathbf{u}_k \tag{11}$$

$$\mathbf{x}^{(i)} = \sum_k \sigma_k^{(2i)} \alpha_k \mathbf{v}_k. \tag{12}$$

Clearly, $\mathbf{y}^{(i)}$ will be dominated by $\mathbf{u}_1$, and $\mathbf{x}^{(i)}$ by $\mathbf{v}_1$ as $i \to \infty$. If, after each iteration, normalization is applied, e.g.,

$$\mathbf{x}^{(i)} := \frac{\mathbf{x}^{(i)}}{||\mathbf{x}^{(i)}||_2} \tag{13}$$

$$\mathbf{y}^{(i)} := \frac{\mathbf{y}^{(i)}}{||\mathbf{y}^{(i)}||_2} \tag{14}$$

where $||.||_2$ denotes the L^2 norm, then $\mathbf{x}^{(i)} \to \mathbf{v}_1$ and $\mathbf{y}^{(i)} \to \mathbf{u}_1$ as $i \to \infty$ up to multiplication of unit norm complex numbers. The uplink party will therefore hold $\mathbf{u}_1$, and the downlink will have $\mathbf{v}_1$.

C. Estimating K Singular Vector Pairs Without Communication Data

Now, assume that some estimates $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{v}}_1$ are known. The following iterations are then used to estimate the *second pair of singular vectors*, $\mathbf{u}_2$, $\mathbf{v}_2$:

$$\mathbf{y}_2^{(i)} = (\mathbf{I} - \hat{\mathbf{u}}_1\hat{\mathbf{u}}_1^*)\mathbf{H}\mathbf{x}_2^{(i-1)} \tag{15}$$
$$\mathbf{x}_2^{(i)} = (\mathbf{I} - \hat{\mathbf{v}}_1\hat{\mathbf{v}}_1^*)\mathbf{H}^*\mathbf{y}_2^{(i)} \tag{16}$$

where $\mathbf{x}_2^{(0)}$ and $\mathbf{y}_2^{(0)}$ are again arbitrary. Normalization of the vectors $\mathbf{y}_2^{(i)}$, $\mathbf{x}_2^{(i)}$ must be included in the same way as in (13) and (14). The Gram–Schmidt-like operations on the right of (15) and (16) remove the contribution of the first singular vector pair in the sums. Consequently, $\mathbf{x}_2^{(i)}$ and $\mathbf{y}_2^{(i)}$ will now converge toward the second pair of singular vectors $\mathbf{u}_2$ and $\mathbf{v}_2$, provided that the estimates of the first singular vectors are sufficiently good. With an obvious generalization of the equations above, we can find *all desired K singular vector pairs* by keeping successive estimates perpendicular to each other and of unit length. If one wants to estimate the rth singular vector pair ($1 < r \leq K$), and the previous $r-1$ pairs are already computed, one uses

$$\mathbf{y}_r^{(i)} = \left(\mathbf{I} - \sum_{k=1}^{r-1} \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^*\right) \mathbf{H}\mathbf{x}_r^{(i-1)} \tag{17}$$

$$\mathbf{x}_r^{(i)} = \left(\mathbf{I} - \sum_{k=1}^{r-1} \hat{\mathbf{v}}_k \hat{\mathbf{v}}_k^*\right) \mathbf{H}^*\mathbf{y}_r^{(i)} \tag{18}$$

always with a subsequent normalization.

Introducing blocks, the following algorithm implements the above approach in an efficient way, holding the full set of K singular vectors as columns of $\hat{\mathbf{U}}$, $\hat{\mathbf{V}}$.

1) $\mathbf{X}_0 = \text{random}, i = 1$.
2) $\mathbf{Y}_i = \mathbf{H}\mathbf{X}_{i-1}$.
3) $\mathbf{QR} = \mathbf{Y}_i$, $\hat{\mathbf{U}} := \mathbf{Q}$, $\mathbf{Y}_i := \hat{\mathbf{U}}$.
4) $\mathbf{X}_i = \mathbf{H}^*\mathbf{Y}_i$.
5) $\mathbf{QR} = \mathbf{X}_i$, $\hat{\mathbf{V}} := \mathbf{Q}$, $\mathbf{X}_i := \hat{\mathbf{V}}$.
6) Increase i, and repeat from 2) until convergence.

Here, $\mathbf{X}_i \in C^{M\times K}$ is the UL block, and $\mathbf{Y}_i \in C^{N\times K}$ is the DL block. Note that $\mathbf{QR} = \mathbf{Z}$ denotes the decomposition of a matrix $\mathbf{Z}$ into one unitary matrix $\mathbf{Q}$ and one matrix $\mathbf{R}$, which is upper triangular with real positive diagonal elements. The matrices $\mathbf{X}_i, \mathbf{Y}_i$ converge to the matrices of singular vectors $\mathbf{X}_i \rightarrow \mathbf{V}, \mathbf{Y}_i \rightarrow \mathbf{U}$ up to multiplication of unit norm complex numbers. This follows from an extension of the proof of the method of orthogonal iteration [12].

