IDENTIFICATION OF MULTIPLE-INPUT MULTIPLE-OUTPUT CHANNELS UNDER LINEAR SIDE CONSTRAINTS

Dae Gwan Lee[⋆] *Götz E. Pfander*[⋆] *Volker Pohl*[†] *Weiqi Zhou*[⋆]

[⋆] Mathematisch-Geographische Fakultät
Katholische Universität Eichstätt-Ingolstadt, 85071 Eichstätt, Germany
[†] Lehrstuhl für Theoretische Informationstechnik
Technische Universität München, 80333 München, Germany

ABSTRACT

We investigate the impact of having additional information in a form of linear constraints in the channel identification problem. With those constraints taken into account, the problem turns into solving a linear system that is associated with a block matrix where each submatrix is either a Gabor system matrix or a matrix prescribed by the linear constraints. So, the identifiability hinges on whether one can find some generating windows of the Gabor systems for which the full linear system is solvable. We show that in single-input single-output (SISO) settings as well as in multiple-input multiple-output (MIMO) settings, linear constraints consisting of a single equation are beneficial for channel identification, as there always exist windows for which the corresponding full linear system is solvable. Concerning multiple linear constraints, however, there exists a set of linear constraints with two equations for which the full linear system is singular for all choices of windows. In the SISO case, we also provide some sufficient conditions on the linear side constraints under which the full linear system is solvable.

***Index Terms*—** Channel identification, operator sampling, operator Paley–Wiener space, time-frequency analysis.

1. INTRODUCTION

The identification of time-varying channels is an important problem in communications with a long research history. In a series of papers [1–4] necessary and sufficient conditions for the identifiability of the channel in terms of its spreading function were derived and sufficient conditions for appropriate test signals were given. More recently, ideas from compressive sampling where incorporated to identify sparsely supported channels with unknown support [5, 6] and extensions to stochastic channel models where investigated [7, 8].

During the last decades communication systems with multiple-input and multiple-output (MIMO) antennas gained in importance because the channel capacity of such systems scales, in principle, linearly with the minimum of the number of input and the number of output antennas [9]. However, to achieve this potentially huge gain in channel capacity the different antennas at the input and output have to be uncorrelated. This requires a sufficiently large antenna spacing as well as a sufficiently rich scattering environment of the communication channel [10–12]. Apart from the increased capacity, the deployment of a very large number (up to hundreds or thousands) of antennas, operating coherently and adaptively, will help to simplify the signal processing, and to improve the energy efficiency as well as the reliability of the communication link. This paradigm, known as *massive MIMO* gained much interest over the last years [13].

With respect to the channel identification problem, a MIMO channel is much more demanding than a SISO channel [4]. Assuming a system with N inputs and M outputs, one has to identify $N \cdot M$ individual sub-channels. Nevertheless, due to potential coupling of antennas and due to fading correlations, these sub-channels are often not completely independent. In many cases, the relations between the different channels can be characterized analytically, e.g. in terms of an S-parameter model [14] or due to the particular channel model [15].

This paper investigates the channel identification problem for SISO and MIMO channels under the assumption of known correlations between the individual sub-channels. These correlations are taken into account by linear side constraints. Our approach allows to incorporate all kinds of linear relations between the different channels, including antenna coupling and fading correlations but also side constraints on the spreading function which are not induced by multiple antennas but by some constraints on the spreading function of the channel.

2. BACKGROUND AND PROBLEM FORMULATION

In the following, we shall describe the channel identification problem, introduce the operator classes of interest, and develop the necessary mathematical background.

2.1. Channel identification

To identify a channel prior to using it for communication is a classical problem in electrical engineering. We formalize this as follows.

Definition 1: *A class of linear operators* $\mathcal{H} \subseteq \mathcal{L}(\mathbb{C}^{L_1}, \mathbb{C}^{L_2})$ *is called* identifiable *if there exists a vector* $\boldsymbol{c} \in \mathbb{C}^{L_1}$ *with the property that the map*

$$\boldsymbol{\Phi}_{\boldsymbol{c}} : \mathcal{H} \longrightarrow \mathbb{C}^{L_2}, \;\; H \mapsto H\boldsymbol{c}$$

is injective. Such a vector $\boldsymbol{c}$ *is called an* identifier *for* $\mathcal{H}$.

If $\mathcal{H}$ is an identifiable linear space, then it is necessarily of dimension less or equal to L_2.

This work was supported by the German Research Foundation (DFG) within the priority program *Compressed Sensing in Information Processing (CoSIP)* under Grants PF 450/9–1 and PO 1347/3–1.

