in Proc. IEEE ICASSP-03, Hong Kong, April 2003, vol. IV, pp. 41–44

SPACE-TIME MATRIX MODULATION: EXTENSION TO UNKNOWN DOUBLY SELECTIVE MIMO CHANNELS

Harold Artés and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria (Europe)
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

Most transmission schemes for MIMO channels assume (i) a block fading channel without delay spread and (ii) availability of channel state information at the receiver. Here, we extend the *space-time matrix modulation (STMM)* scheme and the iterative demodulation algorithm that we introduced previously to *unknown, doubly selective* MIMO channels, i.e., delay/Doppler-spread MIMO channels that are unknown to the receiver. We show that the structure inherent in STMM allows perfect reconstruction of the data when transmitting over an unknown doubly selective channel (apparently, this is not currently possible with other transmission schemes). Numerical simulations demonstrate significant performance advantages of STMM over Cayley differential unitary space-time modulation.

1. INTRODUCTION

Background and Motivation. Most transmission schemes for multi-input/multi-output (MIMO) channels assume flat or frequency-selective fading channels and availability of channel state information at the receiver. Only recently, methods have been developed for the case where neither the receiver nor the transmitter possesses any knowledge about the channel (e.g. [1–4]). An example is the *space-time matrix modulation (STMM)* scheme that we proposed in [4–6]. These methods are especially interesting in the case of low SNR that may occur, e.g., when many users are present whose interference can approximately be modeled as white noise.[1] Typically, these methods (including STMM) are formulated for a block fading channel that is constant over one block but allowed to change from block to block.

In this paper, we extend our STMM scheme to *doubly selective* MIMO channels, i.e., delay-spread and Doppler-spread MIMO channels that are time-varying even within a block. The motivation for doing so is twofold:

- Modeling the time-varying fading within blocks leads to an additional source of diversity, namely, *Doppler diversity* [8].
- The conventional block fading model severely restricts the block length in the case of fast fading channels. A smaller block length may be a disadvantage for code design, and it usually implies that the channel has to be estimated more frequently. Furthermore, explicitly modeling the channel's time variations within a longer block typically requires fewer parameters than using individual time-invariant channel models for several shorter blocks.

Main Results and Organization of Paper. Our paper contains two main contributions. First, we present an identifiability (or perfect reconstruction) result stating that the structure of STMM is strong enough to permit joint channel estimation and data detection for doubly selective, unknown MIMO channels. (This result actually applies to all linear space-time codes that are *separable* in that the coding over space and the coding over time are done independently.) Second, we present an iterative demodulation algorithm that performs joint channel estimation and data detection.

Funding by FWF grant P15156.

[1] Indeed, it has been shown (e.g. [7]) that in the low SNR case pilot symbol based channel estimation is highly suboptimal.

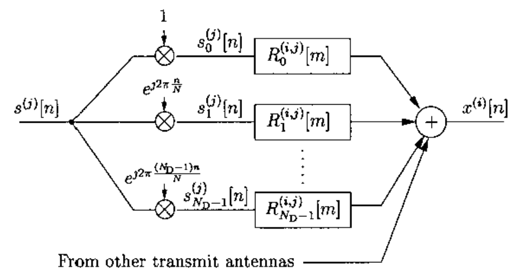


Fig. 1: Multichannel LTI representation of the (i,j)th LTV channel.

The paper is organized as follows. Section 2 introduces a model for the doubly selective MIMO channel. In Section 3, we show how to combine the STMM structure with the Doppler structure of the time-varying channel. Identifiability (or perfect reconstruction) results and an iterative demodulation algorithm for doubly selective channels are presented in Section 4 and 5, respectively. Finally, simulation results provided in Section 6 demonstrate the good performance of STMM and, specifically, significant performance advantages over Cayley differential unitary space-time modulation [9].

2. TIME-VARYING MIMO CHANNEL MODEL

We first present a model for the doubly selective MIMO channel that will be used in what follows. The discrete-time baseband signal received at the ith receive antenna is given by

$$x^{(i)}[n] = \sum_{j=1}^{M_T} \sum_{m=0}^{L-1} h^{(i,j)}[n,m]\, s^{(j)}[n-m]\,, \quad i = 1, 2, \cdots, M_R\,, \tag{1}$$

for $n = 0, 1, \cdots, N-1$ (i.e., the $x^{(i)}[n]$ are observed over the interval $[0, N-1]$). Here, $h^{(i,j)}[n,m]$ denotes the impulse response of the linear, time-varying (LTV) single-input/single-output (SISO) channel that maps the signal at the jth transmit antenna, $s^{(j)}[n]$, into $x^{(i)}[n]$; $L-1$ is the maximum time delay; and M_T and M_R are the numbers of transmit and receive antennas, respectively.

