BLIND EQUALIZATION OF MIMO CHANNELS USING DETERMINISTIC PRECODING

Harold Artés and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
phone: +43 1 58801 38923, fax: +43 1 58801 38999, email: hartes@aurora.nt.tuwien.ac.at
web: http://www.nt.tuwien.ac.at/dspgroup/time.html

ABSTRACT

We present a novel precoding or modulation scheme (*matrix modulation*) that allows parallel transmission of several data signals over an unknown multiple-input multiple-output (MIMO) channel. We first present a theorem on unique signal demodulation and an efficient iterative demodulation algorithm for transmission over an unknown instantaneous-mixture channel. We then generalize our results to an unknown MIMO channel with memory.

1. INTRODUCTION

Using multiple antennas at transmitter and receiver allows a significant increase of data rates. Most transmission schemes (e.g., [1–4]) require the receiver to know the multiple-input multiple-output (MIMO) channel; unfortunately, training symbol based channel estimation significantly reduces the effective data rates. Other methods (e.g., [5–8]) do not require knowledge of the channel.

In this paper, we propose a scheme for transmission over an *unknown* MIMO channel whereby K parallel data streams $d_k[n]$ ($k = 1, \cdots, K$) are "precoded" into M_{T} antenna input signals $s_k[n]$ ($k = 1, \cdots, M_{\mathrm{T}}$) using a novel matrix modulation technique. The receiver demodulates M_{R} antenna output signals $x_k[n]$ ($k = 1, \cdots, M_{\mathrm{R}}$) into data estimates $\hat{d}_k[n]$ (see Fig. 1).

This paper is organized as follows. Section 2 presents a theorem on unique demodulation and an efficient iterative demodulation algorithm for the special case of memoryless MIMO channels (instantaneous mixture channels). Section 3 extends the demodulation method to general MIMO channels. Section 4 presents an improved iterative equalization method for general MIMO channels. Finally, simulation results are provided in Section 5.

2. INSTANTANEOUS MIXTURE CHANNEL

We first consider an instantaneous mixture channel (i.e., a memoryless MIMO channel). The channel's input-output relation is[1]

$$\mathbf{x}[n] = \mathbf{H}\,\mathbf{s}[n]\,, \tag{1}$$

with the transmit vector $\mathbf{s}[n] \triangleq [s_1[n] \cdots s_{M_{\mathrm{T}}}[n]]^T$, the received vector $\mathbf{x}[n] \triangleq [x_1[n] \cdots x_{M_{\mathrm{R}}}[n]]^T$, and the (unknown) $M_{\mathrm{R}} \times M_{\mathrm{T}}$ channel matrix $\mathbf{H}$.

The matrix modulation precoding forces a "modulation structure" on the $M_{\mathrm{T}} \times N$ transmit signal matrix $\mathbf{S} \triangleq [\mathbf{s}[0] \cdots \mathbf{s}[N-1]]$ (with some block length N) according to

$$\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k\,, \tag{2}$$

with the K diagonal $N \times N$ data matrices $\mathbf{D}_k \triangleq \mathrm{diag}\{d_k[0], \cdots, d_k[N-1]\}$ and K "modulation matrices" $\mathbf{M}_k$ of size $M_{\mathrm{T}} \times N$.

Combining N successive received vectors into the matrix $\mathbf{X} \triangleq [\mathbf{x}[0] \cdots \mathbf{x}[N-1]]$ and inserting (2) into (1), we obtain

Funding by FWF grant P12228-TEC.

[1]We consider noiseless transmission; the effect of noise will be studied in Section 5.

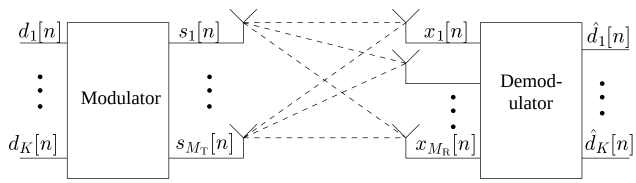

Fig. 1: Multi-input, multi-antenna transmission setup.

$$\mathbf{X} = \mathbf{H}\mathbf{S} = \mathbf{H} \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k\,. \tag{3}$$

Under certain conditions, the structure of $\mathbf{S}$ defined by (2) is strong enough to allow unique reconstruction (up to a constant factor) of the data sequences $d_k[n]$ from $\mathbf{X}$. That is, the received matrix $\mathbf{X} = \mathbf{H} \sum_{k=1}^{K'} \mathbf{M}_k \mathbf{D}_k$ (where $K' \leq K$ is the number of active data streams, i.e., we allow some data streams to be zero) does not permit a different representation $\tilde{\mathbf{H}} \sum_{k=1}^{K} \mathbf{M}_k \tilde{\mathbf{D}}_k$.