2.2. Identifiability of operator Paley–Wiener spaces in finite dimensions

Important communication channels, such as satellite, radio, microwave, infrared and the like, can be modeled in finite dimensions as linear combinations of discrete time-frequency shift operators $\boldsymbol{M}^\ell \boldsymbol{T}^k$ with $k, \ell = 0, \ldots, L-1$ and where $\boldsymbol{T}, \boldsymbol{M} : \mathbb{C}^L \to \mathbb{C}^L$ are the cyclic [1] time shift and frequency shift operators defined by

$$\boldsymbol{T}\boldsymbol{x} = (x_1, \ldots, x_{L-1}, x_0,) \qquad \text{and}$$
$$\boldsymbol{M}\boldsymbol{x} = (\omega^0 x_0, \omega^1 x_1, \ldots, \omega^{L-1} x_{L-1}),$$

respectively, where $\omega = e^{2\pi i/L}$. Since $\{\boldsymbol{M}^\ell \boldsymbol{T}^k\}_{k,\ell=0}^{L-1}$ forms a basis for $\mathcal{L}(\mathbb{C}^L, \mathbb{C}^L)$, every channel can be written as

$$\boldsymbol{H} = \textstyle\sum_{k,\ell=0}^{L-1} \eta(k,\ell) \boldsymbol{M}^\ell \boldsymbol{T}^k, \tag{1}$$

whose characteristics are encoded in its unique coefficients $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{H}) = \{\eta(k,\ell)\}_{k,\ell=0}^{L-1}$ (called the *spreading* coefficients of $\boldsymbol{H}$). Each coefficient $\eta(k,\ell)$ can be seen as a gain factor associated to a transmission path with time-delay k and frequency shift l caused by the Doppler effect. Then the channel identification problem seeks a vector $\boldsymbol{c} \in \mathbb{C}^L$ for which operators of the form (1) with a certain sparsity prior on $\boldsymbol{\eta}$ can be uniquely recovered from the channel output $\boldsymbol{H}\boldsymbol{c}$. This amounts to a question on degrees of freedom in the coefficients $\boldsymbol{\eta}$ (namely the sparsity level) and linear independence of vectors $\boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{c}$ that correspond to the nonzero entries of $\boldsymbol{\eta}$.

Definition 2: *For $\Lambda \subseteq \mathbb{Z}_L \times \mathbb{Z}_L$, we define the* single-input single-output operator Paley–Wiener space[2] *by*

$$\begin{aligned} OPW(\Lambda) &= \operatorname{span}\{\boldsymbol{M}^\ell \boldsymbol{T}^k : (k,\ell) \in \Lambda\} \\ &= \{\boldsymbol{H} \in \mathcal{L}(\mathbb{C}^L, \mathbb{C}^L) : \operatorname{supp} \boldsymbol{\eta} \subseteq \Lambda\}. \end{aligned}$$

Establishing identifiability of operator classes such as $OPW(\Lambda)$ is not always trivial, for example, the following result was established for L prime ten years prior to its full resolution.

Theorem 1 ([16,17]): *The space $OPW(\Lambda)$ is identifiable if and only if $|\Lambda| \leq L$.*

Apart from the the SISO model given by (1), we consider also the more general model for systems with multiple inputs and multiple outputs. In the N-input M-output case, the communication channel $\boldsymbol{H}$ consists of MN subchannels:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{1,1} & \cdots & \boldsymbol{H}_{1,N} \\ \vdots & & \vdots \\ \boldsymbol{H}_{M,1} & \cdots & \boldsymbol{H}_{M,N} \end{bmatrix},$$

where each subchannel $\boldsymbol{H}_{m,n}$ is of the form (1). Then the corresponding channel identification problem seeks an identifier $\boldsymbol{c} = (\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ for which every operator can be uniquely recovered from the output

$$\boldsymbol{H}\boldsymbol{c} = \begin{bmatrix} \boldsymbol{H}_{1,1} & \cdots & \boldsymbol{H}_{1,N} \\ \vdots & & \vdots \\ \boldsymbol{H}_{M,1} & \cdots & \boldsymbol{H}_{M,N} \end{bmatrix} \begin{bmatrix} \boldsymbol{c}^{(1)} \\ \vdots \\ \boldsymbol{c}^{(N)} \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^N \boldsymbol{H}_{1,n}\, \boldsymbol{c}^{(n)} \\ \vdots \\ \sum_{n=1}^N \boldsymbol{H}_{M,n}\, \boldsymbol{c}^{(n)} \end{bmatrix}.$$

We will denote by $\boldsymbol{\eta}_{m,n} = [\boldsymbol{\eta}_{m,n}(k,\ell)]_{k,\ell=0}^{L-1} \in \mathbb{C}^{L^2}$ the spreading coefficients of the subchannel $\boldsymbol{H}_{m,n}$.

