in Proc. IEEE ICASSP-2002, Orlando (FL)
May 2002, pp. 2225–2228.

SPACE-TIME MATRIX MODULATION: RANK-DEFICIENT CHANNELS AND MULTI-USER CASE

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 extend a previously introduced space-time modulation scheme (which we call *matrix modulation*) to the practically important cases of rank-deficient channels and multiple users. In particular, we prove an identifiability result and present an efficient demodulation method for these cases. Our matrix modulation technique does not require channel state information at the transmitter or receiver side. Its structure is strong enough to permit joint data and channel estimation without knowledge of the data alphabet. In the multi-user case, matrix modulation yields a flexible transmission scheme that can easily be adapted to different data-rate requirements of individual users. Numerical simulations demonstrate the good performance of the matrix modulation/demodulation technique.

1. INTRODUCTION

Multiple input/multiple output (MIMO) channels offer the possibility of high data rates (e.g. [1]) and/or high diversity gains (e.g. [2, 3]). Most papers assume that the receiver has perfect channel state information (e.g. [1–3]). Only recently has there been increased interest in the case where neither the receiver nor the transmitter possesses any knowledge about the channel (e.g. [4–6]).

In [7], we proposed a linear space-time modulation technique—which we called *matrix modulation*—that allows the receiver to jointly estimate the channel and demodulate the data without prior channel knowledge at the transmitter or at the receiver. We derived this modulation technique under the assumption that the receiver has at least as many antennas as the transmitter and that the MIMO channel has full rank.

In this paper, we first show that the structure of matrix modulation is strong enough to permit joint channel estimation and data demodulation also for the practically important case [8] of a rank-deficient channel. We then use this result to extend our matrix modulation technique to the case of multiple users (see Fig. 1). This extension is made possible by a close relationship between the rank-deficient single-user case and the multi-user case.

This paper is organized as follows. Section 2 reviews matrix modulation. In Section 3, the identifiability theorem and efficient demodulation algorithm from [7] are extended to the rank-deficient case. In Section 4, the matrix modulation technique, identifiability theorem, and demodulation algorithm are extended to the multi-user case. Finally, simulation results are provided in Section 5.

2. REVIEW OF MATRIX MODULATION

We first provide a brief review of matrix modulation for the single-user case [7]. For simplicity of exposition, we assume noiseless transmission; the effect of noise will be considered in Section 5.

The input-output relation of a flat-fading MIMO channel is

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

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 $\mathbf{H}$ (here, M_T and M_R are the numbers of transmit and receive antennas, respectively). The single-user transmission system is shown in Figure 1(b). We consider N_d input data streams $d_1[n], \cdots, d_{N_d}[n]$ with $d_l[n] \in \mathbb{C}$ (i.e., no finite-alphabet assumption is made).

Funding by FWF grant P12228-TEC.

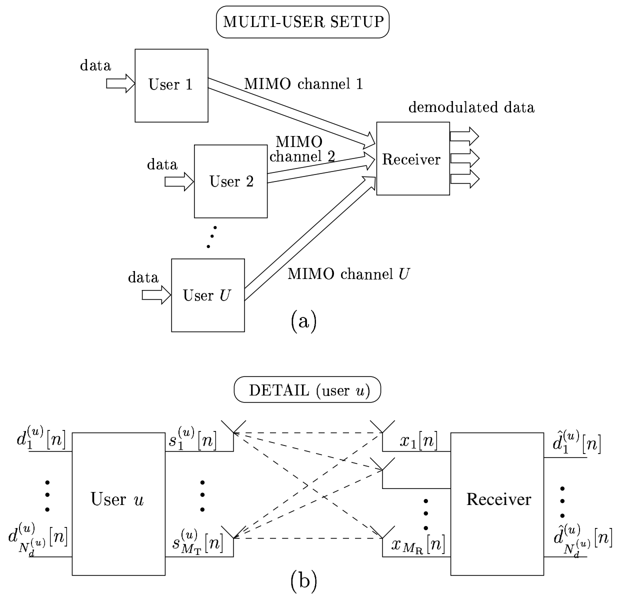


Fig. 1: Multi-user MIMO transmission. (a) Overall setup, (b) MIMO transmission system for user u.

The transmitter forces a "matrix 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

$$\mathbf{S} \,=\, \sum_{l=1}^{N_d} \mathbf{M}_l \mathbf{D}_l\,, \tag{2}$$

with the N_d diagonal $N \times N$ data matrices $\mathbf{D}_l \triangleq \mathrm{diag}\{d_l[0], \cdots, d_l[N-1]\}$ containing the data and with N_d *modulation matrices* $\mathbf{M}_l$ of size $M_\mathrm{T} \times N$. The modulation matrices $\mathbf{M}_l$ determine how the data is mapped (modulated) to the M_T transmit antennas; they are known to the receiver. 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

$$\mathbf{X} \,=\, \mathbf{H}\mathbf{S} \,=\, \mathbf{H}\sum_{l=1}^{N_d} \mathbf{M}_l \mathbf{D}_l\,. \tag{3}$$

From the received matrix $\mathbf{X}$, the receiver jointly estimates the data $d_l[n]$ and the unknown channel $\mathbf{H}$.