Theorem 1. *Let $\mathbf{D}_k$ and $\tilde{\mathbf{D}}_k$ be K' resp. K diagonal matrices of size $N \times N$, with $\mathbf{D}_k$ nonsingular, and let $\mathbf{H}$ and $\tilde{\mathbf{H}}$ be matrices of size $M_{\mathrm{R}} \times M_{\mathrm{T}}$ and with full rank. Let $K' \leq K < M_{\mathrm{T}}$ and $M_{\mathrm{R}} \geq M_{\mathrm{T}}$. Then there exist some N with $\left\lceil \frac{M_{\mathrm{T}}^2 - 1}{M_{\mathrm{T}} - K} \right\rceil \leq N \leq M_{\mathrm{T}}^2 - 1$ and K matrices $\mathbf{M}_k$ ($k = 1, \cdots, K$) of size $M_{\mathrm{T}} \times N$ such that*

$$\tilde{\mathbf{H}} \sum_{k=1}^{K} \mathbf{M}_k \tilde{\mathbf{D}}_k \;=\; \mathbf{H} \sum_{k=1}^{K'} \mathbf{M}_k \mathbf{D}_k \tag{4}$$

(for one given set of data matrices $\mathbf{D}_k$) implies[2] $\tilde{\mathbf{H}}^{\#}\mathbf{H} = c\mathbf{I}$ and

$$\tilde{\mathbf{D}}_k \;=\; \begin{cases} c\,\mathbf{D}_k\,, & k \leq K' \\ \mathbf{0}\,, & K' < k \leq K\,, \end{cases} \tag{5}$$

where $c \in \mathbb{C}$ is an unknown factor.

A sketch of the proof, including one specific construction of the modulation matrices $\mathbf{M}_k$, is given in the Appendix.

Demodulation can be performed similarly to [9]. From (3),

$$\mathbf{x}[n] = \mathbf{H} \sum_{k=1}^{K} d_k[n]\,\mathbf{m}_k[n]\,, \qquad n = 0, \cdots, N-1\,, \tag{6}$$

where $\mathbf{m}_k[n]$ is the nth column of $\mathbf{M}_k$. Multiplying (6) by $\mathbf{H}^{\#}$ from the left and using $\mathbf{H}^{\#}\mathbf{H} = \mathbf{I}$, the N linear equations in (6) can be rewritten as $\mathbf{Q}\mathbf{y} = \mathbf{0}$, where $\mathbf{Q}$ is an $M_{\mathrm{T}}N \times (M_{\mathrm{T}}M_{\mathrm{R}} + KN)$ matrix that contains the known quantities $\mathbf{x}[n]$ and $\mathbf{m}_k[n]$, and $\mathbf{y}$ is an $(M_{\mathrm{T}}M_{\mathrm{R}} + KN) \times 1$ vector that contains the unknowns $(\mathbf{H}^{\#})_{k,l}$

[2]Here, $\mathbf{I}$ is the identity matrix and $\tilde{\mathbf{H}}^{\#}$ is the pseudo-inverse of $\tilde{\mathbf{H}}$.

and $d_k[n]$. The least-squares solution, $\mathbf{y}_{\rm LS} = \operatorname{argmin}_{\|\mathbf{y}\|=1} \|\mathbf{Q}\mathbf{y}\|$, is given by the right singular vector of $\mathbf{Q}$ corresponding to the smallest singular value. This method is however computationally intensive because of the large size of $\mathbf{Q}$,

POCS demodulation algorithm. As an alternative, we now present an efficient iterative demodulation algorithm that is inspired by [10]. Given a received matrix $\mathbf{X} = \mathbf{H}\mathbf{S}$ and modulation matrices $\mathbf{M}_k$, it follows from Theorem 1 that the matrix $\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k$ and, thus, the K data matrices $\mathbf{D}_k$ are uniquely determined (up to a scalar factor) by the following two properties:

1. $\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k$ with $\mathbf{D}_k$ diagonal;
2. the row span of $\mathbf{S}$ equals the row span of $\mathbf{X}$.

(The second property follows from $\mathbf{X} = \mathbf{H}\mathbf{S}$ with $\mathbf{H}$ full rank). Thus, $\mathbf{S} \in \mathcal{A} \cap \mathcal{B}$ where $\mathcal{A}$ denotes the linear subspace of all matrices $\sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k$ with $\mathbf{M}_k$ given and $\mathbf{D}_k$ diagonal, and $\mathcal{B}$ denotes the linear subspace of all matrices whose row span lies in the row span of $\mathbf{X}$, i.e., of all matrices of the form $\mathbf{B}\mathbf{X}$ with some $M_{\rm T} \times M_{\rm R}$ matrix $\mathbf{B}$. Since both $\mathcal{A}$ and $\mathcal{B}$ are linear subspaces and thus convex, the formulation $\mathbf{S} \in \mathcal{A} \cap \mathcal{B}$ suggests a POCS (projections onto convex sets) algorithm [11] for calculating and demodulating $\mathbf{S}$. This algorithm is iterative and consists in alternately projecting the iterated version of $\mathbf{S}$ onto $\mathcal{A}$ and $\mathcal{B}$.