Multichannel LTI Representation. Using the *(delay-Doppler) spreading function* $R^{(i,j)}[m,l] \triangleq \sum_{n=0}^{N-1} h^{(i,j)}[n,m] e^{-j2\pi\frac{ln}{N}}$ [10] (with $j \triangleq \sqrt{-1}$) and denoting the maximum Doppler shift by $N_D - 1$, the input-output relation (1) can be rewritten as (cf. [11–13])

$$x^{(i)}[n] = \sum_{j=1}^{M_T} \sum_{l=0}^{N_D-1} \sum_{m=0}^{L-1} R_l^{(i,j)}[m]\, s_l^{(j)}[n-m]\,, \tag{2}$$

with $R_l^{(i,j)}[m] \triangleq R^{(i,j)}[m,l]\, e^{j2\pi\frac{lm}{N}}$ and $s_l^{(j)}[n] \triangleq s^{(j)}[n]\, e^{j2\pi\frac{ln}{N}}$. This expression, which is illustrated in Fig. 1, corresponds to a *multichannel LTI representation* of the LTV channel where each subchannel consists of a modulator (Doppler shift) and an LTI filter.

We conclude that practically arbitrary channels—including channels with diffuse scattering—are characterized by a finite set of LTI filters associated with *uniformly spaced*, discrete Doppler shifts. Formally, this representation is equivalent to the basis expansion models of [14, 15] using a basis of complex exponentials with uniformly spaced frequencies. Note that all SISO channels are assumed to have the same parameters L and N_{D}; this assumption is reasonable if the transmit antennas and, similarly, the receive antennas are located sufficiently close to each other.

In what follows, we suppose that P of the N_{D} subchannels in Fig. 1, corresponding to specific Doppler shifts $l_p \in [0, N_{\mathrm{D}}-1]$ with $p = 1, 2, \cdots, P$, are active. That is, only the subchannel impulse responses $R_{l_p}^{(i,j)}[m]$ are nonzero. (This is no restriction since we allow $P = N_{\mathrm{D}}$.) Thus, (2) becomes

$$x^{(i)}[n] = \sum_{j=1}^{M_{\mathrm{T}}} \sum_{p=1}^{P} \sum_{m=0}^{L-1} R_{l_p}^{(i,j)}[m]\, s_{l_p}^{(j)}[n-m]. \tag{3}$$

For channels satisfying the *wide-sense stationary uncorrelated scattering* (WSSUS) assumption, the "active Doppler shifts" l_p can be deduced from the channel's scattering function [10, 16]. Because the scattering function of a WSSUS channel does not change with time, it is much easier to estimate than the channel itself [17, 18]. Hereafter, we assume that the l_p are known to the receiver.

Matrix Formulation. For a compact formulation of (3), we define

$$\mathbf{R}^{(j)}[m] \triangleq \begin{bmatrix} R_{l_1}^{(1,j)}[m] & \cdots & R_{l_P}^{(1,j)}[m] \\ \vdots & & \vdots \\ R_{l_1}^{(M_{\mathrm{R}},j)}[m] & \cdots & R_{l_P}^{(M_{\mathrm{R}},j)}[m] \end{bmatrix},$$

and furthermore $\mathbf{R}[m] \triangleq [\mathbf{R}^{(1)}[m] \cdots \mathbf{R}^{(M_{\mathrm{T}})}[m]]$ and $\mathbf{R} \triangleq [\mathbf{R}[0] \cdots \mathbf{R}[L-1]]$. We also define the vector of modulated input samples $\mathbf{s}[n] \triangleq [s_{l_1}^{(1)}[n] \cdots s_{l_P}^{(1)}[n] \cdots s_{l_1}^{(M_{\mathrm{T}})}[n] \cdots s_{l_P}^{(M_{\mathrm{T}})}[n]]^T$ and the following block-Toeplitz input matrix of size $M_{\mathrm{T}}PL \times N$,

$$\mathbf{S} \triangleq \begin{bmatrix} \mathbf{s}[0] & \mathbf{s}[1] & \cdots & \mathbf{s}[N-1] \\ \mathbf{s}[-1] & \mathbf{s}[0] & \cdots & \mathbf{s}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{s}[-L+1] & \mathbf{s}[-L+2] & \cdots & \mathbf{s}[N-L+1] \end{bmatrix}.$$