Note that one part of the job (recording, orthogonalization, conjugation, and resending) is carried out by one party and the corresponding (but independent) part by the other one. This gives an operational framework where only the left set of singular vectors $\{\mathbf{u}_i\}$ or the right set $\{\mathbf{v}_i\}$ are known by each party, which is sufficient for operating the channel modes.

D. Simultaneous Singular Vector Estimation and Two-Way Communication

The above algorithm is generalized so that two-way transmission of information occurs simultaneously with blind tracking and estimation of the singular vectors. To introduce the complete algorithm, assume first that the singular vector pairs $\{\mathbf{u}_i\}$, $\{\mathbf{v}_i\}$ are *known*. One could then use the *transmit vectors*

$$UL: \quad \mathbf{x}_s = \sum_{j=1}^{K} c_j^x \mathbf{v}_j \tag{19}$$

$$DL: \quad \mathbf{y}_s = \sum_{j=1}^{K} c_j^y \mathbf{u}_j \tag{20}$$

where $c_j^x, c_j^y \in C$ are the information-bearing data symbols to be transmitted uplink and downlink. When these vectors are transmitted, and assuming no noise for now, they will be received as

$$\mathbf{y}_r = \mathbf{H}\mathbf{x}_s = \mathbf{H}\sum_{j=1}^{K} c_j^x \mathbf{v}_j = \sum_{j=1}^{K} \sigma_j c_j^x \mathbf{u}_j \tag{21}$$

$$\mathbf{x}_r = \mathbf{H}^*\mathbf{y}_s = \mathbf{H}^*\sum_{j=1}^{K} c_j^y \mathbf{u}_j = \sum_{j=1}^{K} \sigma_j c_j^y \mathbf{v}_j. \tag{22}$$

The symbols are then decoded using corresponding sets of the singular vectors:

$$\hat{c_j^x} = \mathbf{u}_j^*\mathbf{y}_r = \sigma_j c_j^x \tag{23}$$

$$\hat{c_j^y} = \mathbf{v}_j^*\mathbf{x}_r = \sigma_j c_j^y. \tag{24}$$

Using a slicer, the (scaled) symbols c_j^x and c_j^y can be decided. Note that to both transmit and receive, each party needs only *one* set of singular vectors—either $\{\mathbf{u}_i\}$ or $\{\mathbf{v}_i\}$—but not both.

In the proposed algorithm, the true singular vectors are replaced by their estimates using the approach described in the previous section. Introduce first the symbol blocks $\mathbf{C}_i^x, \mathbf{C}_i^y \in C^{K\times n}$ (uplink and downlink, respectively) corresponding to a slot size $n \geq K$. These blocks contain n vectors of size K to be transmitted over the K independent channel modes. The basic version of the proposed algorithm is given below. A flowchart is shown in Fig. 1, where the complex conjugation trick has been replaced by the actual operations taking place on each side.

BIMA - Blind Iterative MIMO Algorithm

1) $\hat{\mathbf{U}}_0 \in C^{N\times K} = \text{random}$, $\hat{\mathbf{V}}_0 \in C^{M\times K} = \text{random}$, $\mathbf{X}_0^{\text{Send}} = \mathbf{V}_0\mathbf{C}_0^x, i = 1$.

---

2) $\mathbf{Y}_i^{\text{Rec}} = \mathbf{H}\mathbf{X}_{i-1}^{\text{Send}}$.

---

3) Decide $\hat{\mathbf{C}}_{i-1}^x$ from $\hat{\mathbf{U}}_{i-1}^*\mathbf{Y}_i^{\text{Rec}}$.
4) $\mathbf{QR} = \mathbf{Y}_i^{\text{Rec}}\hat{\mathbf{C}}_{i-1}^{x\dagger}$, $\hat{\mathbf{U}}_i = \mathbf{Q}$.
5) $\mathbf{Y}_i^{\text{Send}} = \hat{\mathbf{U}}_i\mathbf{C}_i^y$.

---

6) $\mathbf{X}_i^{\text{Rec}} = \mathbf{H}^*\mathbf{Y}_i^{\text{Send}}$.