Projection onto $\mathcal{A}$: The projection onto $\mathcal{A}$ amounts to forming $\mathbf{S}^{(i)} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k^{(i)}$, where the nonzero (diagonal) elements of $\mathbf{D}_k^{(i)}$ can be shown to be given by

$$\big(\mathbf{D}_k^{(i)}\big)_{n,n} \,=\, \frac{1}{M_{\rm T}} \sum_{l=1}^{M_{\rm T}} \big(\mathbf{S}^{(i-1)}\big)_{l,n} \big(\mathbf{M}_k^{+}\big)_{l,n} \,. \tag{7}$$

Here, $\mathbf{S}^{(i-1)}$ is the result of the previous iteration (i.e., the projection onto $\mathcal{B}$, see below) and the $M_{\rm T} \times N$ matrices $\mathbf{M}_k^{+}$ are defined such that $(\mathbf{m}_k^{+}[n])^T$, the transpose of the nth column of $\mathbf{M}_k^{+}$, equals the kth row of the $K \times M_{\rm T}$ matrix $\big[\mathbf{m}_1[n] \cdots \mathbf{m}_K[n]\big]^{\#}$. If the vectors $\mathbf{m}_1[n], \cdots, \mathbf{m}_K[n]$ are orthonormal, then there is simply $\mathbf{M}_k^{+} = \mathbf{M}_k^{*}$ where $\mathbf{M}_k^{*}$ is the complex conjugate of $\mathbf{M}_k$.

Projection onto $\mathcal{B}$: The projection onto $\mathcal{B}$ amounts to forming $\mathbf{S}^{(i)} = \mathbf{B}^{(i)}\mathbf{X}$, where it can be shown that $\mathbf{B}^{(i)} = \mathbf{S}^{(i-1)}\mathbf{X}^{\#}$. Here, $\mathbf{S}^{(i-1)}$ is the result of the previous iteration (i.e., the projection onto $\mathcal{A}$, see above) and $\mathbf{X}^{\#}$ is the pseudo-inverse of $\mathbf{X}$, which can be pre-calculated before starting the iteration.

The POCS algorithm is guaranteed to converge to an intersection point, i.e., $\mathbf{S}^{(\infty)} \in \mathcal{A} \cap \mathcal{B}$ [11]. Thus, $\mathbf{S}^{(\infty)} = c\,\mathbf{S}$ and $\mathbf{D}_k^{(\infty)} = c\mathbf{D}_k$ where the $\mathbf{D}_k$ are the true data matrices and $c \in \mathbb{C}$. The convergence speed depends on the initialization, $\mathbf{S}^{(0)}$. In the semiblind case, some known input symbols can be used to calculate a good initialization. Another way to speed up convergence is to use *relaxation* [11] and/or knowledge of the data symbol alphabet. (cf. [12]). The latter approach, however, introduces a nonconvex set and thus convergence to the desired solution is no longer guaranteed. For large N, the POCS method typically is much more efficient than the demodulation method discussed previously.

3. GENERAL MIMO CHANNEL

We will now extend our method to a MIMO channel with memory (intersymbol interference). Here, the input-output relation is

$$\mathbf{x}[n] \,=\, \sum_{m=0}^{L-1} \mathbf{H}[m]\,\mathbf{s}[n-m]\,, \tag{8}$$

where the $M_{\rm R} \times M_{\rm T}$ matrices $\mathbf{H}[m]$ constitute the channel's impulse response and $L-1$ is the channel's maximum time delay.

The matrix modulation precoding is still given by (2), however with the $M_{\rm T} \times (N+L-1)$ transmit signal matrix $\mathbf{S} \triangleq [\mathbf{s}[-L+1] \cdots \mathbf{s}[N-1]]$ and the $(N+L-1) \times (N+L-1)$ diagonal data matrices $\mathbf{D}_k \triangleq \operatorname{diag}\{d_k[-L+1], \cdots, d_k[N-1]\}$.

Next, we will write the input-output relation (8) in block matrix form. Setting $\mathbf{H}' \triangleq \big[\mathbf{H}[0] \cdots \mathbf{H}[L-1]\big]$, we define the following channel block matrix of size $M_{\rm R}p \times M_{\rm T}(L+p-1)$, in which $\mathbf{H}'$ is stacked p times with shifts to the left by $M_{\rm T}$ positions each (the stacking parameter p is called *smoothing factor* [13, 14]),

$$\boldsymbol{\mathcal{H}} \triangleq \begin{bmatrix} \mathbf{0} & & \boxed{\mathbf{H}'} \\ & \boxed{\mathbf{H}'} & \\ & \iddots & \\ \boxed{\mathbf{H}'} & & \mathbf{0} \end{bmatrix}.$$

We also form the following transmit block matrix of size $M_{\rm T}(L+p-1) \times (N-p+1)$,

$$\boldsymbol{\mathcal{S}} \triangleq \begin{bmatrix} \mathbf{s}[p-1] & \mathbf{s}[p] & \cdots & \mathbf{s}[N-1] \\ \mathbf{s}[p-2] & \mathbf{s}[p-1] & \cdots & \mathbf{s}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{s}[-L+1] & \mathbf{s}[-L+2] & \cdots & \mathbf{s}[N-L-p+1] \end{bmatrix}.$$

This is a block-Toeplitz matrix that is "generated" by the columns of $\mathbf{S} = [\mathbf{s}[-L+1] \cdots \mathbf{s}[N-1]]$. Thus, we shall call $\mathbf{S}$ the *generating matrix* of $\boldsymbol{\mathcal{S}}$. Finally, we form the following received block-Hankel matrix of size $M_{\rm R}p \times (N-p+1)$,

$$\boldsymbol{\mathcal{X}} \triangleq \begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \cdots & \mathbf{x}[N-p-1] & \mathbf{x}[N-p] \\ \mathbf{x}[1] & \mathbf{x}[2] & \cdots & \mathbf{x}[N-p] & \mathbf{x}[N-p+1] \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ \mathbf{x}[p-1] & \mathbf{x}[p] & \cdots & \mathbf{x}[N-2] & \mathbf{x}[N-1] \end{bmatrix}.$$