Definition 3: *For $\boldsymbol{\Lambda} = [\Lambda_{m,n}]_{m=1}^{M}{}_{n=1}^{N}$ with $\Lambda_{m,n} \subseteq \mathbb{Z}_L \times \mathbb{Z}_L$, define the* MIMO operator Paley–Wiener space *$OPW(\boldsymbol{\Lambda})$ by*

$$OPW(\boldsymbol{\Lambda}) = [\,OPW(\Lambda_{m,n})\,]_{m=1}^{M}{}_{n=1}^{N}.$$

By definition, the operator space $OPW(\boldsymbol{\Lambda})$ is identifiable if there exists a $\boldsymbol{c} = (\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that the linear map $\boldsymbol{H} \mapsto \boldsymbol{H}\boldsymbol{c}$ is injective on $OPW(\boldsymbol{\Lambda})$, i.e., $\boldsymbol{H} \in OPW(\boldsymbol{\Lambda})$ and $\boldsymbol{H}\boldsymbol{c} = 0$ imply $\boldsymbol{H} = 0$.

Theorem 2 ([4,18]): *The space $OPW(\boldsymbol{\Lambda})$ is identifiable if and only if $\sum_{n=1}^N |\Lambda_{m,n}| \leq L$ for all $m = 1, \ldots, M$.*

The identifiability results presented above depend highly on the invertibility of submatrices of Gabor system matrices.

2.3. Gabor system matrices

Given a window $\boldsymbol{c} \in \mathbb{C}^L$, the *full Gabor system matrix $\boldsymbol{G}(\boldsymbol{c})$* is the $L \times L^2$ matrix whose columns are the time-frequency shifts $\boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{c}$ where $k, \ell = 0, \ldots, L-1$. That is

$$\boldsymbol{G}(\boldsymbol{c}) = [\ \boldsymbol{D}_0 \boldsymbol{W}_L \mid \boldsymbol{D}_1 \boldsymbol{W}_L \mid \cdots \mid \boldsymbol{D}_{L-1} \boldsymbol{W}_L\ ],$$

wherein $\boldsymbol{D}_k = \operatorname{diag}(\boldsymbol{T}^k \boldsymbol{c}) = \operatorname{diag}(c_k, \ldots, c_{L-1}, c_0, \ldots, c_{k-1})$ and $\boldsymbol{W}_L = (e^{2\pi i n m/L})_{n,m=0}^{L-1}$ is the $L \times L$ Fourier matrix. For any $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$, we denote by $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ the submatrix of $\boldsymbol{G}(\boldsymbol{c})$ formed with columns indexed by Λ, i.e., $\boldsymbol{G}(\boldsymbol{c})|_\Lambda = [\boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{c}]_{(k,\ell)\in\Lambda}$.

The *spark* of a matrix $\boldsymbol{A} \in \mathbb{C}^{m \times N}$ with $m < N$ is the size of the smallest linearly dependent subset of columns, i.e., $\operatorname{spark}(\boldsymbol{A}) = \min\{\|\boldsymbol{z}\|_0 : \boldsymbol{A}\boldsymbol{z} = 0,\ \boldsymbol{z} \neq 0\}$. We say that the matrix $\boldsymbol{A}$ has *full spark* if $\operatorname{spark}(\boldsymbol{A}) = m+1$. Here and in the following, $\|\boldsymbol{z}\|_0$ counts the number of nonzero entries in a vector $\boldsymbol{z}$.

Theorem 1 is equivalent to the existence of Gabor matrices with full spark [16, 17], while Theorem 2 is equivalent to the following result.

Proposition 3 ([18,19]): *For every $L, N \in \mathbb{N}$ there exists a dense, open subset $\mathcal{S}_N \subset (\mathbb{C}^L)^N$ of full measure such that the block matrix $\left[\boldsymbol{G}(\boldsymbol{c}^{(1)}) \,\middle|\, \boldsymbol{G}(\boldsymbol{c}^{(2)}) \,\middle|\, \cdots \,\middle|\, \boldsymbol{G}(\boldsymbol{c}^{(N)})\right]$ has full spark for all $(\boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}, \ldots, \boldsymbol{c}^{(N)}) \in \mathcal{S}_N$.*

3. CHANNEL IDENTIFICATION WITH SIDE CONSTRAINTS

In this paper, we aim to recover operators $OPW(\boldsymbol{\Lambda})$ even in the case where $\sum_{n=1}^N |\Lambda_{m,n}| > L$ for some m. This requires to consider additional *a priori* information on the operator that we aim to identify. To motivate our choice of linear side constraints, we give two simple examples.