A set of modulation matrices $\{\mathbf{M}_1, \cdots, \mathbf{M}_{N_d}\}$ will be called *admissible* if the data sequences $d_l[n]$ can be uniquely reconstructed (up to a constant factor) from the received matrix $\mathbf{X}$, without knowledge of $\mathbf{H}$. In [7], we proved the existence of an admissible set of modulation matrices for the case where the channel $\mathbf{H}$ has full rank and the number of receive antennas is at least as large as the number of transmit antennas, i.e., $M_{\mathrm{R}} \geq M_{\mathrm{T}}$. Furthermore, an efficient iterative demodulation algorithm was proposed for that case.

3. THE RANK-DEFICIENT CASE

We shall now extend these results to the practically important case [8] of a rank-deficient channel matrix $\mathbf{H}$.

Unique Reconstruction. We first note that as far as unique reconstruction is concerned, the case $M_{\mathrm{R}} < M_{\mathrm{T}}$ can always be formulated as a rank-deficient channel, and thus it need not be considered separately. Indeed, when $M_{\mathrm{R}} < M_{\mathrm{T}}$, we can append a sufficient number of zero rows to $\mathbf{H}$ to obtain the *square* channel matrix $\mathbf{H}' = [\mathbf{H}^T, \mathbf{0}^T]^T$ with $\mathrm{rank}\{\mathbf{H}'\} = \mathrm{rank}\{\mathbf{H}\}$. Thus, $\mathbf{H}'$ is a rank-deficient channel with $M'_{\mathrm{R}} = M'_{\mathrm{T}}$. This shows that the case $M_{\mathrm{R}} < M_{\mathrm{T}}$ is equivalent to the case $M_{\mathrm{R}} \geq M_{\mathrm{T}}$ with rank-deficient channel matrix.

The following identifiability theorem states that under certain conditions (which allow rank-deficiency), there exists an admissible set of modulation matrices $\mathbf{M}_l$. More specifically, there exists a set of $\mathbf{M}_l$ such that the received matrix $\mathbf{X} = \mathbf{H}\sum_{l=1}^{N'_d}\mathbf{M}_l\mathbf{D}_l$ (with $N'_d \leq N_d$ since we allow some of the data streams to be zero) does not permit a different representation $\hat{\mathbf{H}}\sum_{l=1}^{N_d}\mathbf{M}_l\hat{\mathbf{D}}_l$ with $\hat{\mathbf{H}}$ and $\hat{\mathbf{D}}_l$ nontrivially different from $\mathbf{H}$ and $\mathbf{D}_l$, respectively. For this theorem, we model $\mathbf{H}$ as a realization of a random matrix. If $\mathbf{H}$ has full rank, its entries are assumed iid Gaussian; if $\mathbf{H}$ is rank-deficient, we model it as the product of a tall matrix $\mathbf{H}_1$ and a wide matrix $\mathbf{H}_2$, both with iid Gaussian entries [8].

Theorem 1. *Let $\mathbf{D}_l$ and $\hat{\mathbf{D}}_l$ be N'_d resp. N_d diagonal matrices of size $N \times N$, with $\mathbf{D}_l$ nonsingular, and let $\mathbf{H}$ and $\hat{\mathbf{H}}$ be matrices of size $M_{\mathrm{R}} \times M_{\mathrm{T}}$, with $\mathbf{H}$ modeled as explained above. Let $N'_d \leq N_d \leq \mathrm{rank}\{\mathbf{H}\} - 1$. Then there exist some N with $\left\lceil \frac{M_{\mathrm{T}}^2-1}{\mathrm{rank}\{\mathbf{H}\}-N_d} \right\rceil \leq N \leq M_{\mathrm{T}}^2 - 1$ and a set of N_d matrices $\mathbf{M}_l$ of size $M_{\mathrm{T}} \times N$ such that*

$$\hat{\mathbf{H}}\sum_{l=1}^{N_d}\mathbf{M}_l\hat{\mathbf{D}}_l \,=\, \mathbf{H}\sum_{l=1}^{N'_d}\mathbf{M}_l\mathbf{D}_l\,, \tag{4}$$

for one given set of data matrices $\mathbf{D}_l$, implies $\hat{\mathbf{H}} = c\mathbf{H}$ (with $c \in \mathbb{C}$ an unknown factor) and, with probability 1,

$$\hat{\mathbf{D}}_l \,=\, \begin{cases}(1/c)\mathbf{D}_l\,, & l \leq N'_d\\ \mathbf{0}\,, & N'_d\!+\!1 \leq l \leq N_d.\end{cases} \tag{5}$$

The proof of this theorem is sketched in Appendix A.

Demodulation Algorithm. Next, we propose an efficient iterative demodulation algorithm for a potentially rank-deficient channel. This algorithm is inspired by the demodulation algorithm for the full-rank case presented in [7].