Finally, we define the output vector $\mathbf{x}[n] \triangleq [x_1[n] \cdots x_{M_{\mathrm{T}}}[n]]^T$ and the output matrix $\mathbf{X} \triangleq [\mathbf{x}[0] \cdots \mathbf{x}[N-1]]$ of size $M_{\mathrm{R}} \times N$. Now (3) can be written as

$$\mathbf{X} = \mathbf{R}\mathbf{S}. \tag{4}$$

For later use, we define the "generator matrix" of $\mathbf{S}$ as the $M_{\mathrm{T}}P \times (N+L-1)$ matrix $\mathbf{S}_g \triangleq [\mathbf{s}[-L+1]\ \mathbf{s}[-L+2] \cdots \mathbf{s}[N-1]]$. Furthermore we define an operator $\mathcal{T}_L\{\cdot\}$ that maps the generator matrix to the corresponding block-Toeplitz matrix with L block-rows, i.e., $\mathcal{T}_L\{\mathbf{S}_g\} = \mathbf{S}$.

3. COMBINING THE STRUCTURES

As a basis for Section 4, we now show how the structure of STMM and the Doppler structure of the LTV channel model discussed in the previous section can be combined. We start by briefly reviewing the STMM scheme [4].

Review of STMM. We consider K input data streams $d_1[n], \cdots, d_K[n]$ with $d_k[n] \in \mathbb{C}$ (i.e., *no finite-alphabet assumption is made*). These data streams are mapped to the M_{T} transmit antennas as

$$\mathbf{s}_t[n] = \sum_{k=1}^{K} \mathbf{m}_k[n]\, d_k[n], \tag{5}$$

with the $M_{\mathrm{T}} \times 1$ vectors $\mathbf{s}_t[n] \triangleq [s^{(1)}[n] \cdots s^{(M_{\mathrm{T}})}[n]]^T$ and time-varying "modulation vectors" $\mathbf{m}_k[n]$ that are known to the receiver.

Defining the transmit signal matrix $\mathbf{S}_t \triangleq [\mathbf{s}_t[-L+1] \cdots \mathbf{s}_t[N-1]]$ of size $M_{\mathrm{T}} \times (N+L-1)$, the K *modulation matrices* $\mathbf{M}_k \triangleq [\mathbf{m}_k[-L+1] \cdots \mathbf{m}_k[N-1]]$ of size $M_{\mathrm{T}} \times (N+L-1)$, and the K diagonal data matrices $\mathbf{D}_k \triangleq \operatorname{diag}\{d_k[-L+1], \cdots, d_k[N-1]\}$ of size $(N+L-1) \times (N+L-1)$, we can formulate (5) as

$$\mathbf{S}_t = \sum_{k=1}^{K} \mathbf{M}_k \mathbf{D}_k. \tag{6}$$

A set of modulation matrices $\{\mathbf{M}_k\}$ will be called *admissible* if the data sequences $d_k[n]$ can be uniquely reconstructed (up to a common constant factor) from the received matrix $\mathbf{X}$ in (4), without knowledge of $\mathbf{R}$. In [5], we showed that for a flat fading MIMO channel $\mathbf{H}$ with $\operatorname{rank}\{\mathbf{H}\} > 1$, admissible sets of modulation matrices can always be found if $K < \operatorname{rank}\{\mathbf{H}\}$.