Example 1: In the two-input single-output case ($N = 2$, $M = 1$), Theorem 2 states the necessary condition $|\Lambda_{1,1}| + |\Lambda_{1,2}| \leq L$ for $OPW(\boldsymbol{\Lambda})$ to be identifiable. However, if some components of the subchannels $\boldsymbol{H}_{1,1}$ and $\boldsymbol{H}_{1,2}$ are known to be identical, for example, if $(0,0) \in \Lambda_{1,1} \cap \Lambda_{1,2}$ and $\boldsymbol{\eta}_{1,1}(0,0) = \boldsymbol{\eta}_{1,2}(0,0)$, then we may expect that $OPW(\boldsymbol{\Lambda})$ is identifiable even if $|\Lambda_{1,1}| + |\Lambda_{1,2}| = L+1$ by counting the degrees of freedom. Below we shall show that this is indeed true, leading to identifiability of the corresponding space $OPW(\boldsymbol{\Lambda})$ with the prescribed constraints.

[1] Cyclic time shifts are certainly not an accurate representation of time delays that occur in communication channels. The transition from non-cyclic to cyclic shifts is achieved by applying a cyclic prefix.

[2] The terminology operator Paley–Wiener space stems from the analogous time-continuous identification problem. Here, $OPW(S)$ denotes the space of Hilbert–Schmidt operators on $L^2(\mathbb{R})$ with the property that its Kohn–Nierenberg symbol is bandlimited to $S \subset \mathbb{R}^2$, that is, the operators spreading function is supported on S. The results presented in this paper generalize to results on operator identification and operator sampling for SISO and MIMO operator Paley–Wiener spaces in the continuous setting.

Example 2: In the single-input two-output case ($N = 1$, $M = 2$), Theorem 2 gives the necessary condition $|\Lambda_{1,1}| \le L$ and $|\Lambda_{2,1}| \le L$ for $OPW(\boldsymbol{\Lambda})$ to be identifiable. However, if the subchannels $\boldsymbol{H}_{1,1}$ and $\boldsymbol{H}_{2,1}$ are known to be partially identical, say, $\boldsymbol{\eta}_{1,1}|_S = \boldsymbol{\eta}_{2,1}|_S$ for some $S \subset \Lambda_{1,1} \cap \Lambda_{2,1}$, then it turns out that $OPW(\boldsymbol{\Lambda})$ is identifiable if $\max\{|\Lambda_{1,1}\backslash S|, |\Lambda_{2,1}|\} \le L$, or $\max\{|\Lambda_{1,1}|, |\Lambda_{2,1}\backslash S|\} \le L$, or $\max\{|\Lambda_{1,1}\backslash S| + |\Lambda_{2,1}\backslash S|, |S|\} \le L$. This allows the identification of $OPW(\boldsymbol{\Lambda})$ even in the case $S = \Lambda_{1,1} \subset \Lambda_{2,1}$ with $|\Lambda_{1,1}| = L$ and $|\Lambda_{2,1}| = 2L$, where the classical requirements are clearly violated.

These two examples demonstrate that *a priori* knowledge on correlations of subchannels can contribute to the channel identification.

3.1. Problem formulation

In the SISO case, we have for a fixed $\boldsymbol{c} \in \mathbb{C}^L$

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{c} = \sum_{k,\ell=0}^{L-1} \eta(k,\ell)\boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{c} = \boldsymbol{G}(\boldsymbol{c})\boldsymbol{\eta}, \tag{2}$$

and we formulate the linear side constraints as $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$. Combining these equations, one obtains the linear system

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix} \boldsymbol{\eta}.$$

In the MIMO case, rewriting each $\boldsymbol{H}_{m,n}\boldsymbol{c}^{(n)}$ as in (2) yields that

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_M \end{bmatrix} = \boldsymbol{H}\boldsymbol{c} = \begin{bmatrix} \sum_{n=1}^N \boldsymbol{H}_{1,n}\, \boldsymbol{c}^{(n)} \\ \vdots \\ \sum_{n=1}^N \boldsymbol{H}_{M,n}\, \boldsymbol{c}^{(n)} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{n=1}^N \boldsymbol{G}(\boldsymbol{c}^{(n)})\boldsymbol{\eta}_{1,n} \\ \vdots \\ \sum_{n=1}^N \boldsymbol{G}(\boldsymbol{c}^{(n)})\boldsymbol{\eta}_{M,n} \end{bmatrix} = \begin{bmatrix} \widetilde{\boldsymbol{G}} & 0 & \cdots & 0 \\ 0 & \widetilde{\boldsymbol{G}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{\boldsymbol{G}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \vdots \\ \boldsymbol{\eta}_M \end{bmatrix},$$

where $\widetilde{\boldsymbol{G}} = [\boldsymbol{G}(\boldsymbol{c}^{(1)}) \,|\, \cdots \,|\, \boldsymbol{G}(\boldsymbol{c}^{(N)})] \in \mathbb{C}^{L\times NL^2}$ and $\boldsymbol{\eta}_m = \{\boldsymbol{\eta}_{m,n}\}_{n=1}^N \in (\mathbb{C}^{L^2})^N$. Similarly, we formulate the linear side constraints as $\boldsymbol{b} = \sum_{m=1}^M \boldsymbol{A}_m \boldsymbol{\eta}_m$, where $\boldsymbol{\eta}_m = \{\boldsymbol{\eta}_{m,n}\}_{n=1}^N \in (\mathbb{C}^{L^2})^N$. Combining all these equations, we have