Given a received matrix $\mathbf{X} = \mathbf{HS}$ and assuming admissible modulation matrices $\mathbf{M}_l$, Theorem 1 states that if the receiver is able to find matrices $\hat{\mathbf{H}}$ and $\hat{\mathbf{S}}$ that satisfy the two properties

1. $\hat{\mathbf{H}}\hat{\mathbf{S}} = \mathbf{X}$
2. $\hat{\mathbf{S}} = \sum_{l=1}^{N_d}\mathbf{M}_l\hat{\mathbf{D}}_l$ with $\hat{\mathbf{D}}_l$ diagonal,

then the matrices $\hat{\mathbf{D}}_l$ contain the correct data up to a constant factor. This motivates an iterative algorithm that consists in alternately executing two different steps. The ith iteration is as follows.

Step 1 enforces Property 1. That is, given $\hat{\mathbf{S}}_2^{(i-1)}$ as a result of Step 2 from the previous iteration (see below), we wish to find $\hat{\mathbf{H}}^{(i)}$ and $\hat{\mathbf{S}}_1^{(i)}$ such that $\hat{\mathbf{H}}^{(i)}\hat{\mathbf{S}}_1^{(i)}$ best approximates $\mathbf{X}$.

As a first substep, we calculate $\hat{\mathbf{H}}^{(i)}$ such that $\hat{\mathbf{H}}^{(i)}\hat{\mathbf{S}}_2^{(i-1)}$ best approximates $\mathbf{X}$ in the least-squares (LS) sense. This gives $\hat{\mathbf{H}}^{(i)} = \mathbf{X}\hat{\mathbf{S}}_2^{(i-1)\,\#}$, where $\hat{\mathbf{S}}_2^{(i-1)\,\#}$ denotes the pseudo-inverse of $\hat{\mathbf{S}}_2^{(i-1)}$. As a second substep, we calculate $\hat{\mathbf{S}}_1^{(i)}$ such that $\hat{\mathbf{H}}^{(i)}\hat{\mathbf{S}}_1^{(i)}$ best approximates $\mathbf{X}$ in the LS sense. This gives the final result

$$\hat{\mathbf{S}}_1^{(i)} \,=\, \hat{\mathbf{H}}^{(i)\,\#}\mathbf{X} \,=\, \big(\mathbf{X}\hat{\mathbf{S}}_2^{(i-1)\,\#}\big)^{\#}\mathbf{X}\,.$$

Step 2 enforces Property 2. That is, given $\hat{\mathbf{S}}_1^{(i)}$ from Step 1 above, we calculate an $\hat{\mathbf{S}}_2^{(i)}$ with modulation structure, i.e., $\hat{\mathbf{S}}_2^{(i)} = \sum_{l=1}^{N_d}\mathbf{M}_l\hat{\mathbf{D}}_l^{(i)}$, with the $\hat{\mathbf{D}}_l^{(i)}$ chosen such that $\hat{\mathbf{S}}_2^{(i)}$ best approximates $\hat{\mathbf{S}}_1^{(i)}$ in the LS sense. It can be shown that the nonzero (diagonal) elements of $\hat{\mathbf{D}}_l^{(i)}$ are given by

$$\big(\hat{\mathbf{D}}_l^{(i)}\big)_{n,n} \,=\, \frac{1}{M_{\mathrm{T}}}\sum_{k=1}^{M_{\mathrm{T}}}\big(\hat{\mathbf{S}}_1^{(i)}\big)_{k,n}\big(\mathbf{C}_l^{(i)}\big)_{k,n}\,. \tag{6}$$

Here, the matrices $\mathbf{C}_l^{(i)}$ of size $M_{\mathrm{T}} \times N$ are as follows. Let $\mathbf{m}_l[n]$ denote the nth column of $\mathbf{M}_l = \big[\mathbf{m}_l[0]\cdots\mathbf{m}_l[N-1]\big]$. Furthermore, let $\mathbf{M}[n]$ denote the $M_{\mathrm{T}} \times N_d$ matrix that contains the nth columns of all $\mathbf{M}_l$, i.e., $\mathbf{M}[n] \triangleq \big[\mathbf{m}_1[n]\cdots\mathbf{m}_{N_d}[n]\big]$. Then, the nth column of $\mathbf{C}_l^{(i)}$ equals the lth row of the $N_d \times M_{\mathrm{T}}$ matrix $\big(\hat{\mathbf{H}}^{(i)\,\#}\hat{\mathbf{H}}^{(i)}\mathbf{M}[n]\big)^{\#}$, where $\hat{\mathbf{H}}^{(i)}$ was calculated at Step 1 above.

This algorithm yields an estimate $\hat{\mathbf{H}}^{(i)}$ of the channel matrix in Step 1 and estimates $\hat{\mathbf{D}}_l^{(i)}$ of the data matrices in Step 2. In the noise-free case, we always observed the algorithm to converge to the correct channel and data matrices. Results in the presence of noise will be shown in Section 5.