Combining the Structures. The Doppler structure of the channel (3) is expressed by the relation $s_l^{(j)}[n] = s^{(j)}[n]\, e^{j2\pi \frac{ln}{N}}$. We can rewrite this Doppler structure in a way similar to (5):

$$\mathbf{s}^{(j)}[n] = \mathbf{f}[n]\, s^{(j)}[n],$$

with $\mathbf{s}^{(j)}[n] \triangleq [s_{l_1}^{(j)}[n] \cdots s_{l_P}^{(j)}[n]]^T$ and $\mathbf{f}[n] \triangleq [e^{j2\pi \frac{l_1 n}{N}} \cdots e^{j2\pi \frac{l_P n}{N}}]^T$. Thus, the vectors $\mathbf{s}[n] = [\mathbf{s}^{(1)T}[n] \cdots \mathbf{s}^{(M_{\mathrm{T}})T}[n]]^T$ are given by

$$\mathbf{s}[n] = \mathbf{s}_t[n] \otimes \mathbf{f}[n],$$

where $\otimes$ denotes the Kronecker product [19]. Inserting (5) for $\mathbf{s}_t[n]$, we can finally write

$$\mathbf{s}[n] = \sum_{k=1}^{K} \tilde{\mathbf{m}}_k[n]\, d_k[n], \quad \text{with} \quad \tilde{\mathbf{m}}_k[n] \triangleq \mathbf{m}_k[n] \otimes \mathbf{f}[n]. \tag{7}$$

Equivalently, the generator matrix $\mathbf{S}_g = [\mathbf{s}[-L+1] \cdots \mathbf{s}[N-1]]$ becomes

$$\mathbf{S}_g = \sum_{k=1}^{K} \widetilde{\mathbf{M}}_k \mathbf{D}_k, \tag{8}$$

with the "Doppler-spread modulation matrices" $\widetilde{\mathbf{M}}_k \triangleq [\tilde{\mathbf{m}}_k[-L+1] \cdots \tilde{\mathbf{m}}_k[N-1]]$ of size $M_{\mathrm{T}}P \times (N+L-1)$.

The structure of (7), (8) equals that of (5), (6), which shows that the channel's Doppler structure nicely blends into the STMM structure. We finally note that $\widetilde{\mathbf{M}}_k = \mathbf{M}_k \odot \mathbf{F}$ and $\mathbf{S}_g = \mathbf{S}_t \odot \mathbf{F}$, where $\odot$ denotes the Khatri-Rao product [19] and $\mathbf{F} \triangleq [\mathbf{f}[-L+1] \cdots \mathbf{f}[N-1]]$.

4. PERFECT RECONSTRUCTION

First Reconstruction Result. Using methods from deterministic blind equalization (e.g. [13, 20]), the structure of the doubly selective channel can be exploited to equalize the channel up to an unknown instantaneous-mixture matrix. This matrix ambiguity can in turn be resolved by using the STMM structure (see [4]). Indeed, the following theorem can be shown by simple "concatenation" of results provided in [13, 20] and [4].

Theorem 1. *Let the transmit matrix $\mathbf{S}_t$ possess the STMM structure in (6), and consider the time-varying MIMO channel in (3) or (4). Furthermore, let at least one of the K diagonal data matrices $\mathbf{D}_k$ be nonsingular and let $M_{\mathrm{R}} \geq M_{\mathrm{T}}P$, $N > \max\{\frac{M_{\mathrm{T}}P(M_{\mathrm{R}}L+L-M_{\mathrm{R}})-M_{\mathrm{R}}}{M_{\mathrm{R}}-M_{\mathrm{T}}P}, \frac{(M_{\mathrm{T}}P)^2-1}{M_{\mathrm{T}}P-K}\}$, $K \leq M_{\mathrm{T}}$ for $P \geq 2$, and $K < M_{\mathrm{T}}$ for $P = 1$. Then, there exists a set of K admissible Doppler-spread modulation matrices $\widetilde{\mathbf{M}}_k$.*

Hence, using these admissible $\widetilde{\mathbf{M}}_k$, the diagonal data matrices $\mathbf{D}_k$ (and, thus, the data $d_k[n]$) can be reconstructed from the received matrix $\mathbf{X}$ up to an unknown constant factor $c \in \mathbb{C}$. This reconstruction uses the knowledge of the modulation matrices $\widetilde{\mathbf{M}}_k$ but does not require knowledge of the channel $\mathbf{R}$.