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_M \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \widetilde{\boldsymbol{G}} & 0 & \cdots & 0 \\ 0 & \widetilde{\boldsymbol{G}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{\boldsymbol{G}} \\ \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_M \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \vdots \\ \boldsymbol{\eta}_M \end{bmatrix}.$$

These observations immediately lead to the following result.

Theorem 4: *(a) The SISO operator Paley–Wiener space $OPW(\Lambda)$ with side constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ is identifiable if and only if there exists $\boldsymbol{c} \in \mathbb{C}^L$ for which the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A}|_\Lambda \end{bmatrix}$ is injective.*

(b) The MIMO operator Paley–Wiener space $OPW(\boldsymbol{\Lambda})$ with side constraints $\boldsymbol{b} = \sum_{m=1}^M \boldsymbol{A}_m \boldsymbol{\eta}_m$ is identifiable if and only if there exists $\boldsymbol{c} = (\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ for which the matrix

$$\begin{bmatrix} \widetilde{\boldsymbol{G}}|_{\Lambda_1} & 0 & \cdots & 0 \\ 0 & \widetilde{\boldsymbol{G}}|_{\Lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{\boldsymbol{G}}|_{\Lambda_M} \\ \boldsymbol{A}_1|_{\Lambda_1} & \boldsymbol{A}_2|_{\Lambda_2} & \cdots & \boldsymbol{A}_M|_{\Lambda_M} \end{bmatrix}, \tag{3}$$

where $\widetilde{\boldsymbol{G}} = [\boldsymbol{G}(\boldsymbol{c}^{(1)}) \,|\, \cdots \,|\, \boldsymbol{G}(\boldsymbol{c}^{(N)})] \in \mathbb{C}^{L\times NL^2}$ and $\Lambda_m = \{\Lambda_{m,n}\}_{n=1}^N \in (\mathbb{Z}_L \times \mathbb{Z}_L)^N$, is injective.

Clearly, this theorem leads us to investigate matrices of the form $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A}|_\Lambda \end{bmatrix}$ in the SISO case, and matrices of the form (3) in the MIMO case.

4. IDENTIFIABILITY RESULTS FOR SISO

We first consider the case of SISO channels. Let us begin with an example of a matrix $\boldsymbol{A} \in \mathbb{C}^{L\times L^2}$ such that the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A}|_\Lambda \end{bmatrix}$ is not injective for all $\boldsymbol{c} \in \mathbb{C}^L$ and with $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ of size $2L$.

Example 3: For

$$\boldsymbol{A} = [\boldsymbol{I}_L \,|\, \boldsymbol{M}^{-1} \,|\, \cdots \,|\, \boldsymbol{M}^{-(L-1)}] \quad \in \mathbb{C}^{L\times L^2},$$

the $2L \times L^2$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix}$ is rank deficient for all $\boldsymbol{c} \in \mathbb{C}^L$. Indeed, the sum of all rows of $\boldsymbol{G}(\boldsymbol{c})$ is $\sum_{\ell=0}^{L-1}(\sum_{k=0}^{L-1} \omega^{k\ell} c_k)\boldsymbol{v}_\ell$, where $\boldsymbol{v}_\ell$ denotes the ℓth row vector of $\boldsymbol{A}$ (ordered from top to bottom), therefore, the rows of $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix}$ are linearly dependent.

In the simplest case where the matrix $\boldsymbol{A}$ consists of a single row vector, we show below that given any set $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ of size $L+1$ there exists a $\boldsymbol{c} \in \mathbb{C}^L$ such that $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A}|_\Lambda \end{bmatrix}$ is injective. This fact indicates that linear constraints consisting of a single nontrivial equation always contribute to the SISO channel identification.

In the following, we denote by $\mathcal{S}$ the set of all $\boldsymbol{c} \in \mathbb{C}^L$ for which $\boldsymbol{G}(\boldsymbol{c})$ has full spark.

Theorem 5: *Let $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $L+1 \le |\Lambda| \le 2L$. Then*

$$\operatorname{span}\big\{\ker \boldsymbol{G}(\boldsymbol{c})|_\Lambda : \boldsymbol{c} \in \mathcal{S}\big\} = \mathbb{C}^\Lambda.$$

Corollary 6: *For any $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $|\Lambda| = L+1$ and $\boldsymbol{a} \in \mathbb{C}^\Lambda\backslash\{0\}$ there exists a $\boldsymbol{c} \in \mathbb{C}^L$ for which the $(L+1)\times(L+1)$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible.*

We give a proof for this result but skip the more involved derivation of Theorem 5.