4. THE MULTI-USER CASE

We will now extend our methods and results to the case of multiple users as illustrated in Fig. 1. There are U users, all of them equipped with M_{T} transmit antennas and transmitting simultaneously to a single receiver with M_{R} receive antennas. The channel's input-output relation is

$$\mathbf{X} \,=\, \sum_{u=1}^{U}\mathbf{H}^{(u)}\mathbf{S}^{(u)}, \tag{7}$$

where $\mathbf{H}^{(u)}$ and $\mathbf{S}^{(u)}$ are the (unknown) $M_{\mathrm{R}} \times M_{\mathrm{T}}$ channel matrix and the $M_{\mathrm{T}} \times N$ transmit matrix, respectively, associated with the uth user.

Extending the matrix modulation technique discussed above, we propose to construct the transmit matrices as

$$\mathbf{S}^{(u)} = \sum_{l=1}^{N_d^{(u)}}\mathbf{M}_l^{(u)}\mathbf{D}_l^{(u)}, \tag{8}$$

where, analogously to the single-user case, we consider $N_d^{(u)}$ input data streams $d_l^{(u)}[n]$ with associated $N \times N$ diagonal data matrices $\mathbf{D}_l^{(u)} \triangleq \mathrm{diag}\{d_l^{(u)}[0],\cdots,d_l^{(u)}[N-1]\}$ and $N_d^{(u)}$ modulation matrices $\mathbf{M}_l^{(u)}$ of size $M_{\mathrm{T}} \times N$. Inserting (8) in (7), we obtain the overall input-output relation (which extends (3))

$$\mathbf{X} \,=\, \sum_{u=1}^{U}\mathbf{H}^{(u)}\sum_{l=1}^{N_d^{(u)}}\mathbf{M}_l^{(u)}\mathbf{D}_l^{(u)}. \tag{9}$$

We will now reformulate (9) such that the interrelation between the multi-user case and the rank-deficient single-user case becomes apparent. Let $\widetilde{\mathbf{H}} \triangleq [\mathbf{H}^{(1)}\cdots\mathbf{H}^{(U)}]$ denote the overall channel matrix obtained by stacking the individual channel matrices of all users. Furthermore, let $\widetilde{\mathbf{M}}_l^{(u)} \triangleq [\mathbf{0}^T\cdots\mathbf{0}^T\,\mathbf{M}_l^{(u)\,T}\,\mathbf{0}^T\cdots\mathbf{0}^T]^T$ denote a "zero-padded" modulation matrix obtained by stacking $u-1$

zero matrices $\mathbf{0}$ of size $M_\mathrm{T} \times N$ above $\mathbf{M}_l^{(u)}$ and $U-u$ such zero matrices below $\mathbf{M}_l^{(u)}$. We can then rewrite (9) as

$$\mathbf{X} \,=\, \widetilde{\mathbf{H}} \sum_{u=1}^{U} \sum_{l=1}^{N_d^{(u)}} \widetilde{\mathbf{M}}_l^{(u)} \mathbf{D}_l^{(u)}. \tag{10}$$

Comparing with (3), we see that the multi-user case is equivalent to the single-user case with UM_T transmit antennas, M_R receive antennas, $N_U = \sum_{u=1}^{U} N_d^{(u)}$ data streams, a channel matrix $\widetilde{\mathbf{H}}$ of size $M_\mathrm{R} \times UM_\mathrm{T}$, and modulation matrices $\widetilde{\mathbf{M}}_l^{(u)}$ of size $UM_\mathrm{T} \times N$ that are nonzero only in the M_T rows with indices $uM_\mathrm{T}+1$ through $(u+1)M_\mathrm{T}$. There is $\mathrm{rank}\{\widetilde{\mathbf{H}}\} \le UM_\mathrm{T}$; typically, there will be $M_\mathrm{R} < UM_\mathrm{T}$ and thus $\mathrm{rank}\{\widetilde{\mathbf{H}}\} < UM_\mathrm{T}$ (rank-deficient case).

Unique Reconstruction. Based on this formulation, one can prove[1] the following identifiability theorem. The theorem states that under certain conditions, there exists an *admissible* set of modulation matrices $\mathbf{M}_l^{(u)}$ that allows the user data $d_l^{(u)}[n]$ to be uniquely recovered from the received matrix $\mathbf{X}$ up to unknown factors c_u (one factor per user).

Theorem 2. *Let $\mathbf{D}_l^{(u)}$ and $\hat{\mathbf{D}}_l^{(u)}$ be N'_U resp. N_U diagonal matrices of size $N \times N$, with $\mathbf{D}_l^{(u)}$ nonsingular, and let $\mathbf{H}^{(u)}$ and $\hat{\mathbf{H}}^{(u)}$ be matrices of size $M_\mathrm{R} \times M_\mathrm{T}$, with $\mathbf{H}^{(u)}$ modeled as the matrix $\mathbf{H}$ in Theorem 1. Let $N'_U \le N_U \le \mathrm{rank}\{\widetilde{\mathbf{H}}\} - 1$ and $N_d^{(u)\prime} \le N_d^{(u)} \le \mathrm{rank}\{\mathbf{H}^{(u)}\} - 1$. Then there exist some N with $\left\lceil \frac{U^2 M_\mathrm{T}^2 - U}{\mathrm{rank}\{\widetilde{\mathbf{H}}\} - N_U} \right\rceil \le N \le U^2 M_\mathrm{T}^2 - U$ and a set of N_U matrices $\mathbf{M}_l^{(u)}$ $(l = 1, \cdots, N_d^{(u)};\, u = 1, \cdots, U)$ of size $M_\mathrm{T} \times N$ such that*