Theorem 1 can be reformulated as follows. Let $K' \leq K$ be the number of data streams actually present (K' need not be known to the receiver). Then, the transmit matrix $\mathbf{S}$ occurring in (4) is given by $\mathbf{S} = \mathcal{T}_L\{\sum_{k=1}^{K'} \widetilde{\mathbf{M}}_k \mathbf{D}_k\}$ (cf. (8)), and the received matrix is $\mathbf{X} = \mathbf{RS} = \mathbf{R}\,\mathcal{T}_L\{\sum_{k=1}^{K'} \widetilde{\mathbf{M}}_k \mathbf{D}_k\}$. Suppose there is another couple $(\hat{\mathbf{R}}, \{\hat{\mathbf{D}}_k\})$ that also "explains" $\mathbf{X}$, i.e., we have also $\mathbf{X} = \hat{\mathbf{R}}\,\mathcal{T}_L\{\sum_{k=1}^{K} \widetilde{\mathbf{M}}_k \hat{\mathbf{D}}_k\}$. Now Theorem 1 states that there exist Doppler-spread modulation matrices $\widetilde{\mathbf{M}}_k$ such that the identity

$$\hat{\mathbf{R}}\,\mathcal{T}_L\left\{\sum_{k=1}^{K} \widetilde{\mathbf{M}}_k \hat{\mathbf{D}}_k\right\} = \mathbf{R}\,\mathcal{T}_L\left\{\sum_{k=1}^{K'} \widetilde{\mathbf{M}}_k \mathbf{D}_k\right\}$$

implies $\hat{\mathbf{R}} = c\mathbf{R}$ (with $c \in \mathbb{C}$ an unknown factor) and

$$\hat{\mathbf{D}}_k = \begin{cases} (1/c)\mathbf{D}_k, & k \leq K' \\ \mathbf{0}, & K'+1 \leq k \leq K. \end{cases} \tag{9}$$

Thus, $\mathbf{X}$ can be uniquely factored (up to an unknown constant factor $c \in \mathbb{C}$) into the channel $\mathbf{R}$ and the data matrices $\mathbf{D}_k$.

Unfortunately, the assumption $M_{\mathrm{R}} \geq M_{\mathrm{T}}P$ causes Theorem 1 to be of limited practical interest. Therefore, we will now consider an alternative result that does not rely on this assumption.

Second Reconstruction Result. The next theorem can be derived by exploiting the combined structure offered by the channel model and by the STMM scheme, rather than "concatenating" these structures in the two-step approach that underlies Theorem 1.

Theorem 2. *Let the transmit matrix $\mathbf{S}_t$ possess the STMM structure in (6), and consider the time-varying MIMO channel in (3) or (4). Furthermore, let at least one of the K diagonal data matrices $\mathbf{D}_k$ be nonsingular and let $N \geq \frac{(M_{\mathrm{T}}P)^2-1}{M_{\mathrm{T}}P-K} + \frac{K(L-1)}{M_{\mathrm{T}}-K}$ and $K < \min\{\mathrm{rank}\{\mathbf{R}\}, M_{\mathrm{T}}+1\}$. Then, with probability one[2], there exists a set of K admissible $\widetilde{\mathbf{M}}_k$.*

We emphasize that this theorem is also valid for $M_{\mathrm{R}} < M_{\mathrm{T}}P$. Its proof is more difficult than that of Theorem 1 (mainly because for $M_{\mathrm{R}} < M_{\mathrm{T}}P$, $\mathbf{X}$ contains only a part of the row-span of $\mathbf{S}$, i.e., $\mathbf{R}$ is not left invertible); it cannot be included here because of space constraints. We note that the theorem can be extended to the multi-user case, which however is beyond the scope of this paper.

5. ITERATIVE DEMODULATION ALGORITHM

Next, we propose an iterative demodulation algorithm for STMM transmission over a doubly selective MIMO channel. This algorithm if valid for both $M_{\mathrm{R}} < M_{\mathrm{T}}P$ and $M_{\mathrm{R}} \geq M_{\mathrm{T}}P$. However, for $M_{\mathrm{R}} \geq M_{\mathrm{T}}P$ there exists a more efficient POCS algorithm similar to the one presented in [13].

Given a received matrix $\mathbf{X} = \mathbf{RS}$ and assuming admissible Doppler-spread modulation matrices $\widetilde{\mathbf{M}}_k$, it follows from Section 4 that for any pair of matrices $\hat{\mathbf{R}}$ and $\hat{\mathbf{S}}$ that satisfy the two properties

1. $\hat{\mathbf{R}}\hat{\mathbf{S}} = \mathbf{X}$
2. $\hat{\mathbf{S}} = \mathcal{T}_L\{\sum_{k=1}^{K} \widetilde{\mathbf{M}}_k \hat{\mathbf{D}}_k\}$ with diagonal $\hat{\mathbf{D}}_k$,

the matrices $\hat{\mathbf{D}}_k$ contain the correct data up to a common constant factor. This motivates an iterative algorithm which consists in alternately executing two different steps that enforce one of the above properties. The ith iteration is as follows.