Proof: Suppose that for some subset $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ of size $L+1$ and some $\boldsymbol{a} \in \mathbb{C}^\Lambda\backslash\{0\}$ there exists no such $\boldsymbol{c} \in \mathbb{C}^L$. Then for any $\boldsymbol{c} \in \mathcal{S}$ the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is not invertible, which means that $\boldsymbol{a}^*$ is in the row space of $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ as the rows of $\boldsymbol{G}(\boldsymbol{c})_\Lambda$ are linearly independent for $\boldsymbol{c} \in \mathcal{S}$. The fundamental theorem of linear algebra and Theorem 5 imply

$$\boldsymbol{a} \in \bigcap_{\boldsymbol{c}\in\mathcal{S}} \operatorname{ran} \boldsymbol{G}(\boldsymbol{c})|_\Lambda^* = \bigcap_{\boldsymbol{c}\in\mathcal{S}} \ker \boldsymbol{G}(\boldsymbol{c})|_\Lambda^\perp = \{0\},$$

which is a contradiction. ∎

Unfortunately, Theorem 5 does not allow us to draw conclusions concerning multiple linear constraints. Below we give an example of $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with size $L+2$ and linear constraints of two equations for which the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A} \end{bmatrix}$ is always singular. This clarifies that the statement of Corollary 6 cannot be extended to multiple linear constraints.

Example 4: Let $L = 5$ and $\Lambda = \{(0,0),(0,1),(0,2),(1,0),(1,1)\}$. The matrix

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A} \end{bmatrix} = \begin{bmatrix} c_0 & c_0 & c_0 & c_1 & c_1 \\ c_1 & \omega c_1 & \omega^2 c_1 & c_2 & \omega c_2 \\ c_2 & \omega^2 c_2 & \omega^4 c_2 & c_0 & \omega^2 c_0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

is singular for all $\boldsymbol{c} = (c_0, c_1, c_2)^{\mathrm{T}} \in \mathbb{C}^3$; indeed, the first row is always linearly dependent with the fourth and the fifth row.

Note that increasing the ratio of the size of Λ with the number of constraints may help to keep the intersection of the row space of $\boldsymbol{A}$ and the row space of $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ trivial. However, Example 3 shows that there exists an L-dimensional row space that intersects nontrivially with the row space of $\boldsymbol{G}(\boldsymbol{c})$ for every $\boldsymbol{c} \in \mathcal{S}$.

4.1. Sufficient conditions

Although it is impossible to extend Corollary 6 to multiple linear constraints (Example 4), it is still possible that for *some* $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ and $\boldsymbol{A}$ exists $\boldsymbol{c}$ such that $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A} \end{smallmatrix}\right]$ is injective. In the following, we present some conditions on $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ and $\boldsymbol{A}$ that guarantee the existence of such $\boldsymbol{c}$.

Theorem 7: *Let $\boldsymbol{A} = [\boldsymbol{A}_0 \,|\, \boldsymbol{A}_1 \,|\, \cdots \,|\, \boldsymbol{A}_{L-1}] \in \mathbb{C}^{L \times L^2}$, where each $\boldsymbol{A}_k$ is an $L \times L$ matrix. If $\det \boldsymbol{A}_k \neq (-1)^{L-1} \cdot \det \boldsymbol{A}_{k+1}$ for some k, then the $2L \times L^2$ matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{smallmatrix}\right]$ has full rank for a.e. $\boldsymbol{c} \in \mathbb{C}^L$. Further, if L is a prime, it is sufficient that $\det \boldsymbol{A}_k \neq (-1)^{L-1} \det \boldsymbol{A}_\ell$ for some distinct k and ℓ.*

For the matrix $\boldsymbol{A}$ given in Example 3, the $2L \times L^2$ matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{smallmatrix}\right]$ is rank deficient for all $\boldsymbol{c} \in \mathbb{C}^L$. Note that since $\det(\boldsymbol{M}^{-k}) = (-1)^{k(L-1)}$ for $k = 0, \ldots, L-1$, the conditions of Theorem 7 are not satisfied. However, as soon as one of the submatrices of $\boldsymbol{A}$ is scaled by a non-unit constant, e.g., if

$$\boldsymbol{A} = [2\boldsymbol{I}_L \,|\, \boldsymbol{M}^{-1} \,|\, \cdots \,|\, \boldsymbol{M}^{-(L-1)}] \quad \in \mathbb{C}^{L \times L^2},$$

it follows from Theorem 7 that the matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{smallmatrix}\right]$ has full rank for almost every $\boldsymbol{c} \in \mathbb{C}^L$.