Now (8) can be written as (cf. [13, 14])

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{S}}\,. \tag{9}$$

Blind equalization of the unknown channel (described by $\boldsymbol{\mathcal{H}}$) corresponds to calculation of the transmit matrix $\boldsymbol{\mathcal{S}}$ (or, equivalently, its generating matrix $\mathbf{S}$) from the known received matrix $\boldsymbol{\mathcal{X}}$. For this to be possible, $\boldsymbol{\mathcal{S}}$ must be a wide matrix and the row span of $\boldsymbol{\mathcal{X}}$ must be equal to the row span of $\boldsymbol{\mathcal{S}}$ [13, 14]. This, in turn, requires that $\boldsymbol{\mathcal{H}}$ is a square or tall matrix and has full rank. These requirements lead to the necessary conditions [13] $p \geq \frac{M_{\rm T}(L-1)}{M_{\rm R}-M_{\rm T}}$, with $M_{\rm R} > M_{\rm T}$, and $N > M_{\rm T}L + (M_{\rm T}+1)(p-1)$. Blind equalization can then be done as follows:

Step 1: Using a singular value decomposition (SVD), the row span of $\boldsymbol{\mathcal{S}}$ is calculated from $\boldsymbol{\mathcal{X}}$ [13, 14].

Step 2: Another SVD is used to construct an $M_{\rm T}(L+p-1) \times (N-p+1)$ block-Toeplitz matrix $\boldsymbol{\mathcal{S}}_A$ whose row span equals that of $\boldsymbol{\mathcal{S}}$ [13, 14]. It can be shown [13] that the $M_{\rm T} \times (N+L-1)$ generating matrix of $\boldsymbol{\mathcal{S}}_A$ can be written as $\mathbf{S}_A = \mathbf{A}\mathbf{S}$, where $\mathbf{A}$ is an unknown invertible matrix of size $K \times K$. Due to the SVD construction, the rows of $\mathbf{S}_A$ are orthonormal.

Step 3: The unknown instantaneous mixture defined by $\mathbf{S}_A = \mathbf{A}\mathbf{S}$ is analogous to the instantaneous mixture $\mathbf{X} = \mathbf{H}\mathbf{S}$ in (3). Thus, it can be resolved using the demodulation methods of Section 2, whereby $\mathbf{S}$ and, in turn, the data sequences $d_k[-L+1], d_k[-L+2], \cdots, d_k[N-1]$ are obtained up to a common constant factor.

We next present an alternative POCS method that is computationally more efficient and allows to incorporate *a-priori* knowledge such as the symbol alphabet.

4. POCS EQUALIZATION ALGORITHM

A computationally intensive part of the three-step method of Section 3 is the SVD in Step 2 that is used to construct the generating matrix $\mathbf{S}_A$. This SVD can be avoided by the following approach. As mentioned in Section 3, because of its block-Toeplitz structure $\mathbf{S}$ can be reconstructed from the row span of $\boldsymbol{\mathcal{X}}$ up to an unknown instantaneous mixture. According to Section 2, this instantaneous mixture can be resolved based on the modulation structure $\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k$. Consequently, $\boldsymbol{\mathcal{S}}$ is uniquely determined (up to a scalar factor) by the following two properties:

1. $\boldsymbol{\mathcal{S}}$ is block-Toeplitz and its generating matrix has modulation structure, i.e., $\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k$ with $\mathbf{D}_k$ diagonal;
2. the row span of $\boldsymbol{\mathcal{S}}$ equals the row span of $\boldsymbol{\mathcal{X}}$.

(The second property follows from (9) with $\boldsymbol{\mathcal{H}}$ full rank). Thus, $\boldsymbol{\mathcal{S}} \in \mathcal{A} \cap \mathcal{B}$ where $\mathcal{A}$ is the linear subspace of all block-Toeplitz matrices with generating matrix $\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k$, with $\mathbf{M}_k$ given and $\mathbf{D}_k$ diagonal, and $\mathcal{B}$ is the linear subspace of all matrices whose row span lies in that of $\boldsymbol{\mathcal{X}}$, i.e., of all matrices of the form $\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{X}}$ with some $M_{\mathrm{T}}(L+p-1) \times Mp$ matrix $\boldsymbol{\mathcal{B}}$. This again suggests a POCS algorithm for calculating $\boldsymbol{\mathcal{S}}$ that consists in alternately projecting the iterated version of $\boldsymbol{\mathcal{S}}$ onto $\mathcal{A}$ and $\mathcal{B}$.