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

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

Step 2. This step attempts to enforce Property 2. That is, given $\hat{\mathbf{S}}_1^{(i)}$ from Step 1 above, we calculate a generator matrix $\hat{\mathbf{S}}_g^{(i)}$ with STMM structure, i.e., $\hat{\mathbf{S}}_g^{(i)} = \sum_{k=1}^{K} \widetilde{\mathbf{M}}_k \hat{\mathbf{D}}_k^{(i)}$, where the $\hat{\mathbf{D}}_k^{(i)}$ are chosen such that the product $\hat{\mathbf{R}}^{(i)}\mathcal{T}_L\{\hat{\mathbf{S}}_g^{(i)}\}$ best approximates $\hat{\mathbf{R}}^{(i)}\hat{\mathbf{S}}_1^{(i)}$ in the LS sense. Since $\hat{\mathbf{R}}^{(i)}\hat{\mathbf{S}}_1^{(i)} = \mathbf{X}$ (see Step 1), these $\hat{\mathbf{D}}_k^{(i)}$ are such that $\hat{\mathbf{R}}^{(i)}\mathcal{T}_L\{\hat{\mathbf{S}}_g^{(i)}\}$ best approximates $\mathbf{X}$. To solve this problem, we first rewrite (4) as

$$\mathcal{R}\mathbf{d} = \mathrm{vec}\{\mathbf{X}\},$$

where $\mathbf{d} \triangleq [d_1[-L+1] \cdots d_K[-L+1] \cdots d_1[N-1] \cdots d_K[N-1]]^T$, $\mathrm{vec}\{\mathbf{X}\}$ is the $M_{\mathrm{R}}N \times 1$ vector formed by stacking all columns of $\mathbf{X}$, and the $M_{\mathrm{T}}N \times K(L+N-1)$ matrix $\mathcal{R}$ is defined as

$$\mathcal{R} \triangleq \begin{bmatrix} \mathbf{R}_{L-1}^{(-L+1)} & \mathbf{R}_{L-2}^{(-L+2)} & \cdots & \mathbf{R}_0^{(0)} & & & \\ & \mathbf{R}_{L-1}^{(-L+2)} & \mathbf{R}_{L-2}^{(-L+3)} & \cdots & \mathbf{R}_0^{(1)} & & \mathbf{0} \\ \mathbf{0} & & \ddots & \ddots & & \ddots & \\ & & & \mathbf{R}_{L-1}^{(N-L)} & \mathbf{R}_{L-2}^{(N-L+1)} & \cdots & \mathbf{R}_0^{(N-1)} \end{bmatrix}, \tag{10}$$

where $\mathbf{R}_m^{(n)} \triangleq \mathbf{R}[m]\widetilde{\mathbf{M}}[n]$ with $\widetilde{\mathbf{M}}[n] \triangleq [\tilde{\mathbf{m}}_1[n] \cdots \tilde{\mathbf{m}}_K[n]]$. The above LS approximation problem is thus equivalent to choosing $\hat{\mathbf{d}}^{(i)}$ such that $\hat{\mathcal{R}}^{(i)}\hat{\mathbf{d}}^{(i)}$ best approximates $\mathrm{vec}\{\mathbf{X}\}$ in the LS sense; here, $\hat{\mathcal{R}}^{(i)}$ is defined as in (10) with $\mathbf{R}[m]$ replaced by $\hat{\mathbf{R}}^{(i)}[m]$ (note that $\hat{\mathbf{R}}^{(i)}[m]$ is contained in the matrix $\hat{\mathbf{R}}^{(i)}$ that was calculated in Step 1). The solution is given by

$$\hat{\mathbf{d}}^{(i)} = \hat{\mathcal{R}}^{(i)\#}\mathrm{vec}\{\mathbf{X}\}.$$

We can now calculate $\hat{\mathbf{S}}_g^{(i)} = \sum_{k=1}^{K} \widetilde{\mathbf{M}}_k \hat{\mathbf{D}}_k^{(i)}$, where the $\hat{\mathbf{D}}_k^{(i)}$ correspond to $\hat{\mathbf{d}}^{(i)}$. Finally, the desired matrix $\hat{\mathbf{S}}_2^{(i)}$ (to be used in Step 1 of the next iteration) is obtained as $\hat{\mathbf{S}}_2^{(i)} = \mathcal{T}_L\{\hat{\mathbf{S}}_g^{(i)}\}$.