To state our next result, we need the following definitions.

Definition 4: *To each $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$, we associate an L-tuple $\tau(\Lambda) = \boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_{L-1})$, where $\tau_k = \tau_k(\Lambda) := |\Lambda \cap (\{k\} \times \mathbb{Z}_L)|$ is the number of elements of the form (k, ℓ), $\ell \in \mathbb{Z}_L$ contained in Λ.*

It is clear that for every $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$, the associated L-tuple $\boldsymbol{\tau} = \tau(\Lambda)$ satisfies $\boldsymbol{\tau} \in (\mathbb{Z}_L)^L$ and $\|\boldsymbol{\tau}\|_1 = |\Lambda|$ (referred to as the *size* of τ). We also define a partial order on the set of all L-tuples: $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_{L-1}) \preceq \boldsymbol{\tau}' = (\tau'_0, \tau'_1, \ldots, \tau'_{L-1})$ if and only if $\tau_j \le \tau'_j$ for all j. To each L-tuple of size L, we associate an integer-valued index number defined up to modulo L.

Definition 5: *For any $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_{L-1}) \in (\mathbb{Z}_L)^L$ with $\|\tau\|_1 = L$, the* index number *of $\boldsymbol{\tau}$ is defined as $\mathrm{ind}(\boldsymbol{\tau}) = L(L-1)/2 + \sum_{j=0}^{L-1} j \cdot \tau_j$ modulo L.*

Theorem 8: *Let $\widetilde{\Lambda} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ be of size R ($> L$). Assume that there exists a subset $\Lambda \subset \widetilde{\Lambda}$ of size L with*

(i) $\tau_j(\Lambda) = \tau_j(\widetilde{\Lambda})$ whenever $\tau_j(\Lambda) \neq 0$;

(ii) $\mathrm{ind}(\boldsymbol{\tau}') \neq \mathrm{ind}(\tau(\Lambda))$ for every L-tuple $\boldsymbol{\tau}' \preceq \tau(\widetilde{\Lambda})$ of size L different from $\tau(\Lambda)$.

Given any full spark matrix $\boldsymbol{A}$ of size $(R-L) \times R$, the vectors $\boldsymbol{c} \in \mathbb{C}^L$ for which the $R \times R$ matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{A} \end{smallmatrix}\right]$ is invertible constitute a dense open subset of $\mathbb{C}^L$ with full Lebesgue measure.

To support the result, we give some examples.

Example 5: (a) Let $L = 5$ and $\widetilde{\Lambda} = \{(0,0),(0,1),(0,2),(0,3),(0,4),(1,0),(1,1),(1,2),(1,3)\}$. Then $\tau(\widetilde{\Lambda}) = (5,4,0,0,0)$ and the matrix $\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}$ is given by

$$\left[\begin{array}{ccccc|cccc} c_0 & c_0 & c_0 & c_0 & c_0 & c_1 & c_1 & c_1 & c_1 \\ c_1 & \omega c_1 & \omega^2 c_1 & \omega^3 c_1 & \omega^4 c_1 & c_2 & \omega c_2 & \omega^2 c_2 & \omega^3 c_2 \\ c_2 & \omega^2 c_2 & \omega^4 c_2 & \omega^6 c_2 & \omega^8 c_2 & c_3 & \omega^2 c_3 & \omega^4 c_3 & \omega^6 c_3 \\ c_3 & \omega^3 c_3 & \omega^6 c_3 & \omega^9 c_3 & \omega^{12} c_3 & c_4 & \omega^3 c_4 & \omega^6 c_4 & \omega^9 c_4 \\ c_4 & \omega^4 c_4 & \omega^8 c_4 & \omega^{12} c_4 & \omega^{16} c_4 & c_0 & \omega^4 c_0 & \omega^8 c_0 & \omega^{12} c_0 \end{array}\right].$$

It is easy to see that the set $\Lambda = \{(0,0),(0,1),(0,2),(0,3),(0,4)\}$ satisfies the conditions of Theorem 8. Therefore, given any full spark matrix $\boldsymbol{A}$ of size 4×9, the 9×9 matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{A} \end{smallmatrix}\right]$ is invertible for almost every choice of $\boldsymbol{c}$ in $\mathbb{C}^5$.
(b) Let $L = 3$ and $\widetilde{\Lambda} = \{(0,0),(0,1),(1,0),(2,0)\}$. Then $\tau(\widetilde{\Lambda}) = (2,1,1)$ and

$$\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} = \left[\begin{array}{cc|c|c} c_0 & c_0 & c_1 & c_2 \\ c_1 & \omega c_1 & c_2 & c_0 \\ c_2 & \omega^2 c_2 & c_0 & c_1 \end{array}\right].$$

It is easily seen that the set $\Lambda = \{(0,0),(0,1),(1,0)\}$ fulfills the conditions of Theorem 8. Consequently, given any vector $\boldsymbol{a} \in \mathbb{C}^4$ with no zero entries, the 4×4 matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{a}^* \end{smallmatrix}\right]$ is invertible for almost every choice of $\boldsymbol{c}$ in $\mathbb{C}^3$.