$$\hat{\mathbf{H}}^{(u)} \sum_{l=1}^{N_d^{(u)}} \mathbf{M}_l^{(u)} \hat{\mathbf{D}}_l^{(u)} \,=\, \mathbf{H}^{(u)} \sum_{l=1}^{N_d^{(u)\prime}} \mathbf{M}_l^{(u)} \mathbf{D}_l^{(u)} \quad \forall u\,,$$

for one given set of data matrices $\mathbf{D}_l^{(u)}$, implies $\hat{\mathbf{H}}^{(u)} = c_u \mathbf{H}^{(u)}$ (where the $c_u \in \mathbb{C}$ are unknown factors) and, with probability 1,

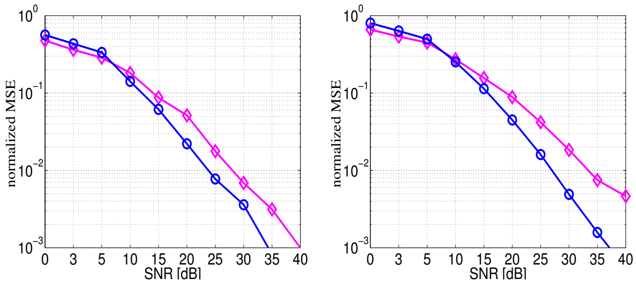
$$\hat{\mathbf{D}}_l^{(u)} \,=\, \begin{cases} (1/c_u)\mathbf{D}_l^{(u)}, & l \le N_d^{(u)\prime} \\ \mathbf{0}\,, & N_d^{(u)\prime}+1 \le l \le N_d^{(u)}. \end{cases}$$

It can furthermore be shown that a set of modulation matrices $\{\mathbf{M}_l^{(u)}\}$ with $u = 1, \cdots, U$ and $l = 1, \cdots, N_d^{(u)}$ that is admissible in the setting of Theorem 2 is also admissible for any single $M_\mathrm{R} \times M_\mathrm{T}$ matrix $\mathbf{H}$ (provided that the rank of the channel $\mathbf{H}$ is sufficient for the N_U data streams), and the same is true for an arbitrary subset of $\{\mathbf{M}_l^{(u)}\}$. This result provides the basis for a flexible multiple-access scheme. Indeed, it means that the modulation matrices of a given admissible set $\{\mathbf{M}_l^{(u)}\}$ can be arbitrarily assigned to the individual users according to their respective data-rate requirements. In the extreme case, all N_U modulation matrices $\mathbf{M}_l^{(u)}$ could be allocated to a single user.

Demodulation Algorithm. The iterative demodulation algorithm from Section 3 can be used also in the multi-user case. We merely have to replace $\hat{\mathbf{H}}$ with $\hat{\widetilde{\mathbf{H}}}$, $\mathbf{M}_l$ with $\widetilde{\mathbf{M}}_l^{(u)}$, and $\hat{\mathbf{D}}_l$ with $\hat{\mathbf{D}}_l^{(u)}$.

5. SIMULATION RESULTS

We conducted two experiments in which we transmitted randomly generated iid Gaussian data signals[2] $d_l[n] \in \mathbb{C}$. The modulation matrices $\mathbf{M}_l^{(u)}$ with block length $N = 200$ were constructed by taking realizations of iid Gaussian random variables as matrix entries and then orthonormalizing the corresponding columns of all $\mathbf{M}_l$ (this always worked in our simulations). The channels 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 $N = 200$.



(a) (b)

Fig. 2: Normalized MSE vs. SNR obtained with matrix modulation using the iterative demodulation algorithm. (a) One data stream, $M_\mathrm{T} = 2$ (◇—◇) or $M_\mathrm{T} = 4$ (○—○) transmit antennas, and $M_\mathrm{R} = 2$ receive antennas; (b) three data streams assigned to a single user (◇—◇) or to three different users (○—○).

In our first experiment, we studied the impact of the number of transmit antennas (i.e., available transmit diversity) in the single-user case. For one data stream ($N_d = 1$) and $M_\mathrm{R} = 2$ receive antennas, we compared the performance obtained with the iterative algorithm from Section 3 for $M_\mathrm{T} = 2$ and $M_\mathrm{T} = 4$ transmit antennas. Fig. 2(a) shows the normalized MSE vs. the SNR[3]. As expected, the additional two antennas lead to better performance.

In our second experiment, we compared the performance obtained in the single-user and multi-user cases. In the single-user case, we transmitted $N_d = 3$ data streams over a channel with $M_\mathrm{T} = 4$ transmit antennas and $M_\mathrm{R} = 6$ receive antennas. In the multi-user case, we assigned each of the $U = 3$ users one data stream (thus $N_U = 3$) which was transmitted by $M_\mathrm{T} = 4$ antennas per user and received by $M_\mathrm{R} = 6$ antennas. Fig. 2(b) shows that for high SNR, we obtain significantly better performance when the data streams are assigned to different users than when one user transmits all data streams. This is true even though we have to estimate three unknown channels in the multi-user case as opposed to only one unknown channel in the single-user case.