Projection onto $\mathcal{A}$: $\boldsymbol{\mathcal{S}}$ being a linear structured matrix [15], it can be shown that the projection onto $\mathcal{A}$ can be performed by the following two steps:

Step 1: Enforce block-Toeplitz property. Let $\boldsymbol{\mathcal{S}}^{(i-1)}$ be the result of the previous iteration (projection onto $\mathcal{B}$, see below). From $\boldsymbol{\mathcal{S}}^{(i-1)}$, which is not block-Toeplitz, we calculate an $M_{\mathrm{T}} \times (N+L-1)$ "pseudo generating matrix" $\tilde{\mathbf{S}}^{(i-1)}$ as follows. The first one of the M_{T} rows of $\tilde{\mathbf{S}}^{(i-1)}$ is obtained by averaging properly aligned and zero-padded versions of the first, $(M_{\mathrm{T}}+1)$st, $(2M_{\mathrm{T}}+1)$st, etc. rows of $\boldsymbol{\mathcal{S}}^{(i-1)}$. More precisely, the first row of $\boldsymbol{\mathcal{S}}^{(i-1)}$ is shifted to the right by one position and added to the $(M_{\mathrm{T}}+1)$st row of $\boldsymbol{\mathcal{S}}^{(i-1)}$, with zeros appended where necessary. The result is again shifted to the right by one position and added to the $(2M_{\mathrm{T}}+1)$st row of $\boldsymbol{\mathcal{S}}^{(i-1)}$, etc. Finally the jth element of the resulting row vector of length $N+L-1$ is divided by the jth element of $(1, 2, \cdots, M_{\mathrm{T}}, M_{\mathrm{T}}, \cdots, M_{\mathrm{T}}, M_{\mathrm{T}}-1, \cdots, 1)$ to yield the first row of $\tilde{\mathbf{S}}^{(i-1)}$. The second row of $\tilde{\mathbf{S}}^{(i-1)}$ is obtained similarly by averaging properly aligned and zero-padded versions of the second, $(M_{\mathrm{T}}+2)$nd, $(2M_{\mathrm{T}}+2)$nd, etc. rows of $\boldsymbol{\mathcal{S}}^{(i-1)}$. In this manner, all M_{T} rows of $\tilde{\mathbf{S}}^{(i-1)}$ are obtained.

Step 2: Enforce modulation structure. Next, we form $\mathbf{S}^{(i)} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k^{(i)}$, where the nonzero (diagonal) elements of $\mathbf{D}_k^{(i)}$ can be shown to be given by (7) with $\mathbf{S}^{(i-1)}$ replaced by $\tilde{\mathbf{S}}^{(i-1)}$ and with $\mathbf{M}_k^{+}$ as defined in Section 2. We then form the block-Toeplitz matrix $\boldsymbol{\mathcal{S}}^{(i)}$ generated by $\mathbf{S}^{(i)}$.

Projection onto $\mathcal{B}$: The projection onto $\mathcal{B}$ amounts to forming $\boldsymbol{\mathcal{S}}^{(i)} = \boldsymbol{\mathcal{B}}^{(i)}\boldsymbol{\mathcal{X}}$, where it can be shown that $\boldsymbol{\mathcal{B}}^{(i)} = \boldsymbol{\mathcal{S}}^{(i-1)}\boldsymbol{\mathcal{X}}^{\#}$. Here, $\boldsymbol{\mathcal{S}}^{(i-1)}$ is the result of the previous iteration (projection onto $\mathcal{A}$, see above). The pseudo-inverse $\boldsymbol{\mathcal{X}}^{\#}$ can be pre-calculated.

The POCS algorithm is guaranteed to converge to an intersection point, i.e., $\boldsymbol{\mathcal{S}}^{(\infty)} \in \mathcal{A} \cap \mathcal{B}$ [11]. Thus, $\boldsymbol{\mathcal{S}}^{(\infty)} = c\boldsymbol{\mathcal{S}}$ and $\mathbf{D}_k^{(\infty)} = c\mathbf{D}_k$ with $c \in \mathbb{C}$. In the semiblind case, some known consecutive input symbols $d_k[n_1], \cdots, d_k[n_2]$ can be used to calculate a good initialization $\boldsymbol{\mathcal{S}}^{(0)}$ with generating matrix $\mathbf{S}^{(0)} = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k^{(0)}$ where $\mathbf{D}_k^{(0)} = \mathrm{diag}\{0, \cdots, 0, d_k[n_1], \cdots, d_k[n_2], 0, \cdots, 0\}$ [16]. Convergence can again be accelerated as outlined on Section 2.

5. SIMULATION RESULTS

We studied the performance of the proposed methods for $M_{\mathrm{T}} = 4$ transmit antennas and $K = 3$ uncoded QPSK data signals $d_k[n]$. The modulation matrices were constructed with $N = 200$ by taking realizations of iid Gaussian variables as matrix entries and then orthonormalizing the corresponding columns of all $\mathbf{M}_k$ (this always worked in our simulations). The channel impulse responses were randomly generated for each simulation run. The channel output signals were corrupted by white Gaussian noise with variance σ^2 and observed over an interval of length $N = 200$.

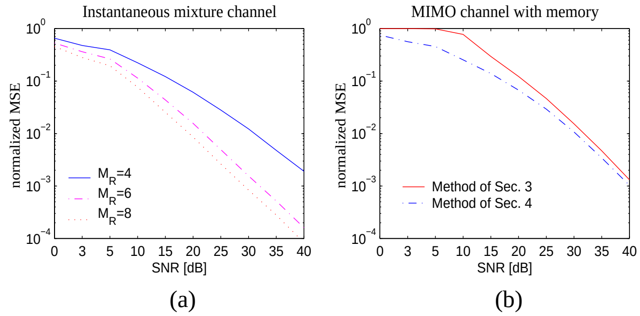

Fig. 2: Normalized MSE vs. SNR for (a) an instantaneous mixture channel, using the method of Section 2, and (b) a MIMO channel with memory, using the methods of Section 3 and Section 4.