---

7) Decide $\hat{\mathbf{C}}_i^y$ from $\hat{\mathbf{V}}_{i-1}^*\mathbf{X}_i^{\text{Rec}}$.
8) $\mathbf{QR} = \mathbf{X}_i^{\text{Rec}}\hat{\mathbf{C}}_i^{y\dagger}$, $\hat{\mathbf{V}}_i = \mathbf{Q}$.
9) $\mathbf{X}_i^{\text{Send}} = \hat{\mathbf{V}}_i\mathbf{C}_i^x$.

---

10) Increase i, repeat from 2).

Here, the symbol † denotes the pseudo-inverse (Moore–Penrose) operator. In the above, $\mathbf{X}_i^{\text{Send}}$ and $\mathbf{Y}_i^{\text{Send}}$ are the data blocks transmitted uplink and downlink, respectively. $\mathbf{Y}_i^{\text{Rec}}$ and $\mathbf{X}_i^{\text{Rec}}$ are the correspondingly received data blocks.

If the symbols are decoded correctly (steps 3 and 7), then $\hat{\mathbf{C}}_i^x\mathbf{C}_i^{x\dagger} = \mathbf{I}$, $\hat{\mathbf{C}}_i^y\mathbf{C}_i^{y\dagger} = \mathbf{I}$ ($n \geq K$), and this algorithm is completely equivalent to the one in the preceding section. Convergence is then assured. For random initial conditions for $\hat{\mathbf{U}}_0$ and $\hat{\mathbf{V}}_0$, convergence has to be determined from simulations, as have the effects of channel variability and noise.

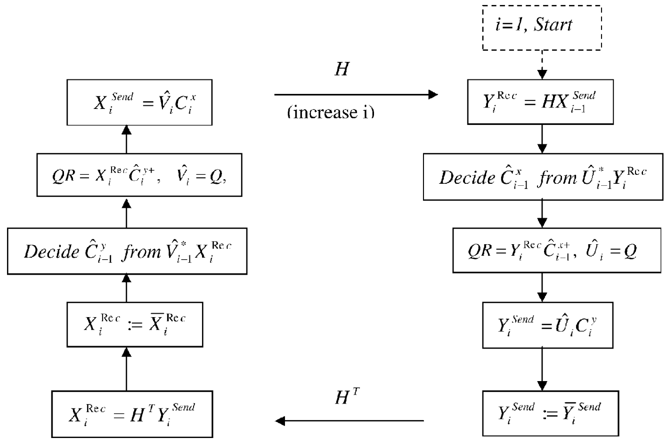

Fig. 1. Overall flowchart of the BIMA algorithm, showing the main computations that each party has to perform.

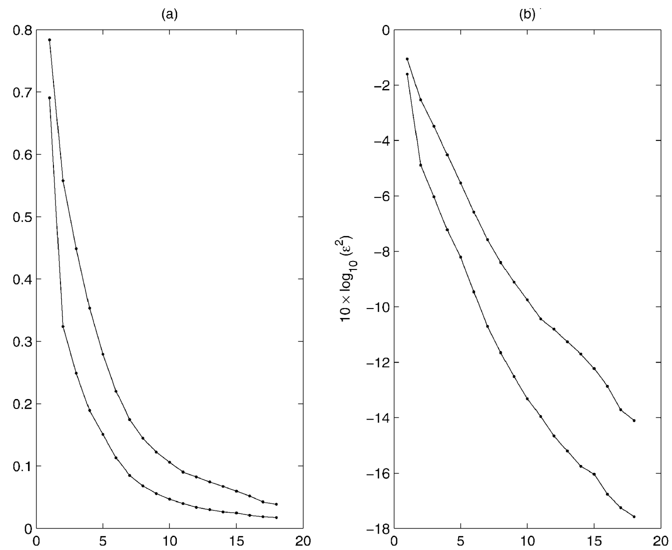

Fig. 2. Ensemble averages of ϵ^2 (now averaged over both left and right singular vectors) for singular vector estimates over 100 constant, randomly chosen 3TX $\times$ 3RX channels as a function of the iteration number. The upper curve is the error for the two second singular vectors (combining $\mathbf{u}_2$, $\mathbf{v}_2$), and the lower is the corresponding measure for the first ones. (a) Errors on a linear scale. (b) Errors on the decibel scale.

The flowchart and the algorithm are slightly simpler than the actual implementation. In practice, subblocks of length n_0, $K < n_0 \leq n$, and of full rank are used to estimate and track the singular vectors within each UL/DL frame. This improves the performance under time-varying conditions and makes the pseudo inverse computationally cheap. In addition, some smoothing is applied to the singular vectors to reduce the influence of noise. This is done by averaging singular vector estimates while making sure that orthogonality is preserved (see [5] for details). In addition, we assume that the K channels are color coded so that the different bit streams can be recognized by the recipient once separated (a typical blind MIMO estimation ambiguity). In addition, the convergence speed is boosted by sorting the column vectors of the matrices to be QR-decomposed by their norm.