6. CONCLUSION

We demonstrated that space-time matrix modulation allows to transmit several data streams from multiple users over *unknown*, possibly rank-deficient MIMO channels. An efficient iterative receiver algorithm exploits the matrix modulation structure to jointly estimate the multi-user data and channels. Unique demodulation in the noise-free case is guaranteed theoretically. The good performance of the space-time matrix modulation/demodulation technique in the presence of noise was demonstrated via numerical simulations.

APPENDIX A: SKETCH OF PROOF OF THEOREM 1

Let $r \triangleq \mathrm{rank}\{\mathbf{H}\}$. We can always factor $\mathbf{H}$ as $\mathbf{H} = \mathbf{U}\mathbf{V}$, with full-rank matrices $\mathbf{U}$ and $\mathbf{V}$ of size $M_\mathrm{R} \times r$ and $r \times M_\mathrm{T}$,

[1]The proof is largely analogous to that of Theorem 1 and will be omitted due to space restrictions. We only remark that as a consequence of the special structure of $\widetilde{\mathbf{M}}_l^{(u)}$, we now have an unknown factor c_u per user.

[2]We do not use a finite source alphabet in order to demonstrate that the matrix modulation/demodulation technique works for arbitrary alphabets. A finite-alphabet case is considered in [9].

[3]In the single-user case, the normalized MSE is defined as $\sum_{l=1}^{N_d} \sum_n \left| d_l[n] - \hat{c}\hat{d}_l[n] \right|^2 / \sum_{l=1}^{N_d} \sum_n |d_l[n]|^2$ averaged over all simulation runs, where $\hat{d}_l[n]$ is the estimate of $d_l[n]$ and $\hat{c}$ is the least-squares fit for the unknown factor c. For the multi-user case, the normalized MSE is defined analogously. The number of simulation runs was chosen between 200 and 4000, depending on the SNR. The SNR is defined as $\frac{1}{NM_\mathrm{R}} \sum_{l=1}^{M_\mathrm{R}} \sum_{n=0}^{N-1} |x_l[n]|^2/\sigma^2$. The same SNR was used for each simulation run.

respectively. Let $\mathbf{P_U} \triangleq \mathbf{U}\mathbf{U}^{\#}$ denote the orthogonal projector on $\mathrm{colspan}\{\mathbf{U}\}$, the column span of $\mathbf{U}$. We can decompose $\hat{\mathbf{H}}$ as $\hat{\mathbf{H}} = \hat{\mathbf{H}}_\mathbf{U} + \hat{\mathbf{H}}_\mathbf{U}^\perp$ with $\hat{\mathbf{H}}_\mathbf{U} \triangleq \mathbf{P_U}\hat{\mathbf{H}}$ and $\mathbf{P_U}\hat{\mathbf{H}}_\mathbf{U}^\perp = \mathbf{0}$.

Statement 1: Equation (4) remains unchanged when we replace $\hat{\mathbf{H}}$ by $\hat{\mathbf{H}}_\mathbf{U}$, i.e., (4) is equivalent to

$$\hat{\mathbf{H}}_\mathbf{U} \sum_{l=1}^{N_d} \mathbf{M}_l \hat{\mathbf{D}}_l \,=\, \mathbf{HS}\,. \tag{11}$$

Proof. With $\mathbf{H} = \mathbf{UV}$ and $\hat{\mathbf{H}} = \hat{\mathbf{H}}_\mathbf{U} + \hat{\mathbf{H}}_\mathbf{U}^\perp$, (4) becomes

$$(\hat{\mathbf{H}}_\mathbf{U} + \hat{\mathbf{H}}_\mathbf{U}^\perp) \sum_{l=1}^{N_d} \mathbf{M}_l \hat{\mathbf{D}}_l \,=\, \mathbf{UVS}\,. \tag{12}$$

Multiplying (12) by $\mathbf{P_U}$ from the left does not change the right-hand side of (12) since $\mathbf{P_U U} = \mathbf{U}$. Hence, also the left-hand side is invariant to multiplication by $\mathbf{P_U}$. This shows that

$$\hat{\mathbf{H}}_\mathbf{U}^\perp \sum_{l=1}^{N_d} \mathbf{M}_l \hat{\mathbf{D}}_l \,=\, \mathbf{0}\,. \tag{13}$$

□

Statement 2: The matrix $\hat{\mathbf{H}}_\mathbf{U}$ can be written as $\hat{\mathbf{H}}_\mathbf{U} = \mathbf{U}\hat{\mathbf{V}}$ with $\mathbf{U}$ as above and $\hat{\mathbf{V}}$ some full-rank matrix of size $r \times M_\mathrm{T}$.