First, we considered three instantaneous mixture channels with $M_{\mathrm{R}} = 4$, 6 and 8 receive antennas. Fig. 2(a) shows the normalized mean-square error (MSE) vs. the SNR[3] obtained with the POCS method of Section 2. It is seen that use of more receive antennas results in better signal estimation.

Next, we considered a MIMO channel with memory ($M_{\mathrm{R}} = 6$ receive antennas, channel impulse response length $L = 3$). Fig. 2(b) shows the normalized MSE vs. the SNR obtained with the methods of Section 3 and Section 4 (smoothing factor $p = 5$). It is seen that for low SNR, the method of Section 4 performs significantly better. This may be due to the fact that in Step 2 of the method of Section 3, a signal subspace/noise subspace allocation of certain singular vectors is required (cf. [13, 14]).

APPENDIX: SKETCH OF PROOF OF THEOREM 1

Setting $\mathbf{G} \triangleq \tilde{\mathbf{H}}^{\#}\mathbf{H}$, we can rewrite (4) as

$$\tilde{\mathbf{v}}_n = \mathbf{G}\mathbf{v}_n \,, \qquad n = 0, \cdots, N-1 \,, \tag{10}$$

with

$$\mathbf{v}_n \triangleq \sum_{k=1}^{K'} d_k[n]\, \mathbf{m}_k[n] \,, \qquad \tilde{\mathbf{v}}_n \triangleq \sum_{k=1}^{K} \tilde{d}_k[n]\, \mathbf{m}_k[n] \tag{11}$$

where $\mathbf{m}_k[n]$ denotes the nth column of $\mathbf{M}_k$. We have to show that there exist vectors $\mathbf{m}_1[n], \cdots, \mathbf{m}_K[n]$ such that (10) with (11) implies $\mathbf{G} = c\mathbf{I}$ with some $c \in \mathbb{C}$. Assuming that the vectors $\mathbf{m}_1[n], \cdots, \mathbf{m}_K[n]$ are linearly independent for n fixed, $\mathbf{G} = c\mathbf{I}$ together with (10) and (11) will imply the desired result in (5), viz., $\tilde{d}_k[n] = c\, d_k[n]$ for $1 \le k \le K'$ and $d_k[n] = 0$ for $K' < k \le K$.

The vectors $\mathbf{v}_n$ and $\tilde{\mathbf{v}}_n$ in (11) are elements of, respectively, the K'-dimensional subspace $\mathcal{C}_n = \mathrm{span}\,\{\mathbf{m}_1[n], \cdots, \mathbf{m}_{K'}[n]\}$ and the K-dimensional subspace $\tilde{\mathcal{C}}_n = \mathrm{span}\,\{\mathbf{m}_1[n], \cdots, \mathbf{m}_K[n]\}$ of $\mathbb{C}^{M_{\mathrm{T}}}$. Since $K' \le K < M_{\mathrm{T}}$, there is $\mathcal{C}_n \subseteq \tilde{\mathcal{C}}_n \subset \mathbb{C}^{M_{\mathrm{T}}}$. Hence, (10) means that $\mathbf{G}$ maps a *specific* vector $\mathbf{v}_n \in \mathcal{C}_n$ (defined by the specific transmitted data $d_k[n]$) to some vector $\tilde{\mathbf{v}}_n \in \tilde{\mathcal{C}}_n$. We have N such transformation relations. The question is, can vectors $\mathbf{m}_1[n], \cdots, \mathbf{m}_K[n]$ ($n = 0, \cdots, N-1$) be found such that these transformation relations imply $\mathbf{G} = c\mathbf{I}$ with some $c \in \mathbb{C}$, and how many such relations do we need (how large is N)?

Let us first study a different problem. Suppose we have I pairs of subspaces $\mathcal{C}_i = \mathrm{span}\,\{\mathbf{c}_1[i], \cdots, \mathbf{c}_{K'}[i]\}$ and $\tilde{\mathcal{C}}_i = \mathrm{span}\{\mathbf{c}_1[i],$

[3] The normalized MSE is defined as $\sum_{k=1}^{K} \sum_n |d_k[n] - \hat{c}\hat{d}_k[n]|^2 / \sum_{k=1}^{K} \sum_n |d_k[n]|^2$ averaged over all simulation runs, where $\hat{d}_k[n]$ is the estimate of $d_k[n]$ obtained with the respective method and $\hat{c}$ is the least-squares fit for the unknown factor c. The number of simulation runs was chosen between 200 and 10000, depending on the SNR. The SNR is defined as $\frac{1}{N M_{\mathrm{R}}} \sum_{k=1}^{M_{\mathrm{R}}} \sum_{n=0}^{N-1} |x_k[n]|^2 / \sigma^2$.