5. IDENTIFIABILITY RESULTS FOR MIMO

Now we consider the identification of MIMO channels. Thereby, we focus on MISO channels, because the identification of an N-input M-output channel is equivalent to the identification of M MISO channels. So the respective results generalize to the MIMO setting. As in the SISO setting, it turns out that linear constraints consisting of a single nontrivial equation can always contribute to the MISO channel identification. This is shown by the following statements, which generalize Theorem 5 and Corollary 6 to the MISO case.

Theorem 9: *Let $\Lambda_1, \Lambda_2, \ldots, \Lambda_N \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $L + 1 \le \sum_{n=1}^N |\Lambda_n| < 2L$, where $L \ge 2$ and $N \ge 1$. Then*

$$\mathrm{span}\left\{\ker\left[\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_1} \ \cdots \ \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda_N}\right] : \boldsymbol{c} \in \mathcal{S}_N\right\} = \mathbb{C}^\Lambda,$$

where $\mathcal{S}_N$ is the set of all $(\boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ for which $[\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_1} \ \cdots \ \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda_N}]$ has full spark (cf. Proposition 3).

Corollary 10: *Let $L \ge 2$ and $N \ge 1$. For any $\Lambda_1, \Lambda_2, \ldots, \Lambda_N \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\sum_{n=1}^N |\Lambda_n| = L+1$ and $\boldsymbol{a} \in \mathbb{C}^{L+1} \backslash \{0\}$, there exists $(\boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ for which the $(L+1) \times (L+1)$ matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_1} \ \cdots \ \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda_N} \\ \boldsymbol{a}^* \end{smallmatrix}\right]$ is invertible.*

Concerning multiple constraints, the SISO case (Example 4) already shows that the statement of Corollary 10 cannot be extended to linear side constraints of two equations.

6. REFERENCES

[1] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Systems*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[2] T. Kailath, "Time-variant communication channels," *IEEE Trans. Inf. Theory*, vol. 9, no. 4, pp. 233–237, Oct. 1963.

[3] G. E. Pfander and D. Walnut, "Measurement of time-variant channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4808–4820, 2006.

[4] G. E. Pfander, "Measurement of time-varying multiple-input multiple-output channels," *Appl. Comput. Harmon. Anal.*, vol. 24, no. 3, pp. 393–401, 2008.

[5] R. Heckel and H. Bölcskei, "Identification of sparse linear operators," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7985–8000, Dec. 2013.

[6] G. E. Pfander and D. F. Walnut, "Sampling and reconstruction of operators," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 435–458, Jan 2016.

[7] G. E. Pfander and P. Zheltov, "Sampling of stochastic operators," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2359–2372, Mar. 2014.

[8] G. E. Pfander and P. Zheltov, "Identification of stochastic operators," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 2, pp. 256–279, 2014.

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

[10] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[11] V. Jungnickel, V. Pohl, and C. von Helmolt, "Capacity of MIMO Systems with Closely-Spaced Antennas," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 361–363, Aug. 2003.

[12] H. N. M. Mbonjo, J. Hansen, and V. Hansen, "MIMO capacity and antenna array design," in *Proc. IEEE Global Telecommun. Conf.*, vol. 5, Nov. 2004, pp. 3155–3159.

[13] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[14] J. W. Wallace and M. A. Jensen, "Mutual coupling in MIMO wireless systems: A rigorous network theory analysis," *IEEE Trans. Wireless Commun*, vol. 3, no. 4, pp. 1317–1325, Jul. 2004.

[15] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *IEEE Global Telecommun. Conf.*, vol. 2, 2000, pp. 1083–1088.

[16] J. Lawrence, G. E. Pfander, and D. Walnut, "Linear independence of Gabor systems in finite dimensional vector spaces," *J. Fourier Anal. Appl.*, vol. 11, no. 6, pp. 715–726, 2005.

[17] R.-D. Malikiosis, "A note on Gabor frames in finite dimensions," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 2, pp. 318–330, 2015.

[18] D. G. Lee, G. E. Pfander, and V. Pohl, "Compressive sampling of multiple-input multiple-output channels," 2017, preprint.

[19] D. G. Lee, G. E. Pfander, V. Pohl, and W. Zhou, "Sampling of correlated channels with multiple inputs and outputs," 2017, preprint.