Proof. Since $\hat{\mathbf{H}}_\mathbf{U} = \mathbf{P_U}\hat{\mathbf{H}}$, we have $\hat{\mathbf{H}}_\mathbf{U} = \mathbf{U}\hat{\mathbf{V}}$ with some $\hat{\mathbf{V}}$ of size $r \times M_\mathrm{T}$. It remains to be proved that $\hat{\mathbf{V}}$ has full rank. It can be shown that a necessary condition for admissible $\mathbf{M}_l$ is that $\mathbf{S} = \sum_{l=1}^{N_d'} \mathbf{M}_l \mathbf{D}_l$ has full rank (this can always be satisfied). Then, $\mathrm{rank}\{\mathbf{HS}\} = \mathrm{rank}\{\mathbf{H}\} = r$ so that for (11) to hold there must be $\mathrm{rank}\{\hat{\mathbf{H}}_\mathbf{U}\} = r$. With $\hat{\mathbf{H}}_\mathbf{U} = \mathbf{U}\hat{\mathbf{V}}$, it then follows that $\mathrm{rank}\{\hat{\mathbf{V}}\} = r$. □

Next, we rewrite (11) column by column as

$$\hat{\mathbf{H}}_\mathbf{U} \tilde{\mathbf{v}}_n = \mathbf{H}\mathbf{v}_n\,, \quad n = 0, \cdots, N-1\,, \tag{14}$$

with

$$\mathbf{v}_n \triangleq \sum_{l=1}^{N_d'} d_l[n]\, \mathbf{m}_l[n]\,, \qquad \tilde{\mathbf{v}}_n \triangleq \sum_{l=1}^{N_d} \hat{d}_l[n]\, \mathbf{m}_l[n]\,, \tag{15}$$

where $\mathbf{m}_l[n]$ denotes the nth column of $\mathbf{M}_l$.

Statement 3: Equation (14) is equivalent to

$$\hat{\mathbf{V}}\tilde{\mathbf{v}}_n = \hat{\mathbf{V}}\mathbf{A}\mathbf{v}_n\,, \quad n = 0, \cdots, N-1\,, \tag{16}$$

with $\hat{\mathbf{V}}$ as above and some matrix $\mathbf{A}$ of size $M_\mathrm{T} \times M_\mathrm{T}$.

Proof. With $\hat{\mathbf{H}}_\mathbf{U} = \mathbf{U}\hat{\mathbf{V}}$ and $\mathbf{H} = \mathbf{UV}$, (14) reads $\mathbf{U}\hat{\mathbf{V}}\tilde{\mathbf{v}}_n = \mathbf{UV}\mathbf{v}_n$. Since $\hat{\mathbf{V}}$ and $\mathbf{V}$ have the same size and full rank, $\mathbf{V} = \hat{\mathbf{V}}\mathbf{A}$ with some $\mathbf{A}$. Thus, (14) becomes $\mathbf{U}\hat{\mathbf{V}}\tilde{\mathbf{v}}_n = \mathbf{U}\hat{\mathbf{V}}\mathbf{A}\mathbf{v}_n$. Since $\mathbf{U}$ is left invertible, this is equivalent to $\hat{\mathbf{V}}\tilde{\mathbf{v}}_n = \hat{\mathbf{V}}\mathbf{A}\mathbf{v}_n$. □

Statement 4: We assume that the vectors $\mathbf{m}_1[n], \cdots, \mathbf{m}_{N_d}[n]$ are linearly independent for n fixed. If $\hat{\mathbf{H}}_\mathbf{U} = c\mathbf{H}$, then (14) with (15) implies (5), i.e., with probability 1 there is $\hat{d}_l[n] = \frac{1}{c}\, d_l[n]$ for $1 \le l \le N_d'$ and $\hat{d}_l[n] = 0$ for $N_d' + 1 \le l \le N_d$.

Proof. $\hat{\mathbf{H}}_\mathbf{U} = c\mathbf{H}$ together with (14) is equivalent to

$$c\mathbf{H}\tilde{\mathbf{v}}_n = \mathbf{H}\mathbf{v}_n\,. \tag{17}$$

Now $\mathbf{H}$ is a realization of a random matrix as discussed in Section 3. Therefore, with probability 1, the vectors $\mathbf{Hm}_1[n], \cdots, \mathbf{Hm}_{N_d}[n]$ will be linearly independent (note that $N_d' \le N_d \le r-1$). Thus, with probability 1, there is a one-to-one correspondence between $\mathbf{H}\mathbf{v}_n = \sum_{l=1}^{N_d'} d_l[n]\, \mathbf{Hm}_l[n]$ and the $d_l[n]$ and similarly between $\mathbf{H}\tilde{\mathbf{v}}_n = \sum_{l=1}^{N_d} \hat{d}_l[n]\, \mathbf{Hm}_l[n]$ and the $\hat{d}_l[n]$. Therefore, it follows from (17) that with probability 1, we have $\hat{d}_l[n] = \frac{1}{c}\, d_l[n]$ for $1 \le l \le N_d'$ and $\hat{d}_l[n] = 0$ for $N_d' + 1 \le l \le N_d$. □