$\cdots, \mathbf{c}_{K'}[i], \cdots, \mathbf{c}_K[i]\}$ ($i = 1, \cdots, I$), with $K' \leq K < M_{\mathrm{T}}$ and $\mathcal{C}_i \subseteq \tilde{\mathcal{C}}_i \subset \mathbb{C}^{M_{\mathrm{T}}}$, and we know that $\mathbf{G}$ maps *any* $\mathbf{v} \in \mathcal{C}_i$ to some $\tilde{\mathbf{v}} \in \tilde{\mathcal{C}}_i$, i.e.,

$$\mathbf{G}\mathcal{C}_i \subseteq \tilde{\mathcal{C}}_i\,, \quad i = 1, \cdots, I\,. \tag{12}$$

How should we construct the pairs of subspaces $\mathcal{C}_i, \tilde{\mathcal{C}}_i$ (or, equivalently, the basis $\{\mathbf{c}_1[i], \cdots, \mathbf{c}_{K'}[i], \cdots, \mathbf{c}_K[i]\}$), and how many such pairs do we need (i.e., how large must we choose I) such that (12) implies $\mathbf{G} = c\,\mathbf{I}$? Let $\tilde{\mathcal{C}}_i^{\perp} = \mathrm{span}\,\{\mathbf{b}_1[i], \cdots, \mathbf{b}_{M_{\mathrm{T}}-K}[i]\} \subset \mathbb{C}^{M_{\mathrm{T}}}$ denote the orthogonal complement space of $\tilde{\mathcal{C}}_i$. Since $\mathbf{c}_k[i] \in \tilde{\mathcal{C}}_i$ and $\mathbf{b}_k[i] \in \tilde{\mathcal{C}}_i^{\perp}$, there must be $\mathbf{c}_l[i] \perp \mathbf{b}_k[i]$, i.e.,

$$\mathbf{b}_k^H[i]\,\mathbf{c}_l[i] = 0\,, \quad k = 1, \cdots, M_{\mathrm{T}} - K\,,\ l = 1, \cdots, K\,, \tag{13}$$

for $i = 1, \cdots, I$. Moreover, since according to (12) $\mathbf{G}\mathbf{c}_k[i] \in \tilde{\mathcal{C}}_i$ for all $\mathbf{c}_k[i]$ with $k = 1, \cdots, K'$, we also have

$$\mathbf{b}_k^H[i]\,\mathbf{G}\,\mathbf{c}_l[i] = 0\,, \quad k = 1, \cdots, M_{\mathrm{T}} - K\,,\ l = 1, \cdots, K'\,, \tag{14}$$

for $i = 1, \cdots, I$. These are $(M_{\mathrm{T}} - K)\,K'I$ equations of the type $\mathbf{x}^H\mathbf{G}\mathbf{y} = 0$ that constrain $\mathbf{G}$. If we choose $\mathbf{x} = \mathbf{u}_m$ and $\mathbf{y} = \mathbf{u}_n$ with $m \neq n$, with $\mathbf{u}_m$ the mth unit vector of size M_{T}, then $\mathbf{x}^H\mathbf{y} = 0$ and $\mathbf{x}^H\mathbf{G}\mathbf{y} = 0$ becomes $G_{m,n} = 0$. Therefore, by choosing for $\mathbf{c}_l[i]$ and $\mathbf{b}_k[i]$ suitable unit vectors, we can force the off-diagonal elements of $\mathbf{G}$ to zero while satisfying (13).

More specifically, we start with $i = 1$, allocating unit vectors for all $\mathbf{c}_l[1]$ and $\mathbf{b}_k[1]$ such that $\mathbf{c}_l[1] \neq \mathbf{b}_k[1]$. Since there are M_{T} unit vectors, we can find $M_{\mathrm{T}}^2 - M_{\mathrm{T}}$ pairs of *different* unit vectors. There are $(M_{\mathrm{T}} - K)\,K'$ pairs $\mathbf{c}_l[1], \mathbf{b}_k[1]$ for which we must allocate pairs of different unit vectors. Since $(M_{\mathrm{T}} - K)\,K' < M_{\mathrm{T}}^2 - M_{\mathrm{T}}$ (due to $K' \leq M_{\mathrm{T}} - 1$), we can force $(M_{\mathrm{T}} - K)\,K'$ off-diagonal elements of $\mathbf{G}$ to zero. Note that by specifying $\mathbf{b}_1[1], \cdots, \mathbf{b}_{M_{\mathrm{T}}-K}[1]$ and $\mathbf{c}_1[1], \cdots, \mathbf{c}_{K'}[1]$, we have specified the first subspace pair $\mathcal{C}_1, \tilde{\mathcal{C}}_1$.

We now proceed to construct the second subspace pair $\mathcal{C}_2, \tilde{\mathcal{C}}_2$. We allocate "new" unit vector pairs (that have not been used previously) for vector pairs $\mathbf{c}_l[2], \mathbf{b}_k[2]$; these allow to force additional off-diagonal elements of $\mathbf{G}$ to zero. It may happen that less than $(M_{\mathrm{T}} - K)\,K'$ new unit vector pairs exist (however, there always exists at least one new unit vector pair). In that case, we have to allocate "old" unit vector pairs for some pairs $\mathbf{c}_l[2], \mathbf{b}_k[2]$ in order to span our subspaces $\mathcal{C}_2, \tilde{\mathcal{C}}_2$; these vector pairs do not serve to annihilate additional off-diagonal elements of $\mathbf{G}$.

We continue this construction for $i = 3$, $i = 4$, etc., until we have forced *all* $M_{\mathrm{T}}^2 - M_{\mathrm{T}}$ off-diagonal elements of $\mathbf{G}$ to zero, i.e., $\mathbf{G}$ is a *diagonal* matrix.