As the singular vectors are determined up to multiplication by a complex number of unit norm, differential coding and decoding have to be added to the BIMA algorithm in order to avoid symbol rotation problem, as in any blind MIMO approach.

III. Results

This section presents results for BIMA in a simulated TDD MIMO environment, using differentially encoded quadrature phase shift keying (QPSK). Several issues are investigated with emphasis on convergence, tracking capabilities, and performance in cases where multiple nonzero singular values occur.

A prior concern is the estimation of the singular vectors. For *constant* channels with separate singular values, the estimation error for, say, a left singular vector estimate $\hat{\mathbf{u}}_k$ of $\mathbf{H}$ can be measured as

$$\epsilon^2 = \|\hat{\mathbf{u}}_k - \mathbf{u}_k\|^2 \tag{25}$$

where $\mathbf{u}_k$ is the true singular vector. If several singular modes ($K \geq 2$) are in operation, the average error is used:

$$\epsilon^2 = \frac{1}{K}\sum_{k=1}^{K} \|\hat{\mathbf{u}}_k - \mathbf{u}_k\|^2. \tag{26}$$

For *time-varying channels*, the situation is slightly different. Then, the singular values will at times approach one another and cross over. This is well known from the perturbation theory of the SVD [12]. For identical singular values, *any* linear combination of singular vectors is itself a singular vector. To determine the performance under such conditions, observe that valid sets of singular vectors $\{\mathbf{u}_k\}$ and $\{\mathbf{v}_k\}$ are always eigenvectors of $\mathbf{HH}^*$ and $\mathbf{H}^*\mathbf{H}$, respectively. For an estimate $\hat{\mathbf{u}}_k$, we then define the error measure

$$\epsilon^2_{\text{Eig}} = \left\| \frac{\mathbf{HH}^*\hat{\mathbf{u}}_k}{\|\mathbf{HH}^*\hat{\mathbf{u}}_k\|} - \hat{\mathbf{u}}_k \right\|^2. \tag{27}$$

It follows that $0 \leq \epsilon^2_{\text{Eig}} \leq 2$ with the lower bound achieved only when $\hat{\mathbf{u}}_k$ is an eigenvector. A corresponding measure is used for $\hat{\mathbf{v}}_k$. If several singular vectors are tracked, an average similar to (26) is used. To compute the performance of the channel over a time-frame $[t_1, t_m]$, evaluate the error $\epsilon^2_{\text{Eig}}(t_i)$ at times $t_1, t_2, \ldots, t_m$, and compute the average

$$\epsilon^2_{\text{Eig}} = \frac{1}{m}\sum_{i=1}^{m} \epsilon^2_{\text{Eig}}(t_i). \tag{28}$$

A. Estimation of Singular Vectors for Constant Channels

Consider first a noise-free case where 100 Monte Carlo simulations were performed for constant, randomly chosen 3TX $\times$ 3RX systems, estimating the two top singular vectors while simultaneously performing differentially encoded QPSK. The rate of convergence for our method is illustrated in Fig. 2, where the ensemble averages of ϵ^2 over the left and right singular vectors are shown on a linear scale (a) and on the decibel scale (b). In such cases, a trained approach would yield correct channel estimates with a few symbols from which the

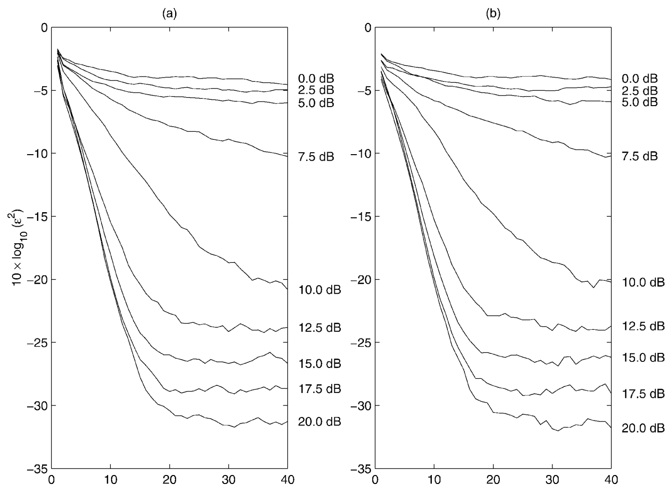

Fig. 3. Convergence of the singular vectors plotted as function of iteration step (slot index) at various SNRs. Here, $K = 2$ is the number of singular modes (channels) used for a fixed 3TX × 3RX system. (a) Ensemble average of $\epsilon^2 = (1/K)\sum_{k=1}^{K} \|\mathbf{u}_k - \hat{\mathbf{u}}_k\|^2$ for the left singular vectors. (b) Corresponding error for the right singular vectors.

exact singular vectors could be found. Although less efficient than training, the figure shows rapid convergence for BIMA.