Statement 5: Let the matrices $\mathbf{M}_l$ be such that for an arbitrary $M_\mathrm{T} \times M_\mathrm{T}$ matrix $\mathbf{A}$, the equation $\tilde{\mathbf{v}}_n - \mathbf{A}\mathbf{v}_n = \mathbf{0}$ with $\mathbf{v}_n$, $\tilde{\mathbf{v}}_n$ as in (15) implies $\mathbf{A} = c'\mathbf{I}$ with some factor c' (as shown in [7], such $\mathbf{M}_l$ can always be found). Then $\{\mathbf{M}_l\}$ is an admissible set for the class of (possibly rank-deficient) $M_\mathrm{R} \times M_\mathrm{T}$ channels $\mathbf{H}$ with $\mathrm{rank}\{\mathbf{H}\} \ge N_d + 1$.

Proof. We can rewrite (16) as

$$\hat{\mathbf{V}}\,(\tilde{\mathbf{v}}_n - \mathbf{A}\mathbf{v}_n) = \mathbf{0}\,, \quad n = 0, \cdots, N-1\,. \tag{18}$$

This shows that either $\tilde{\mathbf{v}}_n - \mathbf{A}\mathbf{v}_n = \mathbf{0}$ or $\tilde{\mathbf{v}}_n - \mathbf{A}\mathbf{v}_n$ is orthogonal to the row span of $\hat{\mathbf{V}}$. With the assumptions of Statement 5, it can then be shown (similarly to [7]) that for (18) to hold, there must be $\mathbf{A} = c'\mathbf{I} + \mathbf{T}$ with an $M_\mathrm{T} \times M_\mathrm{T}$ matrix $\mathbf{T}$ satisfying $\hat{\mathbf{V}}\mathbf{T} = \mathbf{0}$. According to the proof of Statement 3, there is $\mathbf{V} = \hat{\mathbf{V}}\mathbf{A}$ and thus we obtain $\mathbf{V} = \hat{\mathbf{V}}(c'\mathbf{I} + \mathbf{T}) = c'\hat{\mathbf{V}}$ or equivalently $\hat{\mathbf{V}} = c\mathbf{V}$. Inserting in $\hat{\mathbf{H}}_\mathbf{U} = \mathbf{U}\hat{\mathbf{V}}$ (see Statement 2) yields $\hat{\mathbf{H}}_\mathbf{U} = c\mathbf{UV} = c\mathbf{H}$. Using Statement 4, it finally follows that $\{\mathbf{M}_l\}$ is admissible. □

Thus, we have shown that we can find an admissible set $\{\mathbf{M}_l\}$, i.e., a set that allows to uniquely reconstruct the data $d_l[n]$ under the conditions of Theorem 1.

We now prove the second claim in Theorem 1, namely, $\hat{\mathbf{H}} = c\mathbf{H}$. We already showed that $\hat{d}_l[n] = \frac{1}{c}\, d_l[n]$ for $l \le N_d'$ and $\hat{d}_l[n] = 0$ for $N_d' + 1 \le l \le N_d$ (unique reconstruction). Inserting into (13), we obtain $\mathbf{0} = \hat{\mathbf{H}}_\mathbf{U}^\perp \sum_{l=1}^{N_d} \mathbf{M}_l \hat{\mathbf{D}}_l = \frac{1}{c}\hat{\mathbf{H}}_\mathbf{U}^\perp \sum_{l=1}^{N_d'} \mathbf{M}_l \mathbf{D}_l = \frac{1}{c}\hat{\mathbf{H}}_\mathbf{U}^\perp \mathbf{S}$. Since $\mathbf{S}$ has full rank, it follows that $\hat{\mathbf{H}}_\mathbf{U}^\perp = \mathbf{0}$. Inserting this into $\hat{\mathbf{H}} = \hat{\mathbf{H}}_\mathbf{U} + \hat{\mathbf{H}}_\mathbf{U}^\perp$, we obtain (cf. the proof of Statement 5) $\hat{\mathbf{H}} = \hat{\mathbf{H}}_\mathbf{U} = c\mathbf{H}$.

The proof that $\left\lceil \frac{M_\mathrm{T}^2 - 1}{r - N_d} \right\rceil \le N \le M_\mathrm{T}^2 - 1$ will not be included because of space restrictions.

REFERENCES

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

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, March 1998.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Comm.*, vol. 16, pp. 1451–1458, Oct. 1998.

[4] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, pp. 543–564, March 2000.

[5] B. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Comm.*, vol. 48, pp. 2041–2052, Dec. 2000.

[6] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[7] H. Artés and F. Hlawatsch, "Blind equalization of MIMO channels using deterministic precoding," in *Proc. IEEE ICASSP-2001*, (Salt Lake City, UT), pp. 2153–2156, May 2001.

[8] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Comm.*, July 2000. submitted.

[9] H. Artés, G. Matz, D. Schafhuber, and F. Hlawatsch, "Space-time multicarrier matrix modulation for unknown dispersive MIMO channels," in *Proc. 39th Allerton Conf. on Commun., Contr., Comput.*, (Urbana, IL), Oct. 2001.