Compressive Identification of Linear Operators

Reinhard Heckel and Helmut Bölcskei

Department of Information Technology and Electrical Engineering, ETH Zurich

E-mail: {heckel,boelcskei}@nari.ee.ethz.ch

Abstract—We consider the problem of identifying a linear deterministic operator from an input-output measurement. For the large class of continuous (and hence bounded) operators, under additional mild restrictions, we show that stable identifiability is possible if the total support area of the operator's spreading function satisfies $\Delta \leq 1/2$. This result holds for arbitrary (possibly fragmented) support regions of the spreading function, does not impose limitations on the total extent of the support region, and, most importantly, does not require the support region of the spreading function to be known prior to identification. Furthermore, we prove that asking for identifiability of only almost all operators, stable identifiