Linear Algebra and its Applications 581 (2019) 435–470

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Identification of channels with single and multiple inputs and outputs under linear constraints

Dae Gwan Lee [a,*], Götz E. Pfander [a], Volker Pohl [b], Weiqi Zhou [a]

[a] *Mathematisch-Geographische Fakultät, Katholische Universität Eichstätt-Ingolstadt, 85071 Eichstätt, Germany*

[b] *Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Arcisstraße 21, 80333 München, Germany*

A R T I C L E   I N F O

Article history:
Received 27 November 2018
Accepted 19 July 2019
Available online 31 July 2019
Submitted by V. Mehrmann

MSC:
15A15
42C15

Keywords:
Channel identification
Operator sampling
Time-frequency analysis
Gabor matrices
Finite frames
Multiple-input multiple-output channels

A B S T R A C T

The channel identification problem for multiple-input multiple-output (MIMO) channels under linear constraints can be formulated as solving a linear system which involves finite-dimensional Gabor matrices and a pre-determined unstructured matrix that represents the linear constraints. While matrices of the latter type are fixed a priori by the constraints, Gabor matrices depend on the choice of their generating windows which is often chosen by the user. This is important since even if the matrix associated with the linear constraints is ill-conditioned, the full system may be solvable if the windows are designed appropriately.

We prove that linear constraints consisting of a single equation always remove a single degree of freedom in the channel identification problem, in the sense that preknowledge of such constraints allows identification of MIMO channels with support size one greater than the fundamental limit. However, we give an explicit example showing that this statement does not generalize to the case of multiple constraints. In the single-input single-output (SISO) case, we provide some sufficient

* Corresponding author.
E-mail addresses: daegwans@gmail.com (D.G. Lee), pfander@ku.de (G.E. Pfander), volker.pohl@tum.de (V. Pohl), anothergalaas@gmail.com (W. Zhou).

https://doi.org/10.1016/j.laa.2019.07.025
0024-3795/

conditions on the linear side constraints under which the corresponding SISO channels are identifiable.

1. Introduction

System identification is an important and challenging task in a variety of areas of science and engineering such as physics, communications and control theory. The problem is to identify a linear system H from its response Hg to a test signal g. In particular, we ask which systems can be identified by such a probing scheme and how one has to design the test signal g in order to identify a given class of systems.

This paper is motivated by the channel identification problem for continuous-time *multiple-input multiple-output (MIMO)* channels $\boldsymbol{H} : (L^2(\mathbb{R}))^N \to (L^2(\mathbb{R}))^M$,

$$\boldsymbol{H}\begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} = \begin{pmatrix} H_{1,1} & \dots & H_{1,N} \\ \vdots & & \vdots \\ H_{M,1} & \dots & H_{M,N} \end{pmatrix}\begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} = \begin{pmatrix} \sum_{n=1}^N H_{1,n} f_n \\ \vdots \\ \sum_{n=1}^N H_{M,n} f_n, \end{pmatrix}, \tag{1}$$

where each subchannel $H_{m,n}$ is an operator of the form $H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ with

$$(Hf)(x) = \iint_{\mathbb{R}\times\mathbb{R}} \eta_H(t,\nu)\, M_\nu T_t f(x)\, d\nu\, dt, \quad f \in L^2(\mathbb{R}). \tag{2}$$

Here, N and M are the number of inputs and outputs, respectively, $\eta_H(t,\nu)$ is the *spreading function* of H, T_t is translation (time shift) by $t \in \mathbb{R}$, that is, $T_t f(x) = f(x-t)$, and M_ν is modulation (frequency shift) by $\nu \in \mathbb{R}$, that is, $M_\nu f(x) = e^{2\pi i \nu x} f(x)$. If $N = M = 1$, we call $\boldsymbol{H} = \{H_{1,1}\}$ a *Single-Input Single-Output (SISO)* channel.

The identification problem for MIMO channels of the form (1) can be reduced [1,2] to identifying a finite-dimensional system $\boldsymbol{H} : (\mathbb{C}^L)^N \to (\mathbb{C}^L)^M$,

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

where each subsystem $\boldsymbol{H}_{m,n} : \mathbb{C}^L \to \mathbb{C}^L$ is of the form

$$\boldsymbol{H}_{m,n} = \sum_{k,\ell=0}^{L-1} \eta_{m,n}(k,\ell)\, \boldsymbol{M}^\ell \boldsymbol{T}^k. \tag{3}$$

Here, $\boldsymbol{T}$ and $\boldsymbol{M}$ denote translation and modulation on $\mathbb{C}^L$. They are defined by $\boldsymbol{Tx} = (x_1, \dots, x_{L-1}, x_0)$ and $\boldsymbol{Mx} = (\omega^0 x_0, \omega^1 x_1, \dots, \omega^{L-1} x_{L-1})$ with $\omega = e^{2\pi i/L}$, respectively. This reduction allows us to carry out most of the analysis on identifiability in a finite-dimensional setting. Therefore, we focus our attention to the finite-dimensional channel identification problem which is formulated in Section 2. Its relation to the corresponding continuous-time problem is discussed in detail in Section 5.

Systems of the form (2) are of particular importance in communications as Hf in (2) describes the propagation of a signal f through a time-varying dispersive communication channel [3,4]. In this case H is referred to as a linear time-varying channel with the *delay-Doppler* spreading function η_H. Especially in this communications setting, the identification problem for linear systems of the form (2) has drawn much attention in the past. In the SISO case, a series of papers [1,5–7] studied necessary and sufficient conditions for the identifiability of H in terms of the support set of η_H. More recently, ideas from compressive sampling were applied to identify sparsely supported channels [8] and extension to stochastic channel models was considered in [9].

During the last decades MIMO communication systems have gained much importance because the channel capacity of such systems scales, in principle, linearly with the number of input and output antennas [10]. However, to achieve this potential gain in channel capacity, the signals at the different inputs and outputs have to be uncorrelated. In a communication setting, this requires a sufficiently large antenna spacing as well as a sufficiently rich scattering environment of the communications channel [11–13]. Apart from increasing the channel capacity, deploying a very large number (up to hundreds or thousands) of antennas that operate coherently and adaptively helps the signal processing and improves the energy efficiency and reliability of the communication link. This paradigm, known as *massive MIMO*, has gained much interest over the last years [14]. On the other hand, the problem of identifying a system certainly gets more demanding as the number of inputs and outputs increases [7], simply because for a system with N inputs and M outputs, one has to identify $N \cdot M$ individual subsystems. Nevertheless, due to potential coupling of antennas and due to fading correlations, these subchannels are often not completely independent. In many cases, the relation between subchannels can be characterized analytically for specific channel models, see for example, [15,16].

In this paper, we formulate and study the channel identification of deterministic SISO and MIMO systems whose subchannels have some a priori known linear relationship. The linear relations are expressed in terms of linear constraints that can accommodate relations between subchannels (for example, antenna coupling and fading correlations) as well as relations within each subchannel. We would like to emphasize that the focus of this paper is in the deterministic setting. Also, we assume that the spreading support of each subchannel is known [2,17,18]. The corresponding problem in the stochastic setting is to be treated in a forthcoming work.

This paper is organized as follows. In Section 2, we give a detailed review on the finite-dimensional channel identification problem and then formulate our main problem on the identification of SISO and MIMO channels under linear side constraints. Section 3 is devoted to the SISO case, in particular, it is shown that linear side constraints consisting of a single equation always remove a single degree of freedom from the SISO channel identification problem. We also provide some sufficient conditions on the linear side constraints under which the corresponding SISO channels are identifiable. We extend the SISO results to the MIMO setting in Section 4. Applications in the continuous-time setting are discussed in Section 5, followed by some concluding remarks in Section 6. All proofs are collected in Section 7.

2. Background and problem formulation

In this section, we shall first review the finite-dimensional channel identification problem and then formulate our main problem which takes into consideration linear side constraints. We shall also develop the necessary mathematical background for the problem at hand. The transition to continuous-time channels is discussed later in Section 5.

2.1. Channel identification in finite dimensions

Identifying a channel prior to using it for communication is a classical problem in electrical engineering. Communication channels, such as satellite, radio, microwave, can be modeled in finite dimensions as linear combinations of discrete time-frequency shift operators $\boldsymbol{M}^\ell \boldsymbol{T}^k$, $k, \ell = 0, \dots, L-1$, where $\boldsymbol{T}, \boldsymbol{M} : \mathbb{C}^L \to \mathbb{C}^L$ are the cyclic[1] time shift and frequency shift operators defined respectively by

$$\begin{aligned} \boldsymbol{T}\boldsymbol{x} &= (x_1, \dots, x_{L-1}, x_0), \\ \boldsymbol{M}\boldsymbol{x} &= (\omega^0 x_0, \omega^1 x_1, \dots, \omega^{L-1} x_{L-1}) \quad \text{with } \omega = e^{2\pi i/L}. \end{aligned}$$

Since $\{\boldsymbol{M}^\ell \boldsymbol{T}^k\}_{k,\ell=0}^{L-1}$ forms a basis of $\mathcal{L}(\mathbb{C}^L, \mathbb{C}^L)$, every linear single-input single-output (SISO) channel can be expanded in the form of (3) with *unique* coefficients $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{H}) = \{\eta(k,\ell)\}_{k,\ell=0}^{L-1}$ called the *spreading coefficients* of $\boldsymbol{H}$ [19, Lemma 1]. These coefficients encode all the characteristics of $\boldsymbol{H}$, for example, each coefficient $\eta(k,\ell)$ can be seen as a gain factor associated with a transmission path with respective time-delay k and frequency shift ℓ caused by the Doppler effect.

In this setting, the channel identification problem asks whether there exists a vector $\boldsymbol{c} \in \mathbb{C}^L$ such that every matrix of the form (3) with some restrictions (e.g., sparsity) on $\boldsymbol{\eta}$ can be uniquely recovered from $\boldsymbol{H}\boldsymbol{c}$.

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

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

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

[1] Cyclic time shifts are certainly not an accurate representation of time delays that occur in communication channels. To transition from non-cyclic to cyclic shifts, a cyclic prefix may be applied.

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

Using the representation (3), any set of operators $\mathcal{H} \subset \mathcal{L}(\mathbb{C}^L, \mathbb{C}^L)$ can be expressed in terms of spreading coefficients, that is,

$$\mathcal{H} = \left\{ \sum_{k,\ell=0}^{L-1} \eta(k,\ell)\, \boldsymbol{M}^\ell \boldsymbol{T}^k : \{\eta(k,\ell)\}_{k,\ell=0}^{L-1} \in \Omega \right\}$$

where operators in $\mathcal{H}$ and elements in $\Omega \subset \mathbb{C}^{L\times L}$ are in one-to-one correspondence. In particular, we are interested in operator classes with $\Omega \subset \mathbb{C}^{L\times L} = \mathbb{C}^{\mathbb{Z}_L\times\mathbb{Z}_L}$ being of the form $\mathbb{C}^\Lambda \times \{0\}^{(\mathbb{Z}_L\times\mathbb{Z}_L)\setminus\Lambda}$ with $\Lambda \subset \mathbb{Z}_L\times\mathbb{Z}_L$, where B^A denotes the set of functions from A to B. Here and subsequently we will abuse notation and not distinguish the sets $\mathbb{C}^\Lambda \times \{0\}^{(\mathbb{Z}_L\times\mathbb{Z}_L)\setminus\Lambda}$, $\mathbb{C}^\Lambda$, and $\mathbb{C}^{|\Lambda|}$, as they are all isomorphic.

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

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

For example, the class of operators consisting of linear combinations of translation operators $\boldsymbol{T}^k$, $k = 0, 1, \ldots, L-1$, corresponds to having $\eta(k,\ell) = 0$ for all but $\ell = 0$ (i.e., the only possible nonzero coefficients are $\eta(0,0)$, $\eta(1,0)$, $\cdots$, $\eta(L-1,0)$) and therefore can be expressed as $OPW(\mathbb{Z}_L\times\{0\})$.

According to Definition 1, the space $OPW(\Lambda)$ is identifiable if and only if there exists a vector $\boldsymbol{c} \in \mathbb{C}^L$ such that for each $\boldsymbol{H} \in OPW(\Lambda)$ the equation

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{c} = \sum_{(k,\ell)\in\Lambda} \eta(k,\ell)\, \boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{c} = \boldsymbol{G}(\boldsymbol{c})\, \boldsymbol{\eta} \tag{4}$$

is uniquely solvable in $\boldsymbol{\eta} = \{\eta(k,\ell)\}_{(k,\ell)\in\Lambda} \in \mathbb{C}^\Lambda$, where $\boldsymbol{G}(\boldsymbol{c})$ is the Gabor matrix introduced in the next section.

Let us mention that in this paper, we only consider the case where the spreading support $\Lambda \subset \mathbb{Z}_L\times\mathbb{Z}_L$ is known. For the study of identification of operators with unknown spreading support, we refer to [2,17,18] and references therein.

2.2. 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 $\boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{c}$, $k, \ell = 0, \ldots, L-1$, of $\boldsymbol{c}$, that is,

[2] The terminology operator Paley-Wiener space stems from the analogous continuous-time identification problem which is discussed in Section 5.

$$\begin{aligned} \boldsymbol{G}(\boldsymbol{c}) &= \big[\, \boldsymbol{c},\, \boldsymbol{M}\boldsymbol{c},\, \ldots,\, \boldsymbol{M}^{L-1}\boldsymbol{c} \;\big|\; \boldsymbol{T}\boldsymbol{c},\, \boldsymbol{M}\boldsymbol{T}\boldsymbol{c},\, \ldots,\, \boldsymbol{M}^{L-1}\boldsymbol{T}\boldsymbol{c} \;\big| \\ &\qquad \ldots \;\big|\; \boldsymbol{T}^{L-1}\boldsymbol{c},\, \boldsymbol{M}\boldsymbol{T}^{L-1}\boldsymbol{c},\, \ldots,\, \boldsymbol{M}^{L-1}\boldsymbol{T}^{L-1}\boldsymbol{c} \,\big] \\ &= \big[\, \boldsymbol{D}_0\boldsymbol{W}_L \;\big|\; \boldsymbol{D}_1\boldsymbol{W}_L \;\big|\; \cdots \;\big|\; \boldsymbol{D}_{L-1}\boldsymbol{W}_L \,\big], \end{aligned} \tag{5}$$

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

Recall that 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, that is, $\operatorname{spark}(\boldsymbol{A}) = \min\{\|\boldsymbol{z}\|_0 : \boldsymbol{A}\boldsymbol{z} = 0,\ \boldsymbol{z} \neq 0\}$. We say that $\boldsymbol{A}$ has *full spark* if $\operatorname{spark}(\boldsymbol{A}) = m + 1$. Here and in the following, $\|\boldsymbol{z}\|_0$ denotes the number of nonzero entries in a vector $\boldsymbol{z}$.

For $\boldsymbol{c} \neq 0$, the L^2 columns of $\boldsymbol{G}(\boldsymbol{c})$ form a tight frame of $\mathbb{C}^L$ with frame bound $L\,\|\boldsymbol{c}\|_2^2$ (see, e.g., [19]) and therefore $\boldsymbol{G}(\boldsymbol{c})$ always has full rank. Moreover, it is known that there exists $\boldsymbol{c}$ so that $\boldsymbol{G}(\boldsymbol{c})$ has full spark:

Theorem 3 *([19,20]). Let $L \in \mathbb{N}$. For any $\Lambda \subset \mathbb{Z}_L\times\mathbb{Z}_L$ of size L, $\det(\boldsymbol{G}(\boldsymbol{c})|_\Lambda)$ is a* non-trivial *homogeneous polynomial of degree L in the variables $c_0, \ldots, c_{L-1}$. Consequently, there exists a dense open subset $\mathcal{S} \subset \mathbb{C}^L$ of full measure such that $\boldsymbol{G}(\boldsymbol{c})$ has full spark for $\boldsymbol{c} \in \mathcal{S}$.*

Theorem 3 was proved in [19] for L prime by isolating the so-called *lowest index monomial* from the determinant of $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$. Later, [20] extended the result to all positive integers $L \in \mathbb{N}$ by finding the so-called *consecutive index monomial*.

As seen in Equation (4), the Gabor matrix $\boldsymbol{G}(\boldsymbol{c})$ plays a crucial role in operator identification. When restricting Equation (4) to the space $OPW(\Lambda)$ with $\Lambda \subset \mathbb{Z}_L\times\mathbb{Z}_L$, that is, if $\boldsymbol{\eta}$ is supported in Λ, the equation reduces to

$$\boldsymbol{y} = \boldsymbol{G}(\boldsymbol{c})|_\Lambda\, \boldsymbol{\eta}|_\Lambda.$$

This implies that $OPW(\Lambda)$ is identifiable if and only if the matrix $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ has linearly independent columns. As a direct consequence of Theorem 3, we have the following characterization on the identifiability of $OPW(\Lambda)$.

Corollary 4. *The space $OPW(\Lambda)$ is identifiable if and only if $|\Lambda| \leq L$.*

2.3. Multiple-input multiple-output channel operators

In the multiple-input multiple-output (MIMO) setting, say, with N inputs and M outputs, a communication channel $\boldsymbol{H}$ consists of MN subchannels and is represented by the $M\times N$ block matrix

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

where each subchannel $\boldsymbol{H}_{m,n} \in \mathbb{C}^{L\times L}$ is of the form (3). The corresponding channel identification problem asks for the existence of a vector $\boldsymbol{c} = (\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that $\boldsymbol{H}$ can be uniquely recovered from

$$\boldsymbol{H}\boldsymbol{c} = \begin{bmatrix} \boldsymbol{H}_{1,1} & \cdots & \boldsymbol{H}_{1,N} \\ \vdots & & \vdots \\ \boldsymbol{H}_{M,1} & \cdots & \boldsymbol{H}_{M,N} \end{bmatrix} \begin{bmatrix} \boldsymbol{c}^{(1)} \\ \vdots \\ \boldsymbol{c}^{(N)} \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N} \boldsymbol{H}_{1,n}\, \boldsymbol{c}^{(n)} \\ \vdots \\ \sum_{n=1}^{N} \boldsymbol{H}_{M,n}\, \boldsymbol{c}^{(n)} \end{bmatrix} \tag{6}$$

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

Definition 5. For $\boldsymbol{\Lambda} = [\Lambda_{m,n}]_{m=1}^{M}{}_{n=1}^{N}$ with $\Lambda_{m,n} \subset \mathbb{Z}_L \times \mathbb{Z}_L$, define the *MIMO operator Paley-Wiener space* $OPW(\boldsymbol{\Lambda})$ by

$$OPW(\boldsymbol{\Lambda}) = \{\boldsymbol{H} : \ \boldsymbol{H}_{mn} \in OPW(\Lambda_{m,n}),\ m = 1, \ldots, M,\ n = 1, \ldots, N\}.$$

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

For the MIMO operator Paley-Wiener space $OPW(\boldsymbol{\Lambda})$, we have the following identifiability result.[3]

Theorem 6 *([7]). The space $OPW(\boldsymbol{\Lambda})$ is identifiable if and only if $\sum_{n=1}^{N} |\Lambda_{m,n}| \leq L$ for $m = 1, \ldots, M$.*

As a consequence, the space $OPW(\boldsymbol{\Lambda})$ is identifiable if and only if for $m = 1, \ldots, M$ the space $OPW(\boldsymbol{\Lambda}_m)$ with $\boldsymbol{\Lambda}_m = \{\Lambda_{m,n}\}_{n=1}^{N}$ is identifiable. This reflects the fact that N-input M-output channels can be separated into M systems of N-input single-output channels. Certainly, this simplification fails in the case that we focus on, that is, in the case that linear relations between subchannels of different m are present.

Theorem 6 is a direct consequence of the following generalization of Theorem 3.

Theorem 7. *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 matrix*

$$\boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) := \Big[\, \boldsymbol{G}(\boldsymbol{c}^{(1)}) \ \Big|\ \boldsymbol{G}(\boldsymbol{c}^{(2)}) \ \Big|\ \cdots \ \Big|\ \boldsymbol{G}(\boldsymbol{c}^{(N)}) \,\Big] \in \mathbb{C}^{L\times NL^2} \tag{7}$$

has full spark for $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in \mathcal{S}_N$.

[3] Theorem 6 was first proved in [7] using a different proof technique.

We can generalize Theorem 7 further to the following theorem.

Theorem 8. *For every $L, N \in \mathbb{N}$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^{\mathrm{T}} \in \mathbb{C}^N \backslash \{0\}$, there exists a dense open subset $\mathcal{S}_{N,\boldsymbol{\alpha}} \subset (\mathbb{C}^L)^N$ of full measure with the property that the matrix*

$$[\,\boldsymbol{G}(\textstyle\sum_{n=1}^N \alpha_n \boldsymbol{c}^{(n)})|_\Lambda \;\; \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \;\; \cdots \;\; \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda^{(N)}}\,] \;\in \mathbb{C}^{L\times L} \tag{8}$$

is invertible for every $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in \mathcal{S}_{N,\boldsymbol{\alpha}}$ and every sets $\Lambda, \Lambda^{(1)}, \ldots, \Lambda^{(N)} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\Lambda \cap (\Lambda^{(1)} \cup \ldots \cup \Lambda^{(N)}) = \emptyset$ and $|\Lambda| + |\Lambda^{(1)}| + \ldots + |\Lambda^{(N)}| = L$.

An application of Theorem 8 is given in Example 33 below. See Section 7.2 for the proof of Theorems 7 and 8.

2.4. Linear relations between and within subchannels

The necessary and sufficient condition for the identifiability of $OPW(\boldsymbol{\Lambda})$ in Theorem 6 is based on the assumption that all subchannels and their components are independent, in the sense that information about one subchannel does not help to identify another. In case that linear relations between the subchannels (or their components) are known, for example, when transmission and/or receiving antennas are not well separated, one should certainly try to take advantage of such information in channel identification.

Let us now formalize the linear relations in terms of linear constraints. In the SISO setting, we express the linear relations between the entries of $\boldsymbol{\eta}$ by the equation

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}.$$

Including Equation (4), we obtain

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

Further, if $\boldsymbol{\eta} \in \mathbb{C}^{L^2}$ is known to be supported in a set $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$, the system reduces to

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A}|_\Lambda \end{bmatrix} \boldsymbol{\eta}|_\Lambda. \tag{9}$$

In the MIMO setting, by rewriting

$$\boldsymbol{H}_{m,n}\, \boldsymbol{c}^{(n)} = \boldsymbol{G}(\boldsymbol{c}^{(n)})\, \boldsymbol{\eta}_{m,n}, \quad m = 1, \ldots, M, \quad n = 1, \ldots, N,$$

in (6) we obtain the linear system

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

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

where $\boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)}) = [\,\boldsymbol{G}(\boldsymbol{c}^{(1)})\,|\,\cdots\,|\,\boldsymbol{G}(\boldsymbol{c}^{(N)})\,] \in \mathbb{C}^{L\times NL^2}$ and $\boldsymbol{\eta}_m = \{\boldsymbol{\eta}_{m,n}\}_{n=1}^N \in (\mathbb{C}^{L^2})^N$ for $m = 1,\dots,M$. Similarly, we express the linear relations between and within the vectors $\boldsymbol{\eta}_m$, $m = 1,\dots,M$, by the equation

$$\boldsymbol{b} = \sum_{m=1}^{M} \boldsymbol{A}_m \boldsymbol{\eta}_m.$$

Combining the equations leads to[4]

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_M \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)}) & 0 & \cdots & 0 \\ 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)}) \\ \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_M \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \vdots \\ \boldsymbol{\eta}_M \end{bmatrix}.$$

Under the assumption that $\boldsymbol{\eta}_m = \{\boldsymbol{\eta}_{m,n}\}_{n=1}^N \in (\mathbb{C}^{L^2})^N$ is supported in $\Lambda_m = \{\Lambda_{m,n}\}_{n=1}^N \subset (\mathbb{Z}_L \times \mathbb{Z}_L)^N$, i.e., $\operatorname{supp} \boldsymbol{\eta}_{m,n} \subset \Lambda_{m,n}$ for all m and n, the system reduces again to

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_M \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)})|_{\Lambda_1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)})|_{\Lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{G}(\boldsymbol{c}^{(1)},\dots,\boldsymbol{c}^{(N)})|_{\Lambda_M} \\ \boldsymbol{A}_1|_{\Lambda_1} & \boldsymbol{A}_2|_{\Lambda_2} & \cdots & \boldsymbol{A}_M|_{\Lambda_M} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_1|_{\Lambda_1} \\ \boldsymbol{\eta}_2|_{\Lambda_2} \\ \vdots \\ \boldsymbol{\eta}_M|_{\Lambda_M} \end{bmatrix}. \quad (10)$$

Now that the linear relations are incorporated into (9) and (10) for SISO and MIMO, respectively, the identifiability of the corresponding class of channels with the linear relations can be formalized as follows.

[4] Organizing the collection of output vectors in a matrix, that is, choosing the representation

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)}) & \cdots & \boldsymbol{G}(\boldsymbol{c}^{(N)}) \\ \boldsymbol{A}'_1 & \cdots & \boldsymbol{A}'_N \end{bmatrix} [\,\boldsymbol{\eta}_1 \ \cdots \ \boldsymbol{\eta}_M\,] = \begin{bmatrix} \boldsymbol{y}_1 & \cdots & \boldsymbol{y}_M \\ \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_M \end{bmatrix},$$

seems more natural. But unfortunately, side constraints would be of the form $\boldsymbol{b}_1 = \sum_{n=1}^N \boldsymbol{A}'_n \boldsymbol{\eta}_{1,n}$, $\dots$, $\boldsymbol{b}_M = \sum_{n=1}^N \boldsymbol{A}'_n \boldsymbol{\eta}_{M,n}$, which can only encode linear relations within each vector $\boldsymbol{\eta}_m$ but not between $\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_M$.

Definition 9. The SISO operator Paley-Wiener space $OPW(\Lambda)$ with constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ is identifiable if there exists a vector $\boldsymbol{c} \in \mathbb{C}^{L_1}$ such that (9) is uniquely solvable for every $H \in OPW(\Lambda)$.

The MIMO operator Paley-Wiener space $OPW(\boldsymbol{\Lambda})$ with constraints $\boldsymbol{b} = \sum_{m=1}^{M} \boldsymbol{A}_m \boldsymbol{\eta}_m$ is identifiable if there exists $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that (10) is uniquely solvable for every $H \in OPW(\boldsymbol{\Lambda})$.

We will refer to the constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ in the SISO case, respectively $\boldsymbol{b} = \sum_{m=1}^{M} \boldsymbol{A}_m \boldsymbol{\eta}_m$ in the MIMO case, the *side constraints* associated with $OPW(\Lambda)$, respectively $OPW(\boldsymbol{\Lambda})$.

The following proposition shows that the identifiability of $OPW(\Lambda)$, respectively $OPW(\boldsymbol{\Lambda})$, with side constraints depends only on the matrix $\boldsymbol{A}$, respectively matrices $\boldsymbol{A}_m$, $m = 1, \ldots, M$, and not on the choice of $\boldsymbol{b}$. As the proof is simple, we leave it to the reader.

Proposition 10. *(a) The SISO operator Paley-Wiener space $OPW(\Lambda)$ with side constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ is identifiable by $\boldsymbol{c} \in \mathbb{C}^L$ if and only if the matrix*

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\Lambda} \\ \boldsymbol{A}|_{\Lambda} \end{bmatrix} \tag{11}$$

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 by $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ if and only if the matrix

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\boldsymbol{\Lambda}_1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\boldsymbol{\Lambda}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\boldsymbol{\Lambda}_M} \\ \boldsymbol{A}_1|_{\boldsymbol{\Lambda}_1} & \boldsymbol{A}_2|_{\boldsymbol{\Lambda}_2} & \cdots & \boldsymbol{A}_M|_{\boldsymbol{\Lambda}_M} \end{bmatrix}, \tag{12}$$

with $\boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) = [\boldsymbol{G}(\boldsymbol{c}^{(1)}) \mid \cdots \mid \boldsymbol{G}(\boldsymbol{c}^{(N)})] \in \mathbb{C}^{L \times NL^2}$ and $\boldsymbol{\Lambda}_m = \{\Lambda_{m,n}\}_{n=1}^{N} \subset (\mathbb{Z}_L \times \mathbb{Z}_L)^N$ is injective.

This proposition leads us to investigate matrices of the form (11) and (12) for SISO and MIMO, respectively.

Note that choosing the empty set of side constraints gives the standard SISO/MIMO operator Paley-Wiener spaces. In that case, the matrix (11) is reduced to $\boldsymbol{G}(\boldsymbol{c})|_{\Lambda}$ which can be made injective if and only if $|\Lambda| \leq L$ (see Theorem 3 and Corollary 4). Similarly, the matrix (12) with the last row removed can be made injective if and only if $\sum_{n=1}^{N} |\Lambda_{m,n}| \leq L$ for $m = 1, \ldots, M$ (see Theorems 6 and 7).

Remark 11. The identifiability result for the standard operator Paley-Wiener spaces has been reduced by Corollary 4 and Theorem 6 to counting the degrees of freedom involved with the spreading support. Therefore, one could expect that a set of linear constraints with rank k would remove exactly k degrees of freedom from the channel identification problem. As we shall see, this is true in general only for $k = 1$ and even for this case the proof of the argument is not as easy as may be expected. The result cannot be deduced simply by counting the degrees of freedom but requires a careful analysis of the block matrices appearing in Proposition 10. See part (ii) of Remark 12 for additional remarks on degrees of freedom.

Remark 12. (i) It is worthwhile to note that the following statements are equivalent.

(a) The matrix $\begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{A} \end{bmatrix}$ is injective.
(b) $\ker \boldsymbol{G} \cap \ker \boldsymbol{A} = \{0\}$.
(c) The positive semi-definite matrix $\boldsymbol{G}^*\boldsymbol{G} + \boldsymbol{A}^*\boldsymbol{A}$ is strictly positive.

(ii) Let us recall that the intersection of two generic subspaces of dimension r and s of an L dimensional space is trivial if $r + s \le L$. Hence, for generic $\boldsymbol{G} \in \mathbb{C}^{L\times U}$ and $\boldsymbol{A} \in \mathbb{C}^{K\times U}$ with $L, K \le U$, condition (b) holds as long as $(U - L) + (U - K) \le L$, that is, $2U \le 2L + K$. Note further that in our setup, we have $U = |\Lambda|$, so it seems as if we can compensate any size of Λ as long as we have sufficiently many conditions on $\boldsymbol{A}$, that is, as long as K can compensate for large Λ. But this observation is far from concluding our discussion, as both matrices $\boldsymbol{A}$ and $\boldsymbol{G}$ are not generic in our setup. Moreover, the question at hand is whether for a given $\boldsymbol{A}$ (not a generic one) we can find a vector $\boldsymbol{c}$ so that $\ker(\boldsymbol{G}(\boldsymbol{c})|_{\Lambda}) \cap \ker \boldsymbol{A} = \{0\}$.

(iii) In the SISO setting, imposing the side constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ can be understood as restricting the vector $\boldsymbol{\eta} \in \mathbb{C}^{L^2}$ to the affine subspace $\{\boldsymbol{\eta} \in \mathbb{C}^{L^2} : \boldsymbol{A}\boldsymbol{\eta} = \boldsymbol{b}\}$. Side constraints in the MIMO setting can be viewed in a similar manner.

To illustrate how natural and useful the consideration of linear constraints is, we give two examples, one in the multiple-input single-output and one in the single-input multiple-output setting.

Example 13. (i) In the two-input single-output setting ($M = 1$, $N = 2$), Theorem 6 states the necessary and sufficient condition $|\Lambda_{1,1}| + |\Lambda_{1,2}| \le L$ for $OPW(\boldsymbol{\Lambda})$ to be identifiable. However, if the channel of interest is known to have an identical component in its subchannels, for example, if $\boldsymbol{H} = (\boldsymbol{H}_{1,1}, \boldsymbol{H}_{1,2})$ satisfies $(0,0) \in \Lambda_{1,1} \cap \Lambda_{1,2}$ and $\boldsymbol{\eta}_{1,1}(0,0) = \boldsymbol{\eta}_{1,2}(0,0)$, then by counting the degrees of freedom associated with $\boldsymbol{\eta}_{1,1}$ and $\boldsymbol{\eta}_{1,2}$ one can expect that $OPW(\boldsymbol{\Lambda})$ is identifiable even if $|\Lambda_{1,1}| + |\Lambda_{1,2}| = L + 1$. This is indeed true and can be verified using Proposition 10. A detailed argument is given in Section 7.1.

(ii) In the single-input two-output setting ($M = 2$, $N = 1$), Theorem 6 gives the necessary and sufficient condition that $|\Lambda_{1,1}| \leq L$ and $|\Lambda_{2,1}| \leq L$, for $OPW(\mathbf{\Lambda})$ to be identifiable. However, if a channel $\boldsymbol{H} = (\boldsymbol{H}_{1,1}, \boldsymbol{H}_{2,1}) \in OPW(\mathbf{\Lambda})$ is known to have some identical components in its subchannels, i.e., $\boldsymbol{\eta}_{1,1}|_S = \boldsymbol{\eta}_{2,1}|_S$ for some $S \subset \Lambda_{1,1} \cap \Lambda_{2,1}$, then it turns out that $OPW(\mathbf{\Lambda})$ is identifiable if $\max\{|\Lambda_{1,1} \backslash S|, |\Lambda_{2,1}|\} \leq L$, or $\max\{|\Lambda_{1,1}|, |\Lambda_{2,1} \backslash S|\} \leq L$, or $\max\{|\Lambda_{1,1} \backslash S| + |\Lambda_{2,1} \backslash S|, |S|\} \leq L$. This allows the identification of $\boldsymbol{H}$ even if the classical requirements $|\Lambda_{1,1}| \leq L$ and $|\Lambda_{2,1}| \leq L$ are not satisfied. As an extreme case, $\boldsymbol{H}$ is identifiable when $S = \Lambda_{1,1} \subset \Lambda_{2,1}$ with $|\Lambda_{1,1}| = L$ and $|\Lambda_{2,1}| = 2L$. See Section 7.1 for a detailed discussion.

Remark 14. To develop an identification procedure for $OPW(\Lambda)$ with side constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ involves the following steps. For simplicity, we shall assume $0 = \boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$. After ensuring that the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A} \end{bmatrix}$ has trivial kernel for some, and hence for almost all $\boldsymbol{c}$, we note that $\ker \boldsymbol{A} = (\operatorname{ran} \boldsymbol{A}^*)^\perp$ is a subspace of $\mathbb{C}^\Lambda$ with dimension $|\Lambda| - \operatorname{rank} \boldsymbol{A}$. Let $\boldsymbol{B} \in \mathbb{C}^{|\Lambda| \times (|\Lambda| - \operatorname{rank} \boldsymbol{A})}$ be a matrix whose columns form an orthonormal basis of $\ker \boldsymbol{A} = (\operatorname{ran} \boldsymbol{A}^*)^\perp \subset \mathbb{C}^\Lambda$. We now pick $\boldsymbol{c} \in \mathbb{C}^L$ so that the embedding

$$\boldsymbol{G}(\boldsymbol{c})|_\Lambda \circ \boldsymbol{B} \ \in \ \mathbb{C}^{L \times (|\Lambda| - \operatorname{rank} \boldsymbol{A})}$$

is bounded below by α and above by β with β/α not too large. The identification problem corresponds now to solving

$$(\boldsymbol{G}(\boldsymbol{c})|_\Lambda \circ \boldsymbol{B})\, \boldsymbol{x} = \boldsymbol{y}$$

for $\boldsymbol{x}$ and setting $\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{x}$. If $L \geq |\Lambda| - \operatorname{rank} \boldsymbol{A}$ and $\boldsymbol{G}(\boldsymbol{c})|_\Lambda \circ \boldsymbol{B}$ is left-invertible, then the space $OPW(\Lambda)$ with side constraints $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ is identifiable. Note that in case $L > |\Lambda| - \operatorname{rank} \boldsymbol{A}$, we can apply linear regression methods to obtain the least squares solution in case that our measurements $\boldsymbol{y}$ are affected by noise.

3. Identification results for SISO channels with constraints

Let us start with SISO channels that have some linearly related components. As in the statement of Proposition 10(a), we shall consider SISO channel operators in the space $OPW(\Lambda)$ and represent the linear relations by the equation $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$. Then, identifiability of $OPW(\Lambda)$ depends on the injectivity of the matrix given in (11). We will show that if $\boldsymbol{A}$ consists of a single row, then the space $OPW(\Lambda)$ with $|\Lambda| = L + 1$ under the linear relation $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\eta}$ is always identifiable. Comparing with Corollary 4, this result overcomes the fundamental limitation on the size of Λ by taking into account the linear side constraints.

Theorem 15. *Let $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $|\Lambda| = L + 1$, and $\boldsymbol{a} \in \mathbb{C}^\Lambda \backslash \{0\}$. There exists $\boldsymbol{c} \in \mathbb{C}^L$ such that the $(L+1) \times (L+1)$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible. Moreover, such vectors $\boldsymbol{c}$*

constitute a dense open subset of $\mathbb{C}^L$ *with full measure. Hence, the space* $OPW(\Lambda)$ *with side constraints* $b = \boldsymbol{a}^*\boldsymbol{\eta}|_\Lambda$, *where* $b \in \mathbb{C}$ *and* $\boldsymbol{\eta} \in \mathbb{C}^{L^2}$, *is identifiable.*

The proof of this result relies on the following lemma.

Lemma 16. *Let* $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ *with* $L+1 \leq R = |\Lambda| \leq L^2$. *Then*

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

where $\mathcal{S}$ *is the set of all* $\boldsymbol{c} \in \mathbb{C}^L$ *such that* $\boldsymbol{G}(\boldsymbol{c})$ *has full spark (cf. Theorem 3).*

Unfortunately, to obtain an identifiability result from this lemma requires to restrict ourselves to $|\Lambda| = L+1$ as in Theorem 15. The proof of Theorem 15 and Lemma 16 are contained in Section 7.3.

Theorem 15 does not allow us to draw conclusions for the case of linear constraints with multiple equations. Indeed, if $\boldsymbol{A}$ has multiple rows, the intersection of the row spaces of $\boldsymbol{A}$ and $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ may depend on the choice of $\boldsymbol{c}$. 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 such that the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{A} \end{bmatrix}$ is singular for all $\boldsymbol{c} \in \mathbb{C}^L$.

Example 17. Let $L = 3$ 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 a linear combination of the fourth and the fifth row.

Below we give 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 every $\boldsymbol{c} \in \mathbb{C}^L$ and $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ of size $2L$.

Example 18. If

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

then the $2L \times L^2$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix}$ is rank deficient for all $\boldsymbol{c} \in \mathbb{C}^L$. To see this, observe that the row vectors of $\boldsymbol{A}$ can be expressed as

$$\boldsymbol{v}_0 = (1,1,\ldots,1)^{\mathrm{T}} \otimes \boldsymbol{e}_0,$$

$$\boldsymbol{v}_1 = (1, \omega^{-1}, \dots, \omega^{-(L-1)})^{\mathrm{T}} \otimes \boldsymbol{e}_1,$$
$$\vdots$$
$$\boldsymbol{v}_{L-1} = (1, \omega^{-(L-1)}, \dots, \omega^{-(L-1)^2})^{\mathrm{T}} \otimes \boldsymbol{e}_{L-1},$$

where $\boldsymbol{e}_\ell$ is the ℓ-th canonical basis vector. It is easy to see that the sum of all rows of $\boldsymbol{G}(\boldsymbol{c})$ is equal to $\sum_{\ell=0}^{L-1}(\sum_{k=0}^{L-1}\omega^{k\ell}c_k)\,\boldsymbol{v}_\ell$, hence, the rows of $\left[\begin{smallmatrix}\boldsymbol{G}(\boldsymbol{c})\\ \boldsymbol{A}\end{smallmatrix}\right]$ are linearly dependent.

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. But Example 18 shows that there exists an L-dimensional subspace in the L^2-dimensional space, which intersects nontrivially with the L-dimensional row space of $\boldsymbol{G}(\boldsymbol{c})$ for every $\boldsymbol{c} \in \mathcal{S}$.

3.1. Sufficient conditions

As seen in Example 17, it is impossible to extend Theorem 15 in full generality to linear constraints with multiple equations. However, for *some* $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ and generic $\boldsymbol{A}$ there generally exists a $\boldsymbol{c}$ 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 an identifier $\boldsymbol{c}$. For this we need some definitions.

Definition 19. With each $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ we associate an L-tuple $\boldsymbol{\tau}(\Lambda) = \boldsymbol{\tau} = (\tau_0, \tau_1, \dots, \tau_{L-1})$, where $\tau_k = \tau_k(\Lambda) := |\Lambda \cap (\{k\} \times \mathbb{Z}_L)|$ is the number of elements in Λ that are of the form (k, ℓ), $\ell \in \mathbb{Z}_L$.

Clearly, any L-tuple $\boldsymbol{\tau} = \boldsymbol{\tau}(\Lambda)$ with $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ satisfies $\boldsymbol{\tau} \in \{0, 1, \dots, L\}^L$ and $\|\boldsymbol{\tau}\|_1 = |\Lambda|$ (referred to as the *size* of $\boldsymbol{\tau}$).

Proposition 20. *Let $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ be of size L, and $\boldsymbol{\tau}(\Lambda) = (\tau_0, \tau_1, \dots, \tau_{L-1})$. For every monomial $c_0^{\alpha_0} c_1^{\alpha_1} \dots c_{L-1}^{\alpha_{L-1}}$ appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_\Lambda)$, we have*

$$\sum_{j=0}^{L-1} j \cdot \alpha_j \;\equiv\; L(L-1)/2 + \sum_{j=0}^{L-1} j \cdot \tau_j \mod L.$$

Proof. When $\boldsymbol{\tau}(\Lambda) = (L, 0, \dots, 0)$, every generalized diagonal of $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ yields the monomial $c_0 c_1 \dots c_{L-1}$, whose indices add up to $0+1+\ldots+(L-1) = L(L-1)/2$. Now consider the general case $\boldsymbol{\tau}(\Lambda) = (\tau_0, \tau_1, \dots, \tau_{L-1})$. Compared to the case $(L, 0, \dots, 0)$, we have τ_1 columns with all indices of c_j's increased by 1 (modulo L), τ_2 columns with all indices of c_j's increased by 2 (modulo L), and so on. Therefore, any monomial produced from a generalized diagonal of $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$ has total index sum $L(L-1)/2+\tau_1+2\tau_2+\ldots+(L-1)\tau_{L-1}$ modulo L. □

Motivated by Proposition 20, we associate with each monomial in $c_0, \dots, c_{L-1}$ (resp., with each L-tuple of size L) an integer-valued index number defined modulo L.

Definition 21. The *index number* of a monomial $m = c_0^{\alpha_0} c_1^{\alpha_1} \dots c_{L-1}^{\alpha_{L-1}}$ is defined by $\operatorname{ind}(m) = \sum_{j=0}^{L-1} j \cdot \alpha_j$ modulo L.

Definition 22. For any $\boldsymbol{\tau} = (\tau_0, \tau_1, \dots, \tau_{L-1}) \in \{0, 1, \dots, L\}^L$ with $\|\boldsymbol{\tau}\|_1 = L$, the *index number* of $\boldsymbol{\tau}$ is defined as $\operatorname{ind}(\boldsymbol{\tau}) = L(L-1)/2 + \sum_{j=0}^{L-1} j \cdot \tau_j$ modulo L.

It is straightforward from Definition 22 that $\operatorname{ind}(T^n \boldsymbol{\tau}) = \operatorname{ind}(\boldsymbol{\tau})$ for $n \in \mathbb{Z}_L$. This is consistent with the fact that given $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ of size L and $n \in \mathbb{Z}_L$, there exists a constant $\alpha \neq 0$ with $\det([M^\ell T^k c]_{(k,\ell) \in (n,0)+\Lambda}) = \alpha \cdot \det([M^\ell T^k c]_{(k,\ell) \in \Lambda})$, where both sides are understood as polynomials in $c_0, \dots, c_{L-1}$ (cf. [20, Lemma 4.1]).

Example 23. Let $L = 3$. All the possible L-tuples $\boldsymbol{\tau} = (\tau_0, \tau_1, \tau_2)$ with $\tau_0 + \tau_1 + \tau_2 = 3$ are $(3, 0, 0)$, $(2, 1, 0)$, $(2, 0, 1)$, $(1, 1, 1)$, up to cyclic shifts. Consider any $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\boldsymbol{\tau}(\Lambda) = \boldsymbol{\tau}$.
(i) $\boldsymbol{\tau} = (3, 0, 0), (0, 3, 0), (0, 0, 3)$: The only monomial appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_\Lambda)$ is $c_0 c_1 c_2$, and $\operatorname{ind}(\boldsymbol{\tau}) = 0$.
(ii) $\boldsymbol{\tau} = (2, 1, 0), (0, 2, 1), (1, 0, 2)$: Monomials appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_\Lambda)$ are $c_0^2 c_1$, $c_0 c_2^2$, $c_1^2 c_2$, and $\operatorname{ind}(\boldsymbol{\tau}) = 1$.
(iii) $\boldsymbol{\tau} = (2, 0, 1), (1, 2, 0), (0, 1, 2)$: Monomials appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_\Lambda)$ are $c_0^2 c_2$, $c_0 c_1^2$, $c_1 c_2^2$, and $\operatorname{ind}(\boldsymbol{\tau}) = 2$.
(iv) $\boldsymbol{\tau} = (1, 1, 1)$: Monomials appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_\Lambda)$ are c_0^3, $c_0 c_1 c_2$, c_1^3, c_2^3, and $\operatorname{ind}(\boldsymbol{\tau}) = 0$.

We are now in position to formulate a sufficiency result for the identification of channels with constraints.

Theorem 24. *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 $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix}$ 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.*

Example 25 *(Example 18 revisited).* As seen in Example 18, the $2L \times L^2$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix}$ with

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

is rank deficient for all $\boldsymbol{c} \in \mathbb{C}^L$. Notice that since $\det(\boldsymbol{M}^{-k}) = (-1)^{k(L-1)}$ for $k = 0, \dots, L-1$, the conditions of Theorem 24 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 24 that the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}) \\ \boldsymbol{A} \end{bmatrix}$ has full rank for almost every $\boldsymbol{c} \in \mathbb{C}^L$.

We define a partial order on the set of all L-tuples by

$$\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_{L-1}) \preceq \boldsymbol{\tau}' = (\tau_0', \tau_1', \ldots, \tau_{L-1}') \quad \Leftrightarrow \quad \tau_j \le \tau_j',\ j = 0, \ldots, L-1.$$

Theorem 26. *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) $\mathrm{ind}(\boldsymbol{\tau}') \neq \mathrm{ind}(\boldsymbol{\tau}(\Lambda))$ *for every L-tuple $\boldsymbol{\tau}' \preceq \boldsymbol{\tau}(\widetilde{\Lambda})$ of size L different from* $\boldsymbol{\tau}(\Lambda)$.

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

To support this theorem, we give some examples.

Example 27. (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 $\boldsymbol{\tau}(\widetilde{\Lambda}) = (5,4,0,0,0)$ and the matrix $\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}$ is given by

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

The L-tuple $\boldsymbol{\tau} = (5,0,0,0,0)$ satisfies the conditions (i), (ii) of Theorem 26. Indeed, (i) is clear, and to verify (ii), let $\boldsymbol{\tau}' = (\tau_0', \tau_1', \tau_2', \tau_3', \tau_4') \preceq \boldsymbol{\tau}(\widetilde{\Lambda})$ with $\|\boldsymbol{\tau}'\|_1 = 5$, be an L-tuple different from $\boldsymbol{\tau}$. The only possible $\boldsymbol{\tau}'$ are $(1,4,0,0,0)$, $(2,3,0,0,0)$, $(3,2,0,0,0)$, $(4,1,0,0,0)$, whose respective indices are 4, 3, 2, 1, while $\mathrm{ind}(\boldsymbol{\tau}) = 0$. This verifies (ii). Theorem 26 then implies that with any full spark matrix $\boldsymbol{A}$ of size 4×9, the 9×9 matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{A} \end{bmatrix}$ 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 $\boldsymbol{\tau}(\widetilde{\Lambda}) = (2,1,1)$ 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 easy to check that the L-tuples $\boldsymbol{\tau} = (2,1,0)$ and $\boldsymbol{\tau} = (2,0,1)$ satisfy the conditions (i), (ii) of Theorem 26 (also see Example 23). Consequently, with any vector $a \in \mathbb{C}^4$ with no zero entries, the 4×4 matrix $\left[\begin{smallmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\tilde{\Lambda}} \\ \boldsymbol{a}^* \end{smallmatrix}\right]$ is invertible for almost every choice of $\boldsymbol{c}$ in $\mathbb{C}^3$.

4. Identification of MIMO channels under linear constraints

Let us now extend our SISO results to the multiple-input multiple-output (MIMO) setting. We shall first extend the results to the multiple-input single-output (MISO) setting and then to the MIMO setting.

Lemma 28. *Let $L \geq 2$, $N \geq 1$, and $\Lambda^{(1)}, \ldots, \Lambda^{(N)} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $L+1 \leq R = \sum_{n=1}^{N} |\Lambda^{(n)}| < 2L$. Then*

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

where $\mathcal{S}_N$ is the set of all $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that the matrix $[\boldsymbol{G}(\boldsymbol{c}^{(1)}) \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})]$ has full spark (cf. Theorem 7).

Theorem 29. *Let $L \geq 2$, $N \geq 1$, $\Lambda^{(1)}, \ldots, \Lambda^{(N)} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\sum_{n=1}^{N} |\Lambda^{(n)}| = L+1$, and $\boldsymbol{a} \in \mathbb{C}^{L+1}\backslash\{0\}$. There exists $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that 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. Moreover, such vectors $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})$ constitute a dense open subset of $(\mathbb{C}^L)^N$ with full measure. Hence, the MISO operator Paley-Wiener space $OPW(\boldsymbol{\Lambda}) = [\,OPW(\Lambda^{(n)})\,]_{n=1}^{N}$ with side constraints $b = \boldsymbol{a}^*\boldsymbol{\eta}$, where $b \in \mathbb{C}$, $\boldsymbol{\eta} = [\boldsymbol{\eta}^{(n)}|_{\Lambda^{(n)}}]_{n=1}^{N}$ and $\boldsymbol{\eta}^{(n)} \in \mathbb{C}^{L^2}$ for $n = 1, \ldots, N$, is identifiable.*

As seen in Theorem 6, the MISO operator Paley-Wiener space $OPW(\boldsymbol{\Lambda}) = [\,OPW(\Lambda^{(n)})\,]_{n=1}^{N}$ is identifiable if and only if $\sum_{n=1}^{N} |\Lambda^{(n)}| \leq L$. Theorem 29 shows that one can overcome such limitations on the total size of $\Lambda^{(1)}, \ldots, \Lambda^{(N)}$ by taking linear side constraints into account.

Finally, we generalize the result to the MIMO setting.

Theorem 30. *Let $L \geq 2$, $M, N \geq 1$, $\boldsymbol{\Lambda} = [\Lambda_{m,n}]_{m=1}^{M}{}_{n=1}^{N}$ with $\Lambda_{m,n} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ and $\sum_{m=1}^{M} \sum_{n=1}^{N} |\Lambda_{m,n}| = L+1$, and $\boldsymbol{a} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M)^{\mathrm{T}} \in \mathbb{C}^{L+1}\backslash\{0\}$, where $\boldsymbol{a}_m$ is a vector of dimension $\sum_{n=1}^{N} |\Lambda_{m,n}|$. There exists $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that the $(L+1)\times(L+1)$ matrix*

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_M} \\ \boldsymbol{a}_1^* & \boldsymbol{a}_2^* & \cdots & \boldsymbol{a}_M^* \end{bmatrix} \tag{13}$$

where $\Lambda_m = \{\Lambda_{m,n}\}_{n=1}^{N} \subset (\mathbb{Z}_L \times \mathbb{Z}_L)^N$, *is invertible. Moreover, such vectors* $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})$ *constitute a dense open subset of* $(\mathbb{C}^L)^N$ *with full measure. Hence, the MIMO operator Paley-Wiener space* $OPW(\boldsymbol{\Lambda}) = [OPW(\Lambda_{m,n})]_{m=1}^{M}{}_{n=1}^{N}$ *with side constraints* $b = \sum_{m=1}^{M} \boldsymbol{a}_m^* \boldsymbol{\eta}_m|_{\Lambda_m}$, *where* $b \in \mathbb{C}$ *and* $\boldsymbol{\eta}_m = \{\boldsymbol{\eta}_{m,n}\}_{n=1}^{N} \in (\mathbb{C}^{L^2})^N$ *for* $m = 1, \ldots, M$, *is identifiable.*

Concerning side constraints of multiple equations, Example 17 in the SISO setting clearly indicates that Theorems 29 and 30 cannot be extended to linear side constraints of two or more equations.

5. Applications to continuous-time channel identification and operator sampling

In this section, we discuss the channel identification problem in the continuous-time setting where the subchannels are represented by operators that act on a function space over $\mathbb{R}$. As we shall see, results established in the finite-dimensional (discrete-time) setting can be carried over to the continuous-time setting in a straightforward manner.

5.1. Continuous-time SISO channels

In the continuous-time setting, linear time-variant SISO communication channels are modeled by Hilbert-Schmidt operators of the form

$$\begin{aligned} Hf(x) &= \iint \eta_H(t,\nu)\, M_\nu T_t f(x)\, d\nu\, dt \\ &= \int \sigma_H(x,\xi)\, e^{2\pi i x\xi}\, \widehat{f}(\xi)\, d\xi, \qquad f \in L^2(\mathbb{R}), \end{aligned} \tag{14}$$

where T_t is translation (time shift) by $t \in \mathbb{R}$, that is, $T_t f(x) = f(x-t)$, M_ν is modulation (frequency shift) by $\nu \in \mathbb{R}$, that is, $M_\nu f(x) = e^{2\pi i \nu x} f(x)$, and $\widehat{f}(\xi) = \int f(x)\, e^{-2\pi i x\xi}\, dx$ is the Fourier transform of f. The *spreading function* $\eta_H \in L^2(\mathbb{R}^2)$ and the *Kohn-Nirenberg symbol* $\sigma_H \in L^2(\mathbb{R}^2)$ of H are related by

$$\sigma_H = \mathcal{F}_s \eta_H \quad (\text{equivalently,} \ \ \eta_H = \mathcal{F}_s \sigma_H),$$

where $\mathcal{F}_s$ denotes the *symplectic Fourier transform* defined as

$$(\mathcal{F}_s g)(x,\xi) = \iint g(t,\nu)\, e^{-2\pi i(\nu x - \xi t)}\, dt\, d\nu.$$

Here the operator H being *Hilbert-Schmidt* corresponds to the fact that $\eta_H \in L^2(\mathbb{R}^2)$, or equivalently, that $\sigma_H \in L^2(\mathbb{R}^2)$. Certainly, an operator of the form (14) is a straightforward generalization of (3) to the continuous-time setting.

Similar to the (discrete) support restrictions made for the vector $\boldsymbol{\eta} = \{\eta(k,\ell)\}_{k,\ell=0}^{L-1}$ in Definition 2, we consider (continuous) support restrictions on the spreading function

$\eta_H = \mathcal{F}_s\sigma_H \in L^2(\mathbb{R}^2)$. The *SISO operator Paley-Wiener space* for a set $S \subset \mathbb{R}^2$ is defined by

$$OPW(S) = \{H \in \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R})) : \eta_H \in L^2(\mathbb{R}^2),\ \operatorname{supp} \mathcal{F}_s\sigma_H = \operatorname{supp} \eta_H \subset S\}, \tag{15}$$

which consists of Hilbert-Schmidt operators on $L^2(\mathbb{R})$ whose Kohn-Nirenberg symbol is *bandlimited* to $S \subset \mathbb{R}^2$. In order to avoid pathological sets for $S \subset \mathbb{R}^2$, we shall assume that S is a Jordan domain, that is, a bounded set whose boundary is a Lebesgue zero set. It should be noted that every operator H in $OPW(S)$ is defined a priori on $L^2(\mathbb{R})$, however, its domain can be extended to classes of tempered distributions [21].

We say that a class of operators $\mathcal{H} \subset \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))$ is *identifiable* if there exist a tempered distribution g and constants $0 < A \leq B < \infty$ such that

$$A\|H_1 - H_2\|_{\mathcal{H}} \leq \|H_1 g - H_2 g\|_{L^2(\mathbb{R})} \leq B\|H_1 - H_2\|_{\mathcal{H}}, \quad H_1, H_2 \in \mathcal{H}.$$

If $\mathcal{H}$ is a linear space, this reduces to

$$A\|H\|_{\mathcal{H}} \leq \|Hg\|_{L^2(\mathbb{R})} \leq B\|H\|_{\mathcal{H}}, \quad H \in \mathcal{H}. \tag{16}$$

We refer to operator identification as *operator sampling* if g is a discretely supported distribution, in particular, as *regular operator sampling* if there exists an identifier of the form $g = \sum_{n\in\mathbb{Z}} c_n \delta_{nT}$ for some $T > 0$ and an L-periodic sequence $c = \{c_n\}_{n\in\mathbb{Z}}$ and where δ_{nT} stands for the delta distribution.

The next theorem shows that the identifiability of $OPW(S)$ depends essentially on the size of $S \subset \mathbb{R}^2$.

Theorem 31 *([22,1]). Let $S \subset \mathbb{R}^2$ be a Jordan domain. The SISO operator Paley-Wiener space $OPW(S)$ is identifiable by regular operator sampling if $|S| < 1$, and not identifiable if $|S| > 1$.*

This result is a continuous-time analogue of Corollary 4. In fact, one can immediately deduce the first part of Theorem 31 from Corollary 4. The task involves a discretization, namely, the procedure of finding an appropriate rectification of S and then sampling with respect to the rectification. This reduces the continuous-time equation to a family of $L\times L^2$ linear systems of the form (4) each of which is indexed by (t, ν). We postpone the detailed arguments to Section 7.6. Let us mention that this technique, in general, allows us to carry over results in the finite-dimensional SISO setting to the continuous-time SISO setting in a straightforward way.

5.2. Continuous-time MIMO channels

The corresponding formulations for the continuous-time MIMO setting are as follows. In the N-input M-output case, MIMO channel operators are of the form $\boldsymbol{H} : (L^2(\mathbb{R}))^N \to (L^2(\mathbb{R}))^M$ given by

$$\boldsymbol{H}\begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} = \begin{pmatrix} H_{1,1} & \dots & H_{1,N} \\ \vdots & & \vdots \\ H_{M,1} & \dots & H_{M,N} \end{pmatrix}\begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} = \begin{pmatrix} \sum_{n=1}^N H_{1,n} f_n \\ \vdots \\ \sum_{n=1}^N H_{M,n} f_n, \end{pmatrix},$$

where each subchannel $H_{m,n} \in \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))$ is of the form (14). We shall represent such an operator by the block matrix

$$\boldsymbol{H} = \begin{pmatrix} H_{1,1} & \dots & H_{1,N} \\ \vdots & & \vdots \\ H_{M,1} & \dots & H_{M,N} \end{pmatrix},$$

and define its spreading function and spreading support componentwise, that is,

$$\boldsymbol{\eta}(\boldsymbol{H}) = \begin{pmatrix} \eta(H_{1,1}) & \cdots & \eta(H_{1,N}) \\ \vdots & & \vdots \\ \eta(H_{M,1}) & \cdots & \eta(H_{M,N}) \end{pmatrix},$$

and

$$\operatorname{supp} \boldsymbol{\eta}(\boldsymbol{H}) = \begin{pmatrix} \operatorname{supp}\ \eta(H_{1,1}) & \cdots & \operatorname{supp}\ \eta(H_{1,N}) \\ \vdots & & \vdots \\ \operatorname{supp}\ \eta(H_{M,1}) & \cdots & \operatorname{supp}\ \eta(H_{M,N}) \end{pmatrix} \subset (\mathbb{R}^2)^{M\times N}.$$

The *MIMO operator Paley-Wiener space* for a set $\boldsymbol{S} = \{S_{m,n}\}_{m=1}^{M}{}_{n=1}^{N} \subset (\mathbb{R}^2)^{M\times N}$ is defined by

$$OPW(\boldsymbol{S}) = \{\boldsymbol{H} \in \mathcal{L}((L^2(\mathbb{R}))^N, (L^2(\mathbb{R}))^M) : \boldsymbol{\eta}(\boldsymbol{H}) \in L^2(\mathbb{R}^2)^{M\times N},\ \operatorname{supp} \boldsymbol{\eta}(\boldsymbol{H}) \subset \boldsymbol{S}\Big\}.$$

As in the SISO case, we say that a class of operators $\mathcal{H} \subset \mathcal{L}((L^2(\mathbb{R}))^N, (L^2(\mathbb{R}))^M)$ is *identifiable* if there exist a vector $\boldsymbol{g} = (g_1, \dots, g_N)$ of tempered distributions and constants $0 < A \leq B < \infty$ such that

$$A\|\boldsymbol{H}_1 - \boldsymbol{H}_2\|_{\mathcal{H}} \leq \|\boldsymbol{H}_1 g - \boldsymbol{H}_2 g\|_{L^2(\mathbb{R})} \leq B\|\boldsymbol{H}_1 - \boldsymbol{H}_2\|_{\mathcal{H}}, \quad \boldsymbol{H}_1, \boldsymbol{H}_2 \in \mathcal{H}.$$

If $\mathcal{H}$ is a linear space, this reduces to

$$A\|\boldsymbol{H}\|_{\mathcal{H}} \leq \|\boldsymbol{H} g\|_{L^2(\mathbb{R})} \leq B\|\boldsymbol{H}\|_{\mathcal{H}}, \quad H \in \mathcal{H}.$$

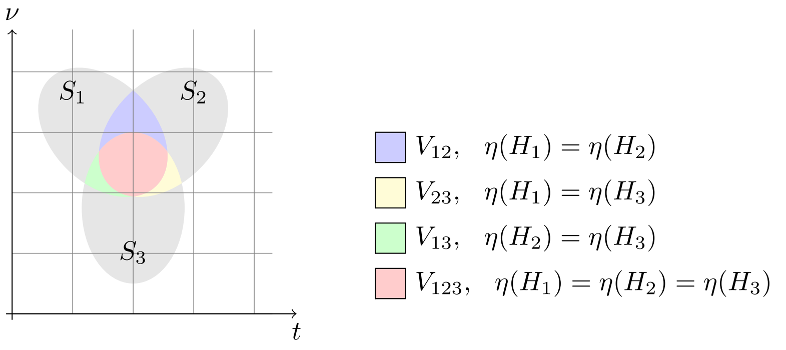

Fig. 1. An example of continuous-time channels with partial linear relation. (For interpretation of the colors in the figure, the reader is referred to the web version of this article.)

In the continuous-time SISO setting, we have already seen that Theorem 31 is a continuous-time analogue of Corollary 4. Similarly, in the continuous-time MIMO setting we have the following as a counterpart of Theorem 6.

Theorem 32 *([7]). Let $\boldsymbol{S} = \{S_{m,n}\}_{m=1}^{M}{}_{n=1}^{N} \subset (\mathbb{R}^2)^{M\times N}$ where each $S \subset \mathbb{R}^2$ is a Jordan domain. If $\sum_{n=1}^{N} \mu(S_{m,n}) < 1$ for $m = 1, \ldots, M$, then $OPW(\boldsymbol{S})$ is identifiable. If $\sum_{n=1}^{N} \mu(S_{m,n}) > 1$ for some m, then $OPW(\boldsymbol{S})$ is not identifiable.*

Using the discretization technique described in Section 5.1, one can also carry over results in the finite-dimensional MIMO setting to the continuous-time MIMO setting. We refer to [17] for detailed arguments. Instead of giving details here, we present a simple example which illustrates the benefit of having linear side constraints in the continuous-time channel identification problem.

Example 33. Consider a three-input single-output channel $\boldsymbol{H} = (H_{1,1}, H_{1,2}, H_{1,3})$ whose subchannels $H_{1,1} \in OPW(S_{1,1})$, $H_{1,2} \in OPW(S_{1,2})$, $H_{1,3} \in OPW(S_{1,3})$ have the property of sharing the same spreading function on their common support sets, that is, $\eta(H_{1,j}) = \eta(H_{1,k})$ on the set $S_{1,j} \cap S_{1,k}$. For notational simplicity, we shall write $H_1 := H_{1,1}$, $S_1 := S_{1,1}$, and so on. It turns out that $\boldsymbol{H} = (H_1, H_2, H_3)$ is identifiable if $\mu(S_1 \cup S_2 \cup S_3) < 1$ and not identifiable if $\mu(S_1 \cup S_2 \cup S_3) > 1$. Clearly, $\mu(S_1 \cup S_2 \cup S_3) \le \mu(S_1) + \mu(S_2) + \mu(S_3)$ where the inequality is strict if any two of the sets S_1, S_2, S_3 have an intersection of positive measure.

Let us give an intuitive explanation and also draw a connection with the finite-dimensional setting. To this end, we partition the support sets S_1, S_2, S_3 as

$$S_1 = V_{123} \uplus V_{12} \uplus V_{13} \uplus V_1,$$

$$S_2 = V_{123} \uplus V_{12} \uplus V_{23} \uplus V_2,$$

$$S_3 = V_{123} \uplus V_{13} \uplus V_{23} \uplus V_3,$$

as shown in Fig. 1, and decompose the operators H_1, H_2, H_2 as

$$\begin{aligned}H_1 &= F_{123} + F_{12} + F_{13} + F_1,\\ H_2 &= F_{123} + F_{12} + F_{23} + F_2,\\ H_3 &= F_{123} + F_{13} + F_{23} + F_3,\end{aligned}$$

where $F_{123} \in OPW(V_{123})$, $F_{12} \in OPW(V_{12})$, $F_{13} \in OPW(V_{13})$, $F_{23} \in OPW(V_{23})$, $F_1 \in OPW(V_1)$, $F_2 \in OPW(V_2)$, $F_3 \in OPW(V_3)$. The response of $\boldsymbol{H} = (H_1, H_2, H_3)$ to the input $\boldsymbol{g} = (g^{(1)}, g^{(2)}, g^{(3)})$ is then given by

$$\begin{aligned}&H_1\, g^{(1)} + H_2\, g^{(2)} + H_3\, g^{(3)}\\ &\quad= F_{123}(g^{(1)} + g^{(2)} + g^{(3)}) + F_{12}(g^{(1)} + g^{(2)}) + F_{13}(g^{(1)} + g^{(3)})\\ &\qquad+ F_{23}(g^{(2)} + g^{(3)}) + F_1 g^{(1)} + F_2 g^{(2)} + F_3 g^{(3)}.\end{aligned}$$

This reformulates the problem of identifying $\boldsymbol{H}$ with the input $\boldsymbol{g}$ to identifying the synthetic 7-input one-output channel $\boldsymbol{F} = (F_{123}, F_{12}, F_{13}, F_{23}, F_1, F_2, F_3)$ with the input $\widetilde{\boldsymbol{g}} = (g^{(1)} + g^{(2)} + g^{(3)}, g^{(1)} + g^{(2)}, g^{(1)} + g^{(3)}, g^{(2)} + g^{(3)}, g^{(1)}, g^{(2)}, g^{(3)})$. In the latter setup, all subchannels of $\boldsymbol{F}$ have disjoint spreading support, but instead the input $\widetilde{\boldsymbol{g}}$ has a limitation on its choice. It turns out that such a limitation hardly affects the identifiability of $\boldsymbol{F}$ and one can deduce that, similarly as in Theorem 32, $\boldsymbol{F}$ is identifiable if $\mu(V_{123}) + \mu(V_{12}) + \mu(V_{13}) + \mu(V_{23}) + \mu(V_1) + \mu(V_2) + \mu(V_3) < 1$ and not identifiable if $\mu(V_{123}) + \mu(V_{12}) + \mu(V_{13}) + \mu(V_{23}) + \mu(V_1) + \mu(V_2) + \mu(V_3) > 1$, that is, $\boldsymbol{H} = (H_1, H_2, H_3)$ is identifiable if $\mu(S_1 \cup S_2 \cup S_3) < 1$ and not identifiable if $\mu(S_1 \cup S_2 \cup S_3) > 1$.

Translating the problem again to the finite-dimensional setting (based on the rectification technique detailed in Section 7.6), the identifiability of $\boldsymbol{F}$ is equivalent to the solvability of the linear system

$$\begin{aligned}\boldsymbol{y} = \Big[&G(\boldsymbol{c}^{(1)} + \boldsymbol{c}^{(2)} + \boldsymbol{c}^{(3)})\;\; G(\boldsymbol{c}^{(1)} + \boldsymbol{c}^{(2)})\;\; G(\boldsymbol{c}^{(1)} + \boldsymbol{c}^{(3)})\;\; G(\boldsymbol{c}^{(2)} + \boldsymbol{c}^{(3)})\\ &\ldots\;\; G(\boldsymbol{c}^{(1)})\;\; G(\boldsymbol{c}^{(2)})\;\; G(\boldsymbol{c}^{(3)})\Big] \begin{bmatrix}\eta_{123}\\ \eta_{12}\\ \eta_{13}\\ \eta_{23}\\ \eta_1\\ \eta_2\\ \eta_3\end{bmatrix},\end{aligned}$$

where $\eta_{123}, \eta_{12}, \ldots, \eta_3 \in \mathbb{C}^{L^2}$ are the L^2-dimensional unknown vectors and $\boldsymbol{c}^{(1)} = \{c_k^{(1)}\}_{k=0}^{L-1}$, $\boldsymbol{c}^{(2)} = \{c_k^{(2)}\}_{k=0}^{L-1}$, $\boldsymbol{c}^{(3)} = \{c_k^{(3)}\}_{k=0}^{L-1}$ are the L-periodic coefficients of the delta trains $g^{(1)}$, $g^{(2)}$, $g^{(3)}$, respectively, that is, for some $T > 0$, $g^{(n)} = \sum_{k\in\mathbb{Z}} c_k^{(n)}\, \delta_{kT}$, $n = 1, 2, 3$. If the matrix associated with the linear system were to have full spark, then the system is uniquely solvable if and only if the vector $(\eta_{123}, \eta_{12}, \eta_{13}, \eta_{23}, \eta_1, \eta_2, \eta_3)^{\mathrm{T}}$ has a known support of size at most L. However, the matrix does not have full spark, indeed, the first column of $G(\boldsymbol{c}^{(1)}+\boldsymbol{c}^{(2)}+\boldsymbol{c}^{(3)})$ is a linear combination of the first columns of $G(\boldsymbol{c}^{(1)})$, $G(\boldsymbol{c}^{(2)})$, $G(\boldsymbol{c}^{(3)})$. Nevertheless, due to the construction of $V_{123}, V_{12}, \dots, V_3$, the vectors η_{123}, η_{12}, $\dots$, $\eta_3 \in \mathbb{C}^{L^2}$ are disjointly supported and therefore the full spark property is not necessary. In fact, to establish the desired identifiability result one only needs that columns of

$$\Big[G(\boldsymbol{c}^{(1)}+\boldsymbol{c}^{(2)}+\boldsymbol{c}^{(3)})|_{\Lambda_{123}}\;\; G(\boldsymbol{c}^{(1)}+\boldsymbol{c}^{(2)})|_{\Lambda_{12}}\;\; G(\boldsymbol{c}^{(1)}+\boldsymbol{c}^{(3)})|_{\Lambda_{13}}$$
$$\dots\;\; G(\boldsymbol{c}^{(2)}+\boldsymbol{c}^{(3)})|_{\Lambda_{23}}\;\; G(\boldsymbol{c}^{(1)})|_{\Lambda_1}\;\; G(\boldsymbol{c}^{(2)})|_{\Lambda_2}\;\; G(\boldsymbol{c}^{(3)})|_{\Lambda_3}\Big]$$

with $\Lambda_{123}, \Lambda_{12}, \Lambda_{13}, \Lambda_{23}, \Lambda_1, \Lambda_2, \Lambda_3 \subset \mathbb{Z}_L \times \mathbb{Z}_L$ being disjoint sets of total size L, are linearly independent. Theorem 8 implies that there exist vectors $\boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}, \boldsymbol{c}^{(3)}$ $\in \mathbb{C}^L$ with this property.

Remark 34. (i) Explicit reconstruction formulas for the continuous-time SISO channel identification are derived in [18] and generalizations to the MIMO setting are addressed in [17].

(ii) Concerning the regular operator sampling of $OPW(S)$, Pfander and Walnut [18] obtained a slightly more general result than Theorem 31 which also covers the critical area case $|S| = 1$. Their result characterizes all spaces $OPW(S)$ that are identifiable by regular operator sampling, in terms of a dual tiling condition on the set S.

(iii) It is possible to extend the space $OPW(S)$ by removing the condition $\sigma_H \in L^2(\mathbb{R}^2)$ from (15). If $\{0\}\times[-\Omega/2, \Omega/2] \subset S$ for some $\Omega > 0$, then the extended space contains the multiplication operators defined by $Hf(x) = m(x)\,f(x)$, where $m(x)$ is an arbitrary function in the classical Paley-Wiener space

$$PW(\Omega) = \{f \in L^2(\mathbb{R}) : \operatorname{supp} \widehat{f} \subset [-\Omega/2, \Omega/2]\}.$$

Such operators can be represented in the form (14) with $\eta_H(t,\nu) = \delta(t)\,\widehat{m}(\nu)$ and $\sigma_H(x,\xi) = m(x)$, neither of which is in $L^2(\mathbb{R}^2)$. Using reconstruction formulas developed for operator sampling of the extended $OPW(S)$, one can immediately deduce the classical sampling reconstruction formula for $PW(\Omega)$. For further details, we refer the reader to [18].

6. Conclusion

This paper studies the channel identification problem for multiple-input multiple-output (MIMO) channels under the assumption that subchannels have some known linear relationship. We formulate the linear relations in terms of linear constraints which can express relations between and within subchannels.

In the single-input single-output (SISO) setting, we show that preknowledge on linear side constraints consisting of a single equation allows the identification of $OPW(\Lambda)$ with

$|\Lambda| = L+1$. By taking into account the constraints, this result overcomes the fundamental limitation that $OPW(\Lambda)$ is identifiable if and only if $|\Lambda| \le L$. Although this result might not be surprising, the proof requires a careful analysis on Gabor submatrices which makes the task surprisingly challenging. The result however does not extend to the case of linear side constraints consisting of multiple equations, in particular, we give an example of linear constraints consisting of two equations such that the space $OPW(\Lambda)$ under the constraints is not identifiable for every Λ with $|\Lambda| = L+2$. Nevertheless, we provide some sufficient conditions for linear constraints of k equations so that $OPW(\Lambda)$ with $|\Lambda| = L+k$ is identifiable. We then extend the SISO results to the MIMO setting. We also discuss the applicability of our results in the continuous-time setting.

7. Proofs

7.1. Proof of Assertions in Example 13

(i) Consider a two-input single-output channel $\boldsymbol{H} = (\boldsymbol{H}_{1,1}, \boldsymbol{H}_{1,2})$ with $(0,0) \in \Lambda_{1,1} \cap \Lambda_{1,2}$ and $\boldsymbol{\eta}_{1,1}(0,0) = \boldsymbol{\eta}_{1,2}(0,0)$. Such constraints can be represented by the equation $\boldsymbol{b}_1 = \boldsymbol{A}_1\boldsymbol{\eta}_1$, where $\boldsymbol{A}_1 = [\,1,\,0,\,\dots,\,0\,|\,-1,\,0,\,\dots,\,0\,] \in \mathbb{C}^{1\times 2L^2}$, $\boldsymbol{\eta}_1 = \begin{bmatrix}\boldsymbol{\eta}_{1,1}\\ \boldsymbol{\eta}_{1,2}\end{bmatrix}$, and $\boldsymbol{b}_1 = 0$. By Gaussian elimination, we have

$$\begin{aligned}\boldsymbol{B} &:= \begin{bmatrix}\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}} & \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda_{1,2}}\\ 1,\,0,\,\dots,\,0 & -1,\,0,\,\dots,\,0\end{bmatrix}\\ &= \begin{bmatrix}\boldsymbol{c}^{(1)} & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\setminus(0,0)} & \boldsymbol{c}^{(2)} & \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda_{1,2}\setminus(0,0)}\\ 1 & 0,\,\dots,\,0 & -1 & 0,\,\dots,\,0\end{bmatrix}\\ &\sim \begin{bmatrix}0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\setminus(0,0)} & \boldsymbol{c}^{(1)}+\boldsymbol{c}^{(2)} & \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda_{1,2}\setminus(0,0)}\\ 1 & 0,\,\dots,\,0 & -1 & 0,\,\dots,\,0\end{bmatrix},\end{aligned}$$

which implies that $\boldsymbol{B}$ is injective if and only if the matrix

$$\left[\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\setminus(0,0)} \mid (\boldsymbol{c}^{(1)}+\boldsymbol{c}^{(2)}) \mid \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda_{1,2}\setminus(0,0)}\right]$$

is injective. Note that this matrix is of the form (8) in Theorem 8 with $\boldsymbol{\alpha} = (1,1)^{\mathrm{T}}$ and $\Lambda = \{(0,0)\}$. With $(\boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}) \in (\mathbb{C}^L)^2$ chosen from the corresponding set $\mathcal{S}_{2,\boldsymbol{\alpha}}$ (defined in Theorem 8), the matrix is injective if and only if $|\Lambda_{1,1}\setminus(0,0)|+|\Lambda_{1,2}\setminus(0,0)|+1 \le L$, i.e., $|\Lambda_{1,1}| + |\Lambda_{1,2}| \le L+1$. Since the $(L+1)\times(|\Lambda_{1,1}| + |\Lambda_{1,2}|)$ matrix $\boldsymbol{B}$ cannot be injective for $|\Lambda_{1,1}| + |\Lambda_{1,2}| > L+1$, we conclude that $OPW(\boldsymbol{\Lambda})$ with such side constraints is identifiable if and only if $|\Lambda_{1,1}| + |\Lambda_{1,2}| \le L+1$.

(ii) Consider a single-input two-output channel $\boldsymbol{H} = (\boldsymbol{H}_{1,1}, \boldsymbol{H}_{2,1}) \in OPW(\boldsymbol{\Lambda})$ whose subchannels $\boldsymbol{H}_{1,1}$ and $\boldsymbol{H}_{2,1}$ are known to be partially identical, i.e., $\boldsymbol{\eta}_{1,1}|_S = \boldsymbol{\eta}_{2,1}|_S$ for some $S \subset \Lambda_{1,1} \cap \Lambda_{2,1}$. Such constraints can be represented by the equation $\boldsymbol{b}_1 = \boldsymbol{A}_1\boldsymbol{\eta}_{1,1} + \boldsymbol{A}_2\boldsymbol{\eta}_{2,1}$, where $\boldsymbol{A}_1 = [0\ I_S\ 0] \in \mathbb{C}^{S\times L^2}$, $\boldsymbol{A}_2 = [0\ -I_S\ 0] \in \mathbb{C}^{S\times L^2}$ and $\boldsymbol{b}_1 = 0$. Applying the Gaussian elimination, we have

$$\boldsymbol{B} := \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}} & 0 \\ 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}} \\ \boldsymbol{A}_1|_{\Lambda_{1,1}} & \boldsymbol{A}_2|_{\Lambda_{2,1}} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_S & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & 0 & 0 \\ 0 & 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_S & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \\ I_S & 0 & -I_S & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & 0 & -\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \\ 0 & 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_S & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \\ I_S & 0 & -I_S & 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & 0 & -\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \\ \boldsymbol{G}(\boldsymbol{c}^{(1)})|_S & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & 0 & 0 \\ I_S & 0 & -I_S & 0 \end{bmatrix}.$$

This shows that $\boldsymbol{B}$ is injective if and only if the $2L\times(|S|+|\Lambda_{1,1}\backslash S|+|\Lambda_{2,1}\backslash S|)$ submatrix

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & 0 & -\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \\ 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_S & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \end{bmatrix}$$

is injective if and only if the $2L\times(|S|+|\Lambda_{1,1}\backslash S|+|\Lambda_{2,1}\backslash S|)$ submatrix

$$\begin{bmatrix} 0 & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & -\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{2,1}\backslash S} \\ \boldsymbol{G}(\boldsymbol{c}^{(1)})|_S & \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{1,1}\backslash S} & 0 \end{bmatrix}$$

is injective. With $\boldsymbol{c}^{(1)} \in \mathbb{C}^L$ chosen so that $\boldsymbol{G}(\boldsymbol{c}^{(1)})$ has full spark (cf. Theorem 3), the first submatrix is injective if either $\max\{|\Lambda_{1,1}\backslash S|, |\Lambda_{2,1}|\} \leq L$ or $\max\{|\Lambda_{1,1}\backslash S| + |\Lambda_{2,1}\backslash S|, |S|\} \leq L$; the second submatrix is injective if either $\max\{|\Lambda_{1,1}\backslash S| + |\Lambda_{2,1}\backslash S|, |S|\} \leq L$ or $\max\{|\Lambda_{1,1}|, |\Lambda_{2,1}\backslash S|\} \leq L$. (Note that these conditions are sufficient but not necessary for the respective injectivity.) Consequently, $\boldsymbol{B}$ is injective if one of the following holds:

(i) $\max\{|\Lambda_{1,1}\backslash S|, |\Lambda_{2,1}|\} \leq L$,
(ii) $\max\{|\Lambda_{1,1}|, |\Lambda_{2,1}\backslash S|\} \leq L$,
(iii) $\max\{|\Lambda_{1,1}\backslash S| + |\Lambda_{2,1}\backslash S|, |S|\} \leq L$.

Note that in any case, we have $|\Lambda_{1,1}\backslash S| + |\Lambda_{2,1}\backslash S| + |S| \leq 2L$, as indicated by the size of $\boldsymbol{B} \in \mathbb{C}^{(2L+|S|)\times(|\Lambda_{1,1}|+|\Lambda_{2,1}|)}$. Therefore, we conclude that $OPW(\boldsymbol{\Lambda})$ under such side constraints is identifiable if one of the conditions (i)-(iii) is satisfied.

7.2. Proof of Theorem 7 and Theorem 8

For the proofs, we need the following lemma.

Lemma 35. *Let $\Lambda^{(1)}, \ldots, \Lambda^{(N)} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\sum_{n=1}^{N} |\Lambda^{(n)}| < 2L$, where $L, N \geq 2$. There exist $(k_n, \ell_n) \in \mathbb{Z}_L \times \mathbb{Z}_L$, $n = 1, \ldots, N$, such that the sets $\Lambda^{(n)} + (k_n, \ell_n)$, $n = 1, \ldots, N$ are mutually disjoint.*

The bound $2L$ in Lemma 35 is sharp because e.g., any translation of two sets $\Lambda^{(1)} = \{0\} \times \mathbb{Z}_L$ and $\Lambda^{(2)} = \mathbb{Z}_L \times \{0\}$ always intersect.

Proof of Lemma 35. The proof is by induction.

(i) First, we show that the claim holds for $N = 2$. Let $\Lambda, \Lambda' \subset \mathbb{Z}_L \times \mathbb{Z}_L$ be such that $|\Lambda| + |\Lambda'| < 2L$. The set

$$\begin{aligned} &\{(k,\ell) \in \mathbb{Z}_L \times \mathbb{Z}_L : \Lambda \cap (\Lambda' + (k,\ell)) \neq \emptyset\} \\ &= \bigcup_{(p,q) \in \Lambda'} \{(k,\ell) \in \mathbb{Z}_L \times \mathbb{Z}_L : (p,q) + (k,\ell) \in \Lambda\} \\ &= \bigcup_{(p,q) \in \Lambda'} (\Lambda - (p,q)) \end{aligned}$$

has cardinality at most $|\Lambda| \cdot |\Lambda'| \leq (|\Lambda| + |\Lambda'|)^2/4 < L^2$ and is therefore a proper subset of $\mathbb{Z}_L \times \mathbb{Z}_L$. For any element (k, ℓ) lying outside this set, we have $\Lambda \cap (\Lambda' + (k, \ell)) = \emptyset$.

(ii) Let $\Lambda^{(1)}, \ldots, \Lambda^{(N)} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\sum_{n=1}^{N} |\Lambda^{(n)}| < 2L$, where $L, N \geq 2$. For $2 \leq K \leq N - 1$ fixed, suppose that $\Lambda^{(n)} + (k_n, \ell_n)$, $n = 1, \ldots, K$ are mutually disjoint for some $\{(k_n, \ell_n)\}_{n=1}^{K} \subset \mathbb{Z}_L \times \mathbb{Z}_L$. Applying (i) to the sets $\Lambda = \cup_{n=1}^{K} (\Lambda^{(n)} + (k_n, \ell_n))$ and $\Lambda' = \Lambda^{(K+1)}$, we find $(k_{K+1}, \ell_{K+1}) \in \mathbb{Z}_L \times \mathbb{Z}_L$ such that $\Lambda \cap (\Lambda' + (k_{K+1}, \ell_{K+1})) = \emptyset$. Then $\Lambda^{(n)} + (k_n, \ell_n)$, $n = 1, \ldots, K+1$ are mutually disjoint.

From (i) and (ii), we conclude that there exists $\{(k_n, \ell_n)\}_{n=1}^{N} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ such that the sets $\Lambda^{(n)} + (k_n, \ell_n)$ are mutually disjoint. $\square$

Proof of Theorem 7. We will only prove the case $N = 2$, as the generalization to $N \geq 3$ is straightforward. For brevity of notation, we write the matrix as $[\boldsymbol{G}(\boldsymbol{c}) \mid \boldsymbol{G}(\boldsymbol{d})] \in \mathbb{C}^{L \times 2L^2}$, where $\boldsymbol{c} = (c_0, \cdots, c_{L-1})^{\mathrm{T}}$, $\boldsymbol{d} = (d_0, \cdots, d_{L-1})^{\mathrm{T}} \in \mathbb{C}^L$, and fix any $\Lambda, \Lambda' \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $|\Lambda| + |\Lambda'| = L$. By Lemma 35, there exists $(a, b) \in \mathbb{Z}_L \times \mathbb{Z}_L$ for which the sets Λ and $\Lambda' + (a, b)$ are disjoint.

Using the commutation relation $\boldsymbol{T}^p \boldsymbol{M}^q = \omega^{pq} \boldsymbol{M}^q \boldsymbol{T}^p$, we have

$$\boldsymbol{M}^{\ell} \boldsymbol{T}^k \boldsymbol{d} = \omega^{(k+a)b} \boldsymbol{M}^{\ell+b} \boldsymbol{T}^{k+a} (\boldsymbol{M}^{-b} \boldsymbol{T}^{-a} \boldsymbol{d}), \quad k, \ell = 0, \ldots, L-1.$$

Applying this to every (k, ℓ) in Λ' and collecting the phase factors, we find $m \in \{0, \ldots, L-1\}$ such that

$$\det[\boldsymbol{G}(\boldsymbol{c})|_{\Lambda} \ \boldsymbol{G}(\boldsymbol{d})|_{\Lambda'}] = \omega^m \det[\boldsymbol{G}(\boldsymbol{c})|_{\Lambda} \ \boldsymbol{G}(\boldsymbol{M}^{-b} \boldsymbol{T}^{-a} \boldsymbol{d})|_{\Lambda' + (a,b)}].$$

Set $\boldsymbol{f} = (f_0, f_1, \ldots, f_{L-1}) = \boldsymbol{M}^{-b} \boldsymbol{T}^{-a} \boldsymbol{d}$. Since $\Lambda \cap (\Lambda' + (a, b)) = \emptyset$, Theorem 3 implies that $\det[\boldsymbol{G}(\boldsymbol{c})|_{\Lambda} \ \boldsymbol{G}(\boldsymbol{c})|_{\Lambda' + (a,b)}]$ is a *nontrivial* homogeneous polynomial of degree L in the variables $(c_0, \ldots, c_{L-1})$. Replacing the $\boldsymbol{c}$ in the second submatrix with $\boldsymbol{f}$, we note that $\det[\boldsymbol{G}(\boldsymbol{c})|_\Lambda \; \boldsymbol{G}(\boldsymbol{f})|_{\Lambda'+(a,b)}]$ is a *nontrivial* homogeneous polynomial of degree L in the variables $(c_0, \ldots, c_{L-1}, f_0, \ldots, f_{L-1})$. Further, since the mapping $\boldsymbol{d} \mapsto \boldsymbol{f} = \boldsymbol{M}^{-b}\boldsymbol{T}^{-a}\boldsymbol{d}$ is a unitary linear transform from $\mathbb{C}^L$ onto $\mathbb{C}^L$, it follows that

$$\det[\boldsymbol{G}(\boldsymbol{c})|_\Lambda \; \boldsymbol{G}(\boldsymbol{d})|_{\Lambda'}] = \omega^m \det[\boldsymbol{G}(\boldsymbol{c})|_\Lambda \; \boldsymbol{G}(\boldsymbol{f})|_{\Lambda'+(a,b)}]$$

is a *nontrivial* homogeneous polynomial of degree L in the variables $(c_0, \ldots, c_{L-1}, d_0, \ldots, d_{L-1})$. Therefore, the zeros of this polynomial constitute a zero-measure subset of $\mathbb{C}^{2L}$ that is closed and has empty interior. The proof is complete by observing that there exist only finitely many choices of $\Lambda, \Lambda' \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $|\Lambda| + |\Lambda'| = L$, and that all the mentioned properties are invariant under finite unions. □

Proof of Theorem 8. We will only prove the case $N = 2$, as the generalization to $N \geq 3$ is straightforward. We may assume without loss of generality that $\alpha_1 \neq 0$. Fix any $\Lambda, \Lambda^{(1)}, \Lambda^{(2)} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ with $\Lambda \cap (\Lambda^{(1)} \cup \Lambda^{(2)}) = \emptyset$ and $|\Lambda| + |\Lambda^{(1)}| + |\Lambda^{(2)}| = L$, and enumerate the elements of $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ by $\Lambda = \{\lambda_1, \ldots, \lambda_r\}$ where $r = |\Lambda|$. For brevity, let us write $\pi(\lambda) = \boldsymbol{M}^\ell \boldsymbol{T}^k$ for $\lambda = (k, \ell) \in \mathbb{Z}_L \times \mathbb{Z}_L$. Using that the determinant of a square matrix is linear in each of the columns, we have

$$\begin{aligned}
&\det[\, \boldsymbol{G}(\alpha_1 \boldsymbol{c}^{(1)} + \alpha_2 \boldsymbol{c}^{(2)})|_\Lambda \;\; \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \;\; \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda^{(2)}} \,] \\
&= \sum_{n_1, \ldots, n_r = 1}^{2} \alpha_{n_1} \ldots \alpha_{n_r} \cdot \det[\, \pi(\lambda_1)\boldsymbol{c}^{(n_1)} \ldots \pi(\lambda_r)\boldsymbol{c}^{(n_r)} \;\; \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \;\; \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda^{(2)}} \,] \\
&= (\alpha_1)^r \det[\, \boldsymbol{G}(\boldsymbol{c}^{(1)})|_\Lambda \;\; \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \;\; \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda^{(2)}} \,] \\
&\quad + \sum_{(n_1, \ldots, n_r) \neq (1, \ldots, 1)} \alpha_{n_1} \ldots \alpha_{n_r} \\
&\qquad\qquad \cdot \det[\, \pi(\lambda_1)\boldsymbol{c}^{(n_1)} \ldots \pi(\lambda_r)\boldsymbol{c}^{(n_r)} \;\; \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \;\; \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda^{(2)}} \,].
\end{aligned} \tag{17}$$

Then, as in the proof of Theorem 7,

$$\det[\, \boldsymbol{G}(\boldsymbol{c}^{(1)})|_\Lambda \;\; \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \;\; \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda^{(2)}} \,] = \det[\, \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda \cup \Lambda^{(1)}} \;\; \boldsymbol{G}(\boldsymbol{c}^{(2)})|_{\Lambda^{(2)}} \,]$$

is a *nontrivial* homogeneous polynomial of degree L in the variables $(c_0^{(1)}, \ldots, c_{L-1}^{(1)}, c_0^{(2)}, \ldots, c_{L-1}^{(2)})$, more precisely, a *nontrivial* linear combination of monomials each of which is formed with $r + |\Lambda^{(1)}|$ elements from $c_0^{(1)}, \ldots, c_{L-1}^{(1)}$ and $|\Lambda^{(2)}|$ elements from $c_0^{(2)}, \ldots, c_{L-1}^{(2)}$. As such monomials do not appear in the remaining terms of (17), no cancellation can occur between the first and the remaining terms of (17). Hence, we conclude that (17) is a *nontrivial* polynomial. The rest of the proof is similar to the proof of Theorem 7. □

7.3. Proof of Theorem 15 and Lemma 16

Proof of Theorem 15. Suppose to the contrary that for some subset $\Lambda \subset \mathbb{Z}_L \times \mathbb{Z}_L$ of size $L+1$ and $\boldsymbol{a} \in \mathbb{C}^\Lambda \backslash \{0\}$ there exists no vector $\boldsymbol{c} \in \mathbb{C}^L$ such that the $(L+1)\times(L+1)$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible. This in particular means that for every $\boldsymbol{c} \in \mathcal{S}$ the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is rank deficient, and since $\boldsymbol{G}(\boldsymbol{c})|_\Lambda \in \mathbb{C}^{L\times(L+1)}$ has the full rank L we deduce that $\boldsymbol{a}^*$ belongs in the row space of $\boldsymbol{G}(\boldsymbol{c})|_\Lambda$, that is,

$$\boldsymbol{a} \in \bigcap_{\boldsymbol{c}\in\mathcal{S}} \operatorname{ran}\left(\boldsymbol{G}(\boldsymbol{c})|_\Lambda\right)^*.$$

However, by the fundamental theorem of linear algebra and Lemma 16 we have

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

which is a contradiction. This shows that there exists a vector $\boldsymbol{c} \in \mathbb{C}^L$ such that the $(L+1)\times(L+1)$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible, which in turn implies that the determinant of $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is a nontrivial polynomial in the variables $c_0, \dots, c_{L-1}$. Hence, the set of all vectors $\boldsymbol{c} \in \mathbb{C}^L$ such that $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_\Lambda \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible is a dense open subset of $\mathbb{C}^L$ with full Lebesgue measure. □

Proof of Lemma 16. Let us enumerate the elements of Λ by $\Lambda = \{(k_1, \ell_1), \dots, (k_R, \ell_R)\} \subset \mathbb{Z}_L \times \mathbb{Z}_L$ and fix any $\boldsymbol{d} \in \mathcal{S}$, so that $\boldsymbol{G}(\boldsymbol{d})$ has full spark. Then $\boldsymbol{G}(\boldsymbol{d})|_\Lambda \in \mathbb{C}^{L\times R}$ has the full rank L and therefore $V = \ker \boldsymbol{G}(\boldsymbol{d})|_\Lambda \subset \mathbb{C}^R$ is an $R - L \geq 1$ dimensional subspace of $\mathbb{C}^R$. The full spark property of $\boldsymbol{G}(\boldsymbol{d})$, and therefore of $\boldsymbol{G}(\boldsymbol{d})|_\Lambda$, implies that every nontrivial vector in V must have at least $L+1$ nonzero entries, that is, $\|\boldsymbol{x}\|_0 \geq L+1$ for $\boldsymbol{x} \in V\backslash\{0\}$. Let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_{R-L}\} \subset \mathbb{C}^R$ be a basis of V. Applying the Gaussian elimination to $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{R-L}$, in general, gives a nontrivial vector in V with at least $R-L-1$ zero entries. If $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{R-L}$ would share a zero entry, then we would get a nontrivial vector in V with at least $R-L$ zero entries, that is, with at most $R-(R-L) = L$ nonzero entries, leading to a contradiction. Therefore, the vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{R-L}$ have no zero entry in common and consequently, we are able to construct a vector $\boldsymbol{x} \in V$ that has no zero entry, that is, $\|\boldsymbol{x}\|_0 = R$.

For $0 \leq p, q \leq L-1$ fixed, the commutation relation $\boldsymbol{T}^k \boldsymbol{M}^\ell = \omega^{k\ell} \boldsymbol{M}^\ell \boldsymbol{T}^k$ with $\omega = e^{2\pi i/L}$ gives

$$\boldsymbol{M}^\ell \boldsymbol{T}^k (\boldsymbol{M}^q \boldsymbol{T}^p \boldsymbol{d}) = \omega^{kq-\ell p} \boldsymbol{M}^q \boldsymbol{T}^p (\boldsymbol{M}^\ell \boldsymbol{T}^k \boldsymbol{d}), \quad k, \ell = 0, \dots, L-1,$$

and since $\boldsymbol{G}(\boldsymbol{d})$ has full spark, it follows that $\boldsymbol{G}(\boldsymbol{M}^q \boldsymbol{T}^p \boldsymbol{d})$ has full spark as well, that is, $\boldsymbol{M}^q \boldsymbol{T}^p \boldsymbol{d} \in \mathcal{S}$. Collecting the equation for $(k, \ell) \in \Lambda$ we obtain

$$\boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d})|_\Lambda = \boldsymbol{M}^q\boldsymbol{T}^p\,\boldsymbol{G}(\boldsymbol{d})|_\Lambda\,\boldsymbol{D}^{(p,q)},$$

where $\boldsymbol{D}^{(p,q)} = \operatorname{diag}(\omega^{k_1q-\ell_1p},\dots,\omega^{k_Rq-\ell_Rp})$. As $\boldsymbol{x}\in\ker\boldsymbol{G}(\boldsymbol{d})|_\Lambda$, this implies that for $p,q=0,\dots,L-1$,

$$\overline{\boldsymbol{D}^{(p,q)}}\boldsymbol{x}\in\ker\boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d})|_\Lambda \quad \text{with} \;\; \boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}\in\mathcal{S}.$$

To prove that $\operatorname{span}\{\ker\boldsymbol{G}(\boldsymbol{c})|_\Lambda : \boldsymbol{c}\in\mathcal{S}\} = \mathbb{C}^R$, it suffices to show that $\operatorname{span}\{\overline{\boldsymbol{D}^{(p,q)}}\boldsymbol{x} : p,q=0,\dots,L-1\} = \mathbb{C}^R$. Since all entries of $\boldsymbol{x}$ are nonzero, this is equivalent to

$$\operatorname{span}\{\boldsymbol{y}^{(p,q)} : p,q=0,\dots,L-1\} = \mathbb{C}^R,$$

where $\boldsymbol{y}^{(p,q)} = (\omega^{k_1q-\ell_1p},\dots,\omega^{k_Rq-\ell_Rp})^{\mathrm{T}}$ for $p,q=0,\dots,L-1$. Note that the vectors

$$\boldsymbol{v}^{(k,\ell)} = [\omega^{kq-\ell p}]_{p,q=0}^{L-1}\in\mathbb{C}^{L^2}, \quad k,\ell=0,\dots,L-1,$$

form an orthogonal basis for $\mathbb{C}^{L^2}$, in particular, the subset of R vectors $\{\boldsymbol{v}^{(k_j,\ell_j)}\}_{j=1}^R$ is linearly independent and therefore the matrix $\boldsymbol{E} = [\boldsymbol{v}^{(k_1,\ell_1)}\,|\,\boldsymbol{v}^{(k_2,\ell_2)}\,|\,\dots\,|\,\boldsymbol{v}^{(k_R,\ell_R)}]\in\mathbb{C}^{L^2\times R}$ has the full rank R. As the vectors $\boldsymbol{y}^{(p,q)}$, $p,q=0,\dots,L-1$, are precisely the rows of $\boldsymbol{E}$, we conclude that these vectors span $\mathbb{C}^R$. $\square$

7.4. *Proof of Theorems 24 and 26*

For the proof, we need the following lemma which can be found in e.g., [19].

Lemma 36 *(Extended Laplace Expansion). Let $\boldsymbol{B} = [B_{i,j}]_{i,j=1}^N$ be an $N\times N$ matrix and let $s=(s_1,s_2,\dots,s_m)$ be a partition of the row indices of $\boldsymbol{B}$, that is, $\bigcup_{j=1}^m s_j = \{0,\dots,N-1\}$ and $s_j\cap s_k=\emptyset$ for $j\neq k$. Then*

$$\det(\boldsymbol{B}) = \sum_t \operatorname{sgn}(s,t)\cdot\det\boldsymbol{B}(s_1,t_1)\cdot\ldots\cdot\det\boldsymbol{B}(s_m,t_m), \tag{18}$$

where $t=(t_1,t_2,\dots,t_m)$ runs through all partitions of the column indices $\{0,\dots,N-1\}$ with $|t_j|=|s_j|$ for $j=1,\dots,m$. Here $\operatorname{sgn}(s,t)=\pm1$ denotes the sign of the permutation $\left(\begin{smallmatrix} s_1 & s_2 & \cdots & s_m \\ t_1 & t_2 & \cdots & t_m\end{smallmatrix}\right)$, and $\boldsymbol{B}(s_j,t_j)$ denotes the submatrix of $\boldsymbol{B}$ formed with the rows indexed by s_j and the columns indexed by t_j, where elements in s_j and t_j are arranged in increasing order.

Remark 37. Given a partition $s=(s_1,s_2,\dots,s_m)$ of $\{0,\dots,N-1\}$, there exist in total $\frac{N!}{|s_1|!\,\dots\,|s_m|!}$ distinct choices of $t=(t_1,t_2,\dots,t_m)$. In fact, each term of the sum in (18) contributes exactly $|s_1|!\,\dots\,|s_m|!$ terms to the sum in the well-known determinant formula

$$\det(\boldsymbol{B}) = \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, B_{0,\sigma_0}\, B_{1,\sigma_1}\, \ldots\, B_{N-1,\sigma_{N-1}},$$

thereby accounting for all $N! = \frac{N!}{|s_1|! \ldots |s_m|!} \times (|s_1|! \ldots |s_m|!)$ terms in this sum.

Proof of Theorem 24. Assume that $\det \boldsymbol{A}_k \neq (-1)^{L-1} \det \boldsymbol{A}_{k+1}$ for some k and let $\boldsymbol{c} \in \mathbb{C}^L$ be arbitrary. It suffices to show that the determinant of the matrix (cf. (5))

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{D}_k \boldsymbol{W}_L & \boldsymbol{D}_{k+1} \boldsymbol{W}_L \\ \boldsymbol{A}_k & \boldsymbol{A}_{k+1} \end{bmatrix}$$

is a nontrivial polynomial in the variables $c_0, \ldots, c_{L-1}$. To see this, we apply Lemma 36 to $\boldsymbol{B}$ with the partition $s = (\{0, \ldots, L-1\}, \{L, \ldots, 2L-1\})$ of $\mathbb{Z}_{2L}$ and obtain that

$$\det(\boldsymbol{B}) = \sum_{t_0} \operatorname{sgn}(s,t) \cdot \det \boldsymbol{B}(\{0, \ldots, L-1\}, t_0) \\ \cdot \det \boldsymbol{B}(\{L, \ldots, 2L-1\}, \mathbb{Z}_{2L} \backslash t_0), \tag{19}$$

where t_0 runs through all subsets of $\mathbb{Z}_{2L}$ with size L. Here $\operatorname{sgn}(s,t)$ is the sign of the permutation $\begin{pmatrix} \{0,\ldots,L-1\} & \{L,\ldots,2L-1\} \\ t_0 & \mathbb{Z}_{2L} \backslash t_0 \end{pmatrix}$. The term corresponding to $t_0 = \{0, \ldots, L-1\}$ is

$$\det(\boldsymbol{D}_k \boldsymbol{W}_L) \det \boldsymbol{A}_{k+1} = c_0 \ldots c_{L-1} \det(\boldsymbol{W}_L) \det \boldsymbol{A}_{k+1}$$

and the one corresponding to $t_0 = \{L, \ldots, 2L-1\}$ is

$$(-1)^L \det(\boldsymbol{D}_{k+1} \boldsymbol{W}_L) \cdot \det \boldsymbol{A}_k = (-1)^L c_0 \ldots c_{L-1} \det(\boldsymbol{W}_L) \det \boldsymbol{A}_k.$$

It follows easily from Proposition 20 that these two terms are the only ones that produce the monomial $c_0 \ldots c_{L-1}$, in fact, all monomials appearing in other terms have indices different from $L(L-1)/2$ modulo L. Hence, the monomial $c_0 \ldots c_{L-1}$ appears in $\det(\boldsymbol{B})$ with the coefficient $\det(\boldsymbol{W}_L) \cdot (\det \boldsymbol{A}_{k+1} + (-1)^L \det \boldsymbol{A}_k) \neq 0$, which shows that $\det(\boldsymbol{B})$ is not identically zero.

Next, assume L is prime and $\det \boldsymbol{A}_k \neq \det \boldsymbol{A}_\ell$ for some $0 \leq k < \ell \leq L-1$. Following the same arguments with

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{D}_k \boldsymbol{W}_L & \boldsymbol{D}_\ell \boldsymbol{W}_L \\ \boldsymbol{A}_k & \boldsymbol{A}_\ell \end{bmatrix},$$

we find two terms in the sum (19): $c_0 \ldots c_{L-1} \det(\boldsymbol{W}_L) \det \boldsymbol{A}_\ell$ corresponding to $t_0 = \{0, \ldots, L-1\}$, and $(-1)^L c_0 \ldots c_{L-1} \det(\boldsymbol{W}_L) \det \boldsymbol{A}_k$ corresponding to $t_0 = \{L, \ldots, 2L-1\}$. Note that since L is prime, we have $p(\ell - k) = 0 \mod L$ if and only if $p = 0 \mod L$. Using this fact, we similarly conclude that the monomial $c_0 \ldots c_{L-1}$ appears in $\det(\boldsymbol{B})$ with the coefficient $\det(\boldsymbol{W}_L) \cdot (\det \boldsymbol{A}_\ell + (-1)^L \det \boldsymbol{A}_k) \neq 0$, hence, $\det(\boldsymbol{B})$ is not identically zero. $\square$

Proof of Theorem 26. Applying Lemma 36 to the matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{A} \end{bmatrix}$ with the partition $s = (\{0, 1, \ldots, L-1\}, \{L, L+1, \ldots, R-1\})$ of $\mathbb{Z}_R$, we have

$$\det\left(\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{A} \end{bmatrix}\right) = \sum_t \operatorname{sgn}(s,t) \cdot \det \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, t_0) \cdot \det \boldsymbol{A}(\mathbb{Z}_{R-L}, \mathbb{Z}_R \backslash t_0), \tag{20}$$

where t_0 runs through all subsets of $\mathbb{Z}_R$ with size L. Here $\operatorname{sgn}(s,t)$ is the sign of the permutation $\begin{pmatrix} \{0,\ldots,L-1\} & \{L,\ldots,R-1\} \\ t_0 & \mathbb{Z}_R\backslash t_0 \end{pmatrix}$. Notice that to each t_0 corresponds a unique subset $\Lambda' \subset \widetilde{\Lambda}$ of size L satisfying $\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, t_0) = \boldsymbol{G}(\boldsymbol{c})|_{\Lambda'}$. Let us denote by $\overline{t_0}$ the one with $\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, \overline{t_0}) = \boldsymbol{G}(\boldsymbol{c})|_{\Lambda}$, and rewrite the righthand side of (20) as

$$\begin{aligned} &\operatorname{sgn}(s,\bar{t}) \cdot \det \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, \overline{t_0}) \cdot \det \boldsymbol{A}(\mathbb{Z}_{R-L}, \mathbb{Z}_R \backslash \overline{t_0}) \\ &+ \sum_{t_0 \neq \overline{t_0}} \operatorname{sgn}(s,t) \cdot \det \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, t_0) \cdot \det \boldsymbol{A}(\mathbb{Z}_{R-L}, \mathbb{Z}_R \backslash t_0). \end{aligned} \tag{21}$$

The condition (i) implies that Λ is the only subset of $\widetilde{\Lambda}$ with size L that is associated with the L-tuple $\boldsymbol{\tau}(\Lambda)$, i.e., $\boldsymbol{\tau}(\Lambda') \neq \boldsymbol{\tau}(\Lambda)$ for any other subset $\Lambda' \subset \widetilde{\Lambda}$ of size L. Together with condition (ii), we have $\operatorname{ind}(\boldsymbol{\tau}(\Lambda')) \neq \operatorname{ind}(\boldsymbol{\tau}(\Lambda))$ for every such $\Lambda' \neq \Lambda$. By Proposition 20, all monomials appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_{\Lambda'})$ have indices equal to $\operatorname{ind}(\boldsymbol{\tau}(\Lambda'))$ and are therefore distinct from the ones appearing in $\det(\boldsymbol{G}(\boldsymbol{c})|_{\Lambda})$, which all have indices equal to $\operatorname{ind}(\boldsymbol{\tau}(\Lambda))$. Equivalently, in terms of t_0, this means that monomials appearing in $\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, t_0)$ with $t_0 \neq \overline{t_0}$ are distinct from the ones appearing in $\boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}}(\mathbb{Z}_L, \overline{t_0})$. Finally, the full spark of $\boldsymbol{A}$ implies $\det \boldsymbol{A}(\mathbb{Z}_{R-L}, \mathbb{Z}_R \backslash t_0) \neq 0$ and therefore the first term of (21) is a *nontrivial* homogeneous polynomial of degree L in the variables $c_0, \ldots, c_{L-1}$ (see Theorem 3). Hence $\det \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c})|_{\widetilde{\Lambda}} \\ \boldsymbol{A} \end{bmatrix}$ is a *nontrivial* polynomial and the desired set can be obtained by excluding the zero set of this polynomial from $\mathbb{C}^L$. □

7.5. Proof of Lemma 28 and Theorems 29 and 30

Proof of Lemma 28. The case $N = 1$ is proved in Lemma 16 (it is even proved for $|\Lambda^{(1)}| \leq L^2$). We will only prove the case $N = 2$, as the generalization to $N \geq 3$ is straightforward.

For brevity of notation, we write $\Lambda^{(1)}, \Lambda^{(2)}$ as Λ, Λ'. By Lemma 35, there exists $(a,b) \in \mathbb{Z}_L \times \mathbb{Z}_L$ such that $\Lambda \cap (\Lambda' + (a,b)) = \emptyset$. Let us enumerate the elements of Λ by $\Lambda = \{(k_1, \ell_1), \ldots, (k_r, \ell_r)\}$ and $\Omega := \Lambda' + (a,b) = \{(k_{r+1}, \ell_{r+1}), \ldots, (k_R, \ell_R)\}$, where $r = |\Lambda|$ and $R = |\Lambda| + |\Omega|$ $(< 2L)$, and choose $(\boldsymbol{d}, \boldsymbol{d}') \in \mathcal{S}_2$ so that $[\boldsymbol{G}(\boldsymbol{d})\ \boldsymbol{G}(\boldsymbol{d}')]$ has full spark.

For $0 \leq p, q \leq L-1$ fixed, the commutation relation $\boldsymbol{T}^k \boldsymbol{M}^\ell = \omega^{k\ell} \boldsymbol{M}^\ell \boldsymbol{T}^k$ gives

$$\boldsymbol{M}^\ell \boldsymbol{T}^k (\boldsymbol{M}^q \boldsymbol{T}^p \boldsymbol{d}) = \omega^{kq} \boldsymbol{M}^{\ell+q} \boldsymbol{T}^{k+p} \boldsymbol{d}, \quad k, \ell = 0, \ldots, L-1,$$

which shows that the columns of $\boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d})$ and $\boldsymbol{G}(\boldsymbol{d})$ are exactly the same up to ordering and phase factors. Therefore, $(\boldsymbol{d}, \boldsymbol{d}') \in \mathcal{S}_2$ implies $(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}, \boldsymbol{M}^{q'}\boldsymbol{T}^{p'}\boldsymbol{d}') \in \mathcal{S}_2$ for every $p, q, p', q' = 0, \dots, L-1$. Using the commutation relation once more, we get

$$\boldsymbol{M}^\ell\boldsymbol{T}^k(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}) = \omega^{kq-\ell p}\boldsymbol{M}^q\boldsymbol{T}^p(\boldsymbol{M}^\ell\boldsymbol{T}^k\boldsymbol{d}), \quad k, \ell = 0, \dots, L-1,$$

which leads to

$$\boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d})|_\Lambda = \boldsymbol{M}^q\boldsymbol{T}^p\, \boldsymbol{G}(\boldsymbol{d})|_\Lambda\, \boldsymbol{D}^{(p,q)},$$

where $\boldsymbol{D}^{(p,q)} = \mathrm{diag}(\omega^{k_1 q - \ell_1 p}, \dots, \omega^{k_r q - \ell_r p}) \in \mathbb{C}^{r\times r}$. Similarly, we have that

$$\boldsymbol{G}(\boldsymbol{M}^b\boldsymbol{T}^a\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}')|_{\Lambda'} = \boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}')|_\Omega\, \boldsymbol{F} = \boldsymbol{M}^q\boldsymbol{T}^p\, \boldsymbol{G}(\boldsymbol{d}')|_\Omega\, \boldsymbol{D'}^{(p,q)}\, \boldsymbol{F},$$

where $\boldsymbol{F} = \mathrm{diag}(\omega^{k_{r+1}b}, \dots, \omega^{k_R b}) \in \mathbb{C}^{(R-r)\times(R-r)}$ and $\boldsymbol{D'}^{(p,q)} = \mathrm{diag}(\omega^{k_{r+1}q-\ell_{r+1}p}, \dots, \omega^{k_R q - \ell_R p}) \in \mathbb{C}^{(R-r)\times(R-r)}$. Combining these together, we obtain

$$\left[\boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d})|_\Lambda\ \boldsymbol{G}(\boldsymbol{M}^b\boldsymbol{T}^a\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}')|_{\Lambda'}\right] = \boldsymbol{M}^q\boldsymbol{T}^p\left[\boldsymbol{G}(\boldsymbol{d})|_\Lambda\ \boldsymbol{G}(\boldsymbol{d}')|_\Omega\right]\boldsymbol{E}^{(p,q)}\begin{bmatrix}\boldsymbol{I}_r & 0\\ 0 & \boldsymbol{F}\end{bmatrix},$$

where $\boldsymbol{E}^{(p,q)} = \mathrm{diag}(\omega^{k_1 q - \ell_1 p}, \dots, \omega^{k_R q - \ell_R p}) \in \mathbb{C}^{R\times R}$.