Consider now a *fixed* 3TX × 3RX system with additive noise. Fig. 3 shows the convergence of the two top singular pairs from 1000 random starting points. Here, $\sigma_1 = 4$, $\sigma_2 = 2$, $\sigma_3 = 1$, and various levels of channel noise are tested. The errors in both singular modes are averaged and convergence for the left and right hand side vectors plotted. Rapid convergence takes place above a certain SNR, and this behavior has been demonstrated over a range of other simulation scenarios. However, the initial bit error rates may be unacceptable, and hence, an initial acquisition period should be used in practice.

B. Tracking and Communication Over Time-Varying Channels

Here, we study a fast time-varying, Rayleigh-fading 4TX × 4RX channel ($D = 50$ Hz Doppler spread), with a transmission rate of $f = 220$ kBit/s per eigen-mode, using $K = 2$ independent channels and a ping-pong time corresponding to 1 ms. This implies a block length $n = 220$ bits. Fig. 4 shows a typical example of the environment in which we are working. The singular values are plotted as functions of time over a period of 100 ms. At this Doppler spread, the channel is completely changed every 20 ms. The two crosses indicate positions where there is a shift in the order of the singular values as the third singular value overtakes the second.

1) Eigenmode Tracking: The error measure $\epsilon^2_{\mathrm{Eig}}$ in (28), taken over the two top singular vectors left and right, are plotted for several SNR scenarios (Fig. 5). This indicates the accuracy with which the singular vectors are tracked for this fast time-varying channel.

2) Bit Error Rates—Comparison with Trained Methods: Ten simulations were performed for each SNR scenario and the mean BER over these simulations taken. In Fig. 6, we plot the BER obtained with our method after an initial acquisition period. For comparison, we also show results for the clairvoyant case (perfectly known SVD) and cases where 5% and 20% of the data, respectively, are used for training. The loss between BIMA and the other cases is in the

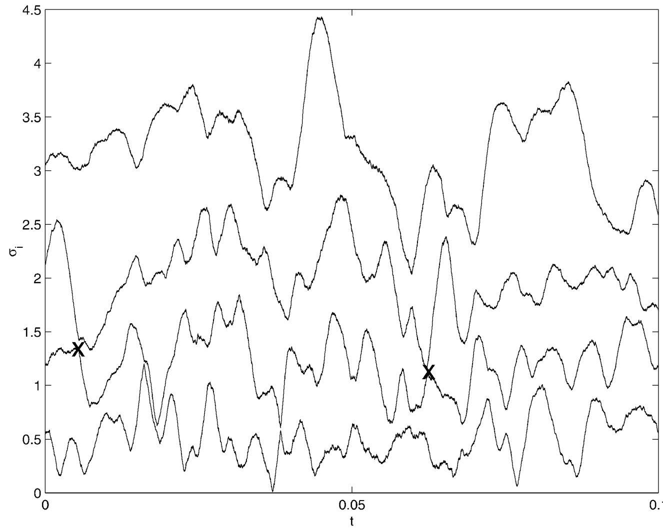

Fig. 4. Example of a time-varying 4TX × 4Rx channel, fading at 50 Hz Doppler over a period of 0.1 s. The singular values are plotted as functions of time. The two crosses indicate shifts in the order of the singular values.

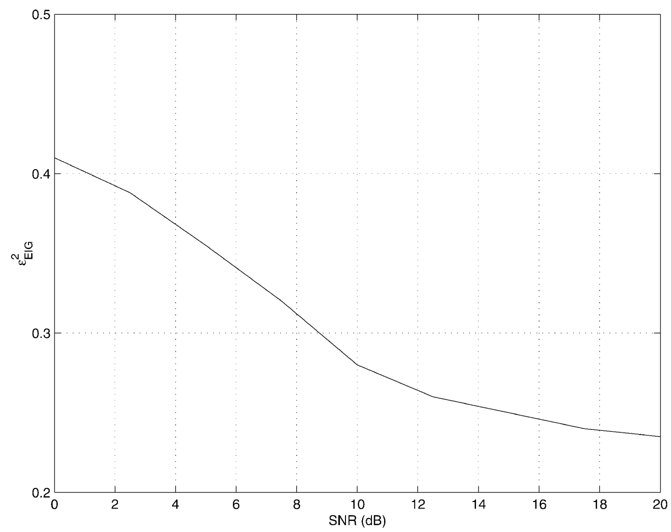

Fig. 5. $\epsilon^2_{\mathrm{Eig}}$ error measure (28) for singular vectors, mean over two top singular modes, right- and left-hand side.

range 2–5 dB, depending on the noise level. (Training-based methods may be improved if knowledge of a certain space-time precoding structure is employed at the transmitter [16]. This is not the case here.) There seems to be a "flooring effect," which is even more prominent for faster fading channels, but which disappears for slower channels (e.g., 10- or 20-Hz Doppler spread). The reason is probably the inherent delay in estimating the singular vectors in highly time-varying channels.