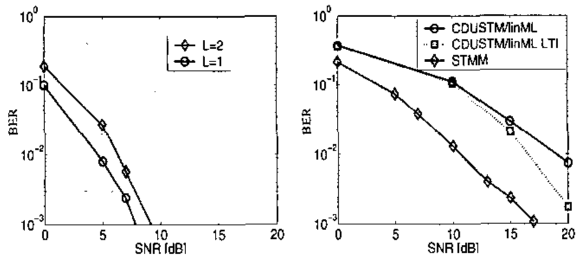
Remarks. This algorithm yields a channel matrix estimate $\hat{\mathbf{R}}^{(i)}$ in Step 1 and data matrix estimates $\hat{\mathbf{D}}_k^{(i)}$ in Step 2. In the noise-free case, we always observed the algorithm to converge to the true channel and data matrices. In the presence of noise, the algorithm converged to matrices that were close to the true channel and data matrices (detailed results in the presence of noise will be shown next). However, the convergence was observed to be rather slow.

6. SIMULATION RESULTS

We conducted two experiments in which a single random data stream $d_1[n]$ was transmitted over a time-varying MIMO channel. For each simulation run, the channel matrix $\mathbf{R}$ was randomly generated with iid complex-valued Gaussian entries. The channel output signals were corrupted by white Gaussian noise and observed over an interval of length $N = 100$. The modulation matrices $\mathbf{M}_k$ were constructed by taking rows of a DFT matrix as the rows of $\mathbf{M}_k$. We assumed two active Doppler shifts $l_1 = -1$ and $l_2 = 1$ (this can be interpreted as a crude approximation to a Jakes Doppler profile).

First Experiment. In our first experiment, we compare the performance of our STMM scheme (using the iterative demodulation technique of Section 5) for two channels with $M_{\mathrm{T}} = M_{\mathrm{R}} = 4$ and different delay spreads. Channel 1 is flat fading ($L = 1$) whereas

[2] For this theorem, the channel $\mathbf{R}$ is modeled as a realization of a random channel that is governed by an arbitrary continuous probability density.



(a) (b)

Fig. 2: BER vs. SNR obtained with STMM using the iterative demodulation algorithm: (a) comparison of channel lengths $L = 1$ and $L = 2$ (for $K = 1$ and $M_T = M_R = 4$), (b) comparison to CDUSTM (for $K = 1$, $M_T = M_R = 2$, and $L = 1$).

Channel 2 has a small delay spread ($L = 2$). We used a single data stream ($K = 1$) and 4-QAM symbols, corresponding to a data rate of 2 bit per channel use. Fig. 2(a) shows the BER[3] vs. the SNR. (The SNR is defined as $\frac{E\{\|X\|^2\}}{M_R N \sigma^2}$, where σ^2 is the noise variance.) The BER performance for Channel 1 is about 2 dB better than for Channel 2 even though Channel 2 has a higher available diversity. We attribute this behavior to the larger number of unknown parameters for Channel 2. Indeed, the channel matrix $\mathbf{R}$ has size 4×8 for Channel 1 and 4×16 for Channel 2, and thus Channel 2 has twice as many parameters that need to be estimated by our algorithm.

Second Experiment. Next, we compare our STMM scheme to the *Cayley differential unitary ST modulation (CDUSTM)* scheme introduced in [9]. We chose $L = 1$ since the CDUSTM scheme assumes a flat fading channel. Furthermore, we used $M_T = M_R = 2$ and a data rate of 2 bit per channel use. For the STMM scheme, this rate was achieved by using a single data stream ($K = 1$) and 4-QAM symbols. For the CDUSTM scheme, we used the optimized codes of [9].