As in the proof of Lemma 16, we can find a vector $\boldsymbol{x}$ in $\ker\left[\boldsymbol{G}(\boldsymbol{d})|_\Lambda\ \boldsymbol{G}(\boldsymbol{d}')|_\Omega\right]$ $(\subset \mathbb{C}^{\Lambda\cup\Omega})$ with no zero entries. Setting

$$\boldsymbol{z} := \begin{bmatrix}\boldsymbol{I}_r & 0\\ 0 & \overline{\boldsymbol{F}}\end{bmatrix}\boldsymbol{x},$$

we note that $\overline{\boldsymbol{E}^{(p,q)}}\boldsymbol{z} \in \ker\left[\boldsymbol{G}(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d})|_\Lambda\ \boldsymbol{G}(\boldsymbol{M}^b\boldsymbol{T}^a\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}')|_{\Lambda'}\right]$ and $(\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d},$ $\boldsymbol{M}^b\boldsymbol{T}^a\boldsymbol{M}^q\boldsymbol{T}^p\boldsymbol{d}') \in \mathcal{S}_2$ for $p, q = 0, \dots, L-1$. It now suffices to show that $\mathrm{span}\{\overline{\boldsymbol{E}^{(p,q)}}\boldsymbol{z} : p, q = 0, \dots, L-1\} = \mathbb{C}^{\Lambda\cup\Omega}$, where $\boldsymbol{z} \in \mathbb{C}^R$ has no zero entries; but this was already seen in the proof of Lemma 16. □

Proof of Theorem 29. We will use similar arguments as in the proof of Theorem 15.

Suppose to the contrary that for some $\Lambda^{(1)}, \dots, \Lambda^{(N)} \subset \mathbb{Z}_L\times\mathbb{Z}_L$ with $\sum_{n=1}^N |\Lambda^{(n)}| = L+1$ and $\boldsymbol{a} \in \mathbb{C}^{L+1}\backslash\{0\}$, there exists no vector $(\boldsymbol{c}^{(1)}, \dots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that the $(L+1)\times(L+1)$ matrix $\begin{bmatrix}\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda^{(N)}}\\ \boldsymbol{a}^*\end{bmatrix}$ is invertible. This in particular means that for every $(\boldsymbol{c}^{(1)}, \dots, \boldsymbol{c}^{(N)}) \in \mathcal{S}_N$ the matrix $\begin{bmatrix}\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda^{(N)}}\\ \boldsymbol{a}^*\end{bmatrix}$ is rank deficient, and since $\left[\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}}\ \cdots\ \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda^{(N)}}\right]$ has the full rank L we deduce that $\boldsymbol{a}^*$ belongs in the row space of $\left[\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}}\ \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda^{(N)}}\right]$, that is,

$$\boldsymbol{a} \in \bigcap_{\boldsymbol{c}\in\mathcal{S}_N} \mathrm{ran}\begin{bmatrix}\boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda^{(1)}} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda^{(N)}}\\ \boldsymbol{a}^*\end{bmatrix}^*.$$

However, by the fundamental theorem of linear algebra and Lemma 28 we have

$$\bigcap_{\boldsymbol{c}\in\mathcal{S}_N} \operatorname{ran} \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda(1)} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda(N)} \\ \boldsymbol{a}^* \end{bmatrix}^* = \bigcap_{\boldsymbol{c}\in\mathcal{S}_N} \left(\ker \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda(1)} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda(N)} \\ \boldsymbol{a}^* \end{bmatrix} \right)^{\perp} = \{0\},$$

which is a contradiction. This shows that there exists a vector $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that the $(L+1)\times(L+1)$ matrix $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda(1)} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda(N)} \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible, which in turn implies that the determinant of $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda(1)} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda(N)} \\ \boldsymbol{a}^* \end{bmatrix}$ is a nontrivial polynomial in the variables $c_0^{(1)}, \ldots, c_{L-1}^{(1)}, \ldots, c_0^{(N)}, \ldots, c_{L-1}^{(N)}$. Hence, the set of all vectors $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that $\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda(1)} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda(N)} \\ \boldsymbol{a}^* \end{bmatrix}$ is invertible is a dense open subset of $(\mathbb{C}^L)^N$ with full Lebesgue measure. □

Proof of Theorem 30. If $|\Lambda_m| = \sum_{n=1}^N |\Lambda_{m,n}| = L+1$ for some m, then the matrix (13) reduces to

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_m} \\ \boldsymbol{a}_m^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)})|_{\Lambda_{m,1}} \cdots \boldsymbol{G}(\boldsymbol{c}^{(N)})|_{\Lambda_{m,N}} \\ \boldsymbol{a}_m^* \end{bmatrix}$$

and the claim follows directly from Theorem 29. Therefore, we may assume that $|\Lambda_m| \le L$ for $m = 1, \ldots, M$. Observe that each of the column spaces of

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_1} \\ 0 \\ \vdots \\ 0 \\ \boldsymbol{a}_1^* \end{bmatrix}, \begin{bmatrix} 0 \\ \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_2} \\ \vdots \\ 0 \\ \boldsymbol{a}_2^* \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \boldsymbol{G}(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)})|_{\Lambda_M} \\ \boldsymbol{a}_M^* \end{bmatrix}$$

has trivial intersection with the others because of the positioning of zero matrices. By Theorem 29, we can find a vector $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that each of the matrices above has full column rank. This in turn implies that the determinant of the $(L+1)\times(L+1)$ matrix (13) is a nontrivial polynomial in the variables $c_0^{(1)}, \ldots, c_{L-1}^{(1)}, \ldots, c_0^{(N)}, \ldots, c_{L-1}^{(N)}$. Hence, the set of all vectors $(\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(N)}) \in (\mathbb{C}^L)^N$ such that the matrix (13) is invertible is a dense open subset of $(\mathbb{C}^L)^N$ with full Lebesgue measure. □

7.6. Proof of the first part of Theorem 31, based on Corollary 4

Assume that $S \subset \mathbb{R}^2$ is a Jordan domain (a bounded set whose boundary is a Lebesgue zero set) with measure $|S| < 1$. Then we can choose $T > 0$ small and $L \in \mathbb{N}$ large so that S is contained in a fundamental domain of the lattice $(LT)\mathbb{Z}\times(1/T)\mathbb{Z}$ and S intersects with at most L rectangles of the form $[kT, (k+1)T)\times[\frac{\ell}{LT}, \frac{\ell+1}{LT})$ with $k, \ell \in \mathbb{Z}$, where the set of all such pairs (k, ℓ) are distinct in $\mathbb{Z}_L\times\mathbb{Z}_L$ (cf. Remark 20 in [18]). Choosing any vector $\boldsymbol{c} = \{c_n\}_{n=0}^{L-1} \in \mathbb{C}^L$ from the set $\mathcal{S}$ in Theorem 3, we design a pilot signal $g = \sum_{n\in\mathbb{Z}} c_n\delta_{nT}$ where $\{c_n\}_{n\in\mathbb{Z}}$ is the L-periodic extension of $\boldsymbol{c}$, i.e., $c_{n+L} = c_n$ for $n \in \mathbb{Z}$.

Since S is contained in a fundamental domain of $(LT)\mathbb{Z}\times(1/T)\mathbb{Z}$, it follows that $Hg \in L^2(\mathbb{R})$ for every $H \in OPW(S)$. Moreover, it holds that for $H \in OPW(S)$,

$$\mathcal{Z}_L^{LT} Hg(t,\nu) = \boldsymbol{G}(\boldsymbol{c})\,\vec{\eta}_H(t,\nu), \tag{22}$$

where $\mathcal{Z}_L^{LT}$ is the L-dimensional vector valued Zak transform

$$\mathcal{Z}_L^{LT} f(t,\nu) = \begin{pmatrix} Z^{LT} f(t,\nu) \\ Z^{LT} f(t+T,\nu)\, e^{-2\pi i T\nu} \\ \vdots \\ Z^{LT} f(t+(L-1)T,\nu)\, e^{-2\pi i (L-1)T\nu} \end{pmatrix}$$

with Z^{LT} defined densely on $L^2(\mathbb{R})$ by

$$Z^{LT} f(t,\nu) = \sum_{n\in\mathbb{Z}} f(t+nLT)\, e^{-2\pi i nLT\nu},$$

and $\vec{\eta}_H(t,\nu) = \{\vec{\eta}_H(t,\nu)_{(k,\ell)}\}_{k,\ell=0}^{L-1}$ is the L^2-dimensional vector valued function defined by

$$\vec{\eta}_H(t,\nu)_{(k,\ell)} = \tfrac{1}{LT} \sum_{p\in\mathbb{Z}} \sum_{q\in\mathbb{Z}} \eta_H\big(t+(pL+k)T, \nu+(qL+\ell)\tfrac{1}{LT}\big)\, e^{-2\pi i (pL+k)T\nu}.$$

For a proof of (22), we refer to Lemma 2.7 and Remark 2.8 in [17] (also see [18, Lemma 44]).

Note that equality (22) holds in locally L^2 sense and therefore pointwise almost everywhere. Further, due to the quasi-periodicity of $\mathcal{Z}_L^{LT} Hg$ and $\vec{\eta}_H$, it is enough to observe (22) for a.e. (t,ν) in $[0,T)\times[0,\frac{1}{LT})$. Therefore, the continuous-time problem is reduced to a family of finite-dimensional problems (22) indexed by a.e. $(t,\nu) \in [0,T)\times[0,\frac{1}{LT})$.

Let us denote by Λ the set of all pairs $(k,\ell) \in \mathbb{Z}\times\mathbb{Z}$ such that S intersects with $[kT,(k+1)T)\times[\frac{\ell}{LT},\frac{\ell+1}{LT})$; then we have $|\Lambda| \le L$. Since all elements of Λ are distinct in $\mathbb{Z}_L\times\mathbb{Z}_L$, the set $\Lambda' = \{(k',\ell') \in \mathbb{Z}_L\times\mathbb{Z}_L : (pL+k', qL+\ell') \in \Lambda \text{ for some } p,q \in \mathbb{Z}\}$ is of cardinality $|\Lambda'| = |\Lambda| \le L$. This implies that $\vec{\eta}_H(t,\nu)_{(k,\ell)} = 0$ for $(t,\nu) \in [0,T)\times[0,\frac{1}{LT})$ and $(k,\ell) \notin \Lambda'$, which reduces equation (4) to

$$\mathcal{Z}_L^{LT} Hg(t,\nu) = \boldsymbol{G}(\boldsymbol{c})|_{\Lambda'}\, \vec{\eta}_H(t,\nu)|_{\Lambda'} \quad a.e.\ \ (t,\nu) \in [0,T)\times[0,\tfrac{1}{LT}).$$

Then for a.e. $(t,\nu) \in [0,T)\times[0,\frac{1}{LT})$ fixed, we can apply the finite dimensional result (Corollary 4) to solve the equation for $\vec{\eta}_H(t,\nu)|_{\Lambda'}$.

Now that we have the L^2-dimensional vector valued function $\vec{\eta}_H(t,\nu)$ for a.e. $(t,\nu) \in [0,T)\times[0,\frac{1}{LT})$, it remains to show that the spreading function η_H can be fully determined from $\vec{\eta}_H$. To see this, we note that $\operatorname{supp}\eta_H \subset S$ is contained in a fundamental domain of $(LT)\mathbb{Z}\times(1/T)\mathbb{Z}$, and thus for each $(t,\nu) \in [0,T)\times[0,\frac{1}{LT})$ and $(k,\ell) \in \mathbb{Z}_L\times\mathbb{Z}_L$ at most one of the values $\eta_H(t+(pL+k)T, \nu+(qL+\ell)\frac{1}{LT})$, $p,q \in \mathbb{Z}$, is nonzero. Therefore, η_H can be recovered directly from $\vec{\eta}_H(t,\nu)$ for a.e. $(t,\nu) \in [0,T)\times[0,\frac{1}{LT})$. Hence, we conclude that $OPW(S)$ is identifiable.

Declaration of Competing Interest

There is no competing interest.

Acknowledgements

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. Part of this work was presented at the 43rd IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Calgary, Canada, April 2018 [23].

References

[1] G.E. Pfander, D. Walnut, Measurement of time-variant channels, IEEE Trans. Inform. Theory 52 (11) (2006) 4808–4820.
[2] R. Heckel, H. Bölcskei, Identification of sparse linear operators, IEEE Trans. Inform. Theory 59 (12) (2013) 7985–8000.
[3] D. Tse, P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, Cambridge, 2005.
[4] G. Matz, F. Hlawatsch, Chapter 1 – Fundamentals of time-varying communication channels, in: F. Hlawatsch, G. Matz (Eds.), Wireless Communications over Rapidly Time-Varying Channels, Academic Press, Oxford, 2011, pp. 1–63.
[5] P. Bello, Characterization of randomly time-variant linear channels, IEEE Trans. Commun. Syst. 11 (4) (1963) 360–393.
[6] T. Kailath, Time-variant communication channels, IEEE Trans. Inform. Theory 9 (4) (1963) 233–237.
[7] G.E. Pfander, Measurement of time-varying multiple-input multiple-output channels, Appl. Comput. Harmon. Anal. 24 (3) (2008) 393–401.
[8] R. Heckel, H. Bölcskei, Identification of sparse linear operators, IEEE Trans. Inform. Theory 59 (12) (2013) 7985–8000.
[9] G.E. Pfander, P. Zheltov, Identification of stochastic operators, Appl. Comput. Harmon. Anal. 26 (2) (2014) 256–279.
[10] E. Telatar, Capacity of multi-antenna gaussian channels, Eur. Trans. Telecommun. 10 (6) (1999) 585–595.
[11] D.-S. Shiu, G.J. Foschini, M.J. Gans, J.M. Kahn, Fading correlation and its effect on the capacity of multielement antenna systems, IEEE Trans. Commun. 48 (3) (2000) 502–513.
[12] V. Jungnickel, V. Pohl, C. von Helmolt, Capacity of MIMO systems with closely-spaced antennas, IEEE Commun. Lett. 7 (8) (2003) 361–363.
[13] H.N.M. Mbonjo, J. Hansen, V. Hansen, MIMO capacity and antenna array design, in: IEEE Global Telecommunications Conference (GLOBECOM), vol. 5, 2004, pp. 3155–3159.
[14] E.G. Larsson, O. Edfors, F. Tufvesson, T.L. Marzetta, Massive MIMO for next generation wireless systems, IEEE Commun. Mag. 52 (2) (2014) 186–195.
[15] J.W. Wallace, M.A. Jensen, Mutual coupling in MIMO wireless systems: a rigorous network theory analysis, IEEE Trans. Wirel. Commun. 3 (4) (2004) 1317–1325.

[16] D. Gesbert, H. Bölcskei, D. Gore, A. Paulraj, MIMO wireless channels: capacity and performance prediction, in: IEEE Global Telecommun. Conf., vol. 2, 2000, pp. 1083–1088.
[17] D.G. Lee, G.E. Pfander, V. Pohl, Sampling and reconstruction of multiple-input multiple-output channels, IEEE Trans. Signal Process. 67 (4) (2019) 961–976.
[18] G.E. Pfander, D.F. Walnut, Sampling and reconstruction of operators, IEEE Trans. Inform. Theory 62 (1) (2016) 435–458.
[19] J. Lawrence, G.E. Pfander, D. Walnut, Linear independence of gabor systems in finite dimensional vector spaces, J. Fourier Anal. Appl. 11 (6) (2005) 715–726.
[20] R.-D. Malikiosis, A note on Gabor frames in finite dimensions, Appl. Comput. Harmon. Anal. 38 (2) (2015) 318–330.
[21] G.E. Pfander, Sampling of operators, J. Fourier Anal. Appl. 19 (3) (2013) 612–650.
[22] W. Kozek, G.E. Pfander, Identification of operators with bandlimited symbols, SIAM J. Math. Anal. 37 (3) (2005) 867–888.
[23] D.G. Lee, G.E. Pfander, V. Pohl, W. Zhou, Identification of multiple-input multiple-output channels under linear side constraints, in: Proc. 43rd Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Calgary, Canada, 2018.