It is of interest to see how successfully the BIMA algorithm retains the MIMO advantage, despite the blindness. Fig. 7 shows bit error rates for time-varying SISO and MIMO channels with the same doppler (50 Hz, as above) and various SNRs. It is seen that the diversity advantage of the MIMO channel is well exploited by the blind algorithm.

C. Crossing Singular Values

As noted, Fig. 4 shows two cases where the third singular value overtakes the second and takes its place. Here, BIMA switches correctly and continues to track the top two singular

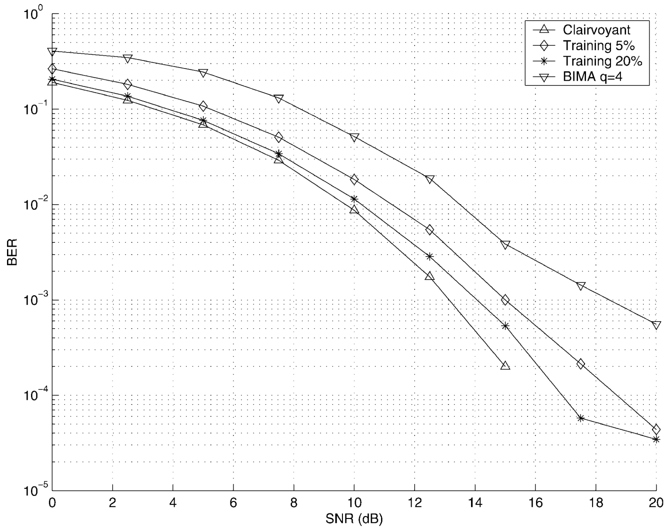

Fig. 6. Mean BER values for various SNRs for a 4TX $\times$ 4RX channel using two eigen-modes, fading at 50 Hz Doppler, and transmitting at $f = 22$ kBits/s. BIMA is compared to cases where 5% and 20% of the data, respectively, are used for training.

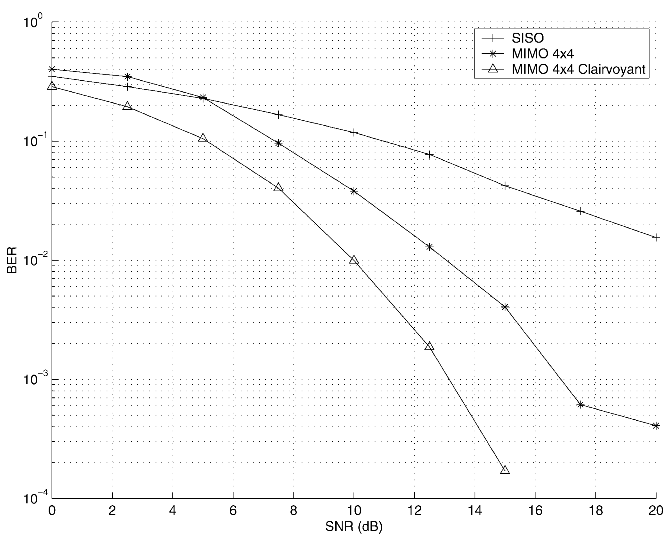

Fig. 7. Comparison of BIMA with a SISO time-varying channel, fading at 50 Hz doppler, with a transmit frequency of $f = 220$ kBits/s and a ping-pong period corresponding to 1 ms.

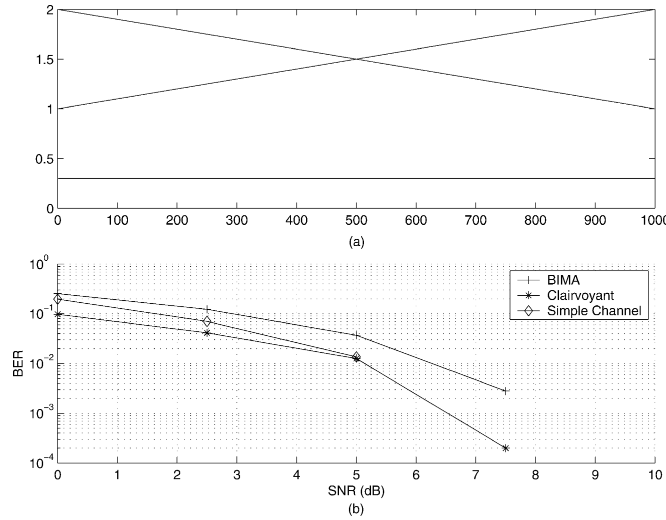

Fig. 8. Crossing singular values. (a)Singular values for a 3TX $\times$ 3RX channel. (b) BER for various SNR scenarios, both for the BIMA algorithm (top), the theroetical optimum (bottom) and BIMA for the constant channel without changing singular values.