To have $\mathbf{G} = c\,\mathbf{I}$, it remains to constrain all diagonal elements $G_{m,m}$ to be equal. For this, we use additional subspace pairs $\mathcal{C}_i, \tilde{\mathcal{C}}_i$ and additional equations (14). Here, $M_{\mathrm{T}} - 1$ vector pairs $\mathbf{c}_l[i], \mathbf{b}_k[i]$ are constructed as follows: Two consecutive elements of $\mathbf{c}_l[i]$ are 1 and -1, with all other elements 0, and the corresponding two consecutive elements of $\mathbf{b}_k[i]$ are both 1, again with all other elements 0. Again, (13) is satisfied. Furthermore, the equations (14) simplify to $G_{m+1,m+1} = G_{m,m}$ for $m = 1, \cdots, M_{\mathrm{T}} - 1$, which means that all diagonal elements $G_{m,m}$ are equal. Thus, finally, $\mathbf{G} = c\,\mathbf{I}$ with some $c \in \mathbb{C}$. It can be shown that the number of subspace pairs required for the entire construction, I, is bounded as $\frac{M_{\mathrm{T}}^2 - 1}{(M_{\mathrm{T}} - K)\,K'} \leq I \leq M_{\mathrm{T}}^2 - 1$.

Having shown how to construct suitable pairs of subspaces to force $\mathbf{G} = c\,\mathbf{I}$, we now return to our original problem. Since (10) maps just individual vectors of each K'-dimensional space $\mathcal{C}_i$ (and not the entire subspace!) into $\tilde{\mathcal{C}}_i$, we have to choose K' linearly independent vectors for each given $\mathcal{C}_i$. Thus, $N = K'I$, which yields the lower bound of $N \geq \frac{M_{\mathrm{T}}^2 - 1}{M_{\mathrm{T}} - K}$ for the number of vectors. Moreover, it can be shown that N can be upper bounded as $N \leq M_{\mathrm{T}}^2 - 1$ (instead of $N \leq K'(M_{\mathrm{T}}^2 - 1)$).

The modulation matrices $\mathbf{M}_k$ can finally be constructed as follows. Since we need K' linearly independent vectors for each subspace pair, we can choose $\mathbf{m}_k[n] = \mathbf{c}_k[1]$ for $k = 1, \cdots, K$ and $n = 0, \cdots, K' - 1$, $\mathbf{m}_k[n] = \mathbf{c}_k[2]$ for $k = 1, \cdots, K$ and $n = K', \cdots, 2K' - 1$, etc., or equivalently $\mathbf{m}_k[n] = \mathbf{c}_k\left[\lfloor n/K' \rfloor + 1\right]$ for $k = 1, \cdots, K$ in general. Now, if any K' successive vectors $\mathbf{d}[n] \triangleq \left[d_1[n] \cdots d_{K'}[n]\right]^T$ with $n = iK', iK' + 1, \cdots, (i+1)K' - 1$ (i.e., starting at integer multiples of K') are linearly independent, the first K' vector constraints in (10) will span $\mathcal{C}_1$, the second K' vector constraints in (10) will span $\mathcal{C}_2$, etc., and thus all our I subspace constraints will be enforced.

ACKNOWLEDGMENT

The authors would like to thank Klaus List and Christian Stiegler for contributions to the proof of Theorem 1.

REFERENCES

[1] J. H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Sel. Areas Comm.*, vol. 5, pp. 871–878, June 1987.

[2] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," *U.S. Patent, no. 5,345,599*, 1994.

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Syst. Tech. J.*, pp. 41–59, 1996.

[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[5] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962–1973, 2000.

[6] J.-F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, pp. 2009–2025, Oct. 1998.

[7] A. Chevreuil and P. Loubaton, "MIMO blind second-order equalization method and conjugate cyclostationarity," *IEEE Trans. Signal Processing*, vol. 47, no. 2, pp. 572–578, 1999.

[8] H. Bölcskei, R. W. Heath Jr., and A. J. Paulraj, "Blind channel estimation in spatial multiplexing systems using nonredundant antenna precoding," in *Asilomar Conf. Signals, Systems, Computers*, (Pacific Grove (CA)), Oct. 1999.

[9] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Processing*, vol. 46, pp. 3003–3013, Nov. 1998.

[10] H. Artés, F. Hlawatsch, and G. Matz, "Efficient POCS algorithms for deterministic blind equalization of time-varying channels," in *Proc. IEEE GLOBECOM-2000*, (San Franscisco, CA), pp. 1031–1035, Nov. 2000.

[11] P. L. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, pp. 181–208, Feb. 1993.

[12] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array — Part I: Algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1184–1197, May 1996.

[13] A. J. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 173–190, Jan. 1997.

[14] H. Liu and G. Xu, "Closed form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 43, pp. 2714–2723, Nov. 1995.

[15] J. A. Cadzow, "Signal enhancement—A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 49–62, Jan. 1988.

[16] J. Laurila, K. Kopsa, R. Schürhuber, and E. Bonek, "Semi-blind separation and detection of co-channel signals," in *Proc. IEEE Int. Conf. Commun. (ICC)*, (Vancouver B.C., Canada), pp. 17–22, June 1999.