Fig. 2(b) shows the BER vs. the SNR for our STMM scheme using the iterative demodulation technique and for the CDUSTM scheme using linearized ML decoding (denoted by 'CDUSTM/linML'). It is seen that STMM outperforms CDUSTM by up to 7 dB, even though our demodulation algorithm is an equalization technique (followed by quantization) and not a detection technique such as ML decoding. We attribute the good performance of our algorithm to the significantly larger block length allowed by STMM and by the time-varying channel model (we used $N = 100$), as compared to the small block length of $N = 4$ required by CDUSTM. Because of the larger N we have many more equations than unknowns, which results in better demodulation results. We also note that STMM is a purely spatial code; typically, it will be augmented by an outer temporal code that can be expected to result in further improvements of performance.

Fig. 2(b) also shows the performance of CDUSTM with linearized ML decoding for the unrealistic case that no Doppler spread is present, i.e., $P = 1$ and $l_1 = 0$ (denoted by 'CDUSTM/linML LTI'). This allows to assess the performance loss suffered by CDUSTM due to the Doppler. While the performance of CDUSTM for the time-invariant channel is better than for the time-varying channel (especially at higher SNR), it is still significantly poorer than the performance of STMM for the time-varying channel.

7. CONCLUSION

Space-time matrix modulation (STMM) is a simple and attractive transmission scheme for *unknown* MIMO channels. In this paper, we showed that both the STMM scheme (including perfect reconstruction for unknown channels) and a corresponding iterative demodulation algorithm can be extended to *doubly selective* (i.e., delay-spread and Doppler spread) MIMO channels. This channel model allows for large block lengths even in the case of fast fading channels. Apparently, STMM is currently the only space-time modulation/demodulation technique that allows perfect data reconstruction for transmission over an unknown doubly selective channel. The good performance of STMM and, specifically, significant performance advantages over Cayley differential unitary space-time modulation were demonstrated by means of simulations.

[3] For calculation of the BER, we assume that the factor c in (9) is known.

ACKNOWLEDGMENT

The authors would like to thank D. Seethaler for help with the simulations.

REFERENCES

[1] 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.

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

[3] 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.

[4] 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.

[5] H. Artés and F. Hlawatsch, 'Space-time matrix modulation: Rank-defi cient channels and multi-user case," in *Proc. IEEE ICASSP-2002*, (Orlando, FL), pp. 2225–2228, May 2002.

[6] H. Artés, G. Matz, D. Schafhuber, and F. Hlawatsch, 'Space-time multicarrier matrix modulation for unknown dispersive MIMO channels," in *Proc. 39th Allerton Conf. Commun., Contr., Comput.*, (Urbana, IL), pp. 466–475, Oct. 2001.

[7] B. Hassibi and B. Hochwald, 'How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, submitted.

[8] X. Ma and G. Giannakis, 'Space-time coding for doubly-selective channels," in *Proc. IEEE ISCAS-02*, vol. 3, pp. 647–650, 2002.

[9] B. Hassibi and B. Hochwald, 'Cayley differential unitary space-time codes," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1485–1503, June 2002.

[10] P. A. Bello, 'Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.

[11] A. M. Sayeed and B. Aazhang, 'Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Comm.*, vol. 47, pp. 123–132, Jan. 1999.

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

[13] H. Artés and F. Hlawatsch, 'Blind multiuser equalization for time-varying channels," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, (Taoyuan, Taiwan (R.O.C.)), pp. 102–105, March 2001.

[14] L. A. Liporace, 'Linear estimation of nonstationary signals," *J. Acoust. Soc. Amer.*, vol. 58, pp. 1288–1295, Dec. 1975.

[15] G. B. Giannakis and C. Tepedelenlioğlu, 'Basis expansion models and diversity techniques for blind identifi cation and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.

[16] R. S. Kennedy, *Fading Dispersive Communication Channels*. New York: Wiley, 1969.

[17] H. Artés, G. Matz, and F. Hlawatsch, 'Unbiased scattering function estimation during data transmission," in *Proc. IEEE VTC-99 Fall*, (Amsterdam, The Netherlands), pp. 1535–1539, Sept. 1999.

[18] C. Tepedelenlioğlu and G. B. Giannakis, 'Precoding for scattering function estimation of mobile channels using output correlations only," in *Proc. 37th Allerton Conf. on Commun., Contr., Comput.*, (Urbana, IL), Sept. 1999.

[19] J. W. Brewer, 'Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits and Systems*, vol. CAS-25, pp. 772–781, Sept. 1978.

[20] 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.