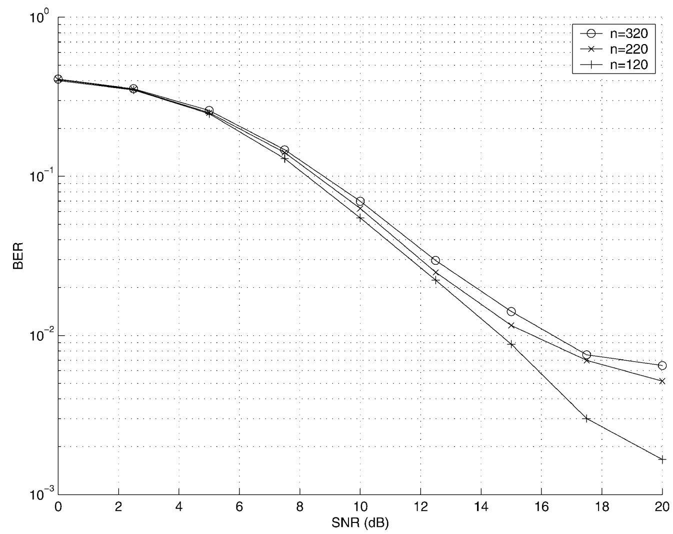

Fig. 9. Bit error rates for various choices of n. The algorithm is stable with respect to this choice, and the loss is only a few decibels for most SNRs.

modes. However, this switching results in a transient increase in the BER.

This situation is further investigated in Fig. 8. For the sake of a simple illustration, one top singular mode of a 3TX by 3RX channel matrix is initially tracked [see Fig. 8(a)]. The corresponding singular value is decreasing while another singular mode with an increasing singular value is also present. At a certain point in time, the latter singular value overtakes the first and becomes the leading one. Fig. 8(b) shows the bit error rates versus the SNR. The upper curve is for BIMA, and the lower is the theoretical optimum (known singular vectors). The middle curve shows the BER for a similar channel but without the swap in the singular values. This illustrates the extra loss associated with crossing singular values.

D. Choice of n

The slot size n is a design parameter closely related to the ping-pong time of the channel through the transmission frequency. We simulated the performance of the algorithm with three different slot sizes $n = 320$, $n = 220$, and $n = 110$, again in the 50-Hz Doppler fading 4TX $\times$ 4RX environment, using $K = 2$ singular modes. The results in Fig. 9 show that BIMA is robust with respect to choice of the slot size; there is a loss of only a few decibels for most relevant SNRs.

IV. Discussion

We have shown that top singular modes of the channel matrix can be estimated and tracked without training data, without the need for a statistical estimate of $\mathbf{H}$, and without performing an actual SVD. These results are based on a channel exhibiting reciprocity. Transmission and retransmission in a MIMO system of an initially arbitrary vector will lead to convergence toward the right and left singular vectors of the main mode. By combining an orthogonalization procedure with this transmission/retransmission scheme, multiple singular vector pairs can be extracted as part of the normal operation.

More work needs to be done both on theoretical and practical aspects, which could lead to further improvements. For example, in numerical linear algebra, the Power method for finding eigenvectors has been replaced by more efficient Krylov methods [12], which do not rely on the convergence of the iterative process but utilize intermediate iterates. Such methods could form the basis for improved estimation of the singular MIMO vectors.

In the simulations, we assumed that color coding was applied to separate the channels at the recipient side. However, as singular value estimates are available for channel marking, it is possible to avoid this extra cost. Then, one must deal with the issue of continuous ordering of the singular values. This can be taken care of by adopting a leader-follower strategy: One party takes responsibility for switching the singular vectors when the difference between singular values is too big to be caused by noise-driven errors. This is explained in more detail in [5].

In other work [4], we develop a method for blind estimation of singular modes in a frequency division duplex (FDD) channel. In this case, the uplink channel matrix is generally not the transpose of the downlink channel, and therefore, BIMA iterations cannot be used. Other techniques are needed to blindly estimate these modes, which are now double in number (one set for the downlink channel, another for the uplink). Initial work [5] shows that singular modes can still be estimated, provided the two parties consider the task to be a joint optimization problem, and devise suitable cooperative strategies.

References

[1] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Selected Areas Commun.*, vol. 18, pp. 2172–2178, Nov. 2000.

[2] P. Comon, "Blind separation of convolutive mixtures, A partial joint diagonalization (PAJOD) approach,", Res. Rep. 2002.04, I3S, 2002.

[3] T. Dahl, N. Christophersen, and D. Gesbert, "BIMA – Blind iterative MIMO algorithm," in *Proc. ICASSP*, 2002, pp. 2365–2368.

[4] T. Dahl, N. Christophersen, O. C. Lingjærde, and N. Lid Hjort, "The game of blind MIMO estimation," Ph.D. dissertation, Dept. Informatics, Univ. Oslo, Oslo, Norway, 2002.

[5] T. Dahl, "From sensory analysis to mobile communication systems," Ph.D. dissertation, Dept. Informatics, Univ. Oslo, 2002.

[6] T. Dahl and N. Christophersen, Convergence Properties of the BIMA Algorithm, to be published.

[7] Z. Ding, "Matrix outer-product decomposition method for blind multiple channel identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 3053–3061, Dec. 1997.

[8] M. Fink, "Time-reversed acoustics," *Phys. Today*, vol. 20, pp. 34–40, 1997.

[9] G. J. Foschini, "Layered space-time architecture for wireless communications in a fading environment," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

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

[11] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniasky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–15, 1999.

[12] G. Golub and C. F. Van Loan, *Matrix Computations*, third ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[13] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.

[14] D. B. Kilfoyle, "Spatial modulation in the underwater acoustic communication channel," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MD, 2000.

[15] C. B. Papadias, "A multi-user kurtosis algorithm for blind source separation," in *Proc. ICASSP*, Istanbul, Turkey, June 5, 2000.

[16] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Trans. Signal Processing*, vol. 50, pp. 2489–2498, Oct. 2002.

[17] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.

[18] A. Touzni, I. Fijalkow, M. G. Larimore, and J. R. Treichler, "A globally convergent approach for blind MIMO adaptive deconvolution," *IEEE Trans. Signal Processing*, vol. 49, pp. 1166–1178, June 2001.

[19] J. K. Tugnait, "Identification and deconvolution of multichannel non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Processing*, vol. 45, pp. 658–672, Mar. 1997.

[20] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155, May 1996.

[21] H. Wold, "Soft modeling by latent variables: The nonlinear iterative partial least squares (NIPALS) approach," *Perspect. Probab. Stat., Pap. Honor M. S. Bartlett Occas. 65th Birthday*, pp. 117–142, 1975.

Tobias Dahl received the M.Sc. and Ph.D. degrees from the Department of Informatics, University of Oslo, Oslo, Norway, in 1998 and 2002, respectively.

He now holds a postdoctoral position, first at Imperial College, London, U.K., and currently at the University of Oslo. His work in array signal processing for wireless communication has its roots in work he carried out earlier related to multivariate data analysis techniques from areas such as chemometrics, sensory analysis, and statistical shape analysis.

Nils Christophersen received the M.Sc. degree in cybernetics from the University of Oslo, Oslo, Norway, in 1972 and the Ph.D. degree from the same university in 1982.

He worked at the Norwegian Defense Research Establishment and The Foundation for Scientific and Industrial Research (SINTEF) in Norway before joining the Department of Informatics as professor in 1991. He spent a year as a visiting scientist at the Centre for Ecology and Hydrology, Wallingford, U.K., from 1988 to 1989 and had a sabbatical stay at the Department of Statistics, Stanford University, Stanford, CA, from 1997 to 1998. His research interests include array signal processing algorithms with applications to wireless communication.

David Gesbert (S'96–M'99) received the Ph.D. degree from Ecole Nationale Superieure des Telecommunications, Paris, France, in 1997.

From 1993 to 1997, he was with France Telecom Research, Paris. From 1997to 1998, he was a Postdoctoral Fellow at the Information Systems Laboratory, Stanford University, Stanford, CA. In October 1998, he co-founded Iospan Wireless (now Intel), Palo Alto, CA: a startup company designing high-speed wireless internet access networks. In 2001, he joined the Signal and Imaging Processing Group, Department of Informatics, at the University of Oslo, Oslo, Norway, as an Associate Professor in parallel to his other activities. He has published more that 50 conference and journal papers and holds several patents in the area of signal processing and communications.

Dr. Gesbert has served on the Technical Program Committee of various IEEE conferences. He is an editor of the IEEE Transactions on Wireless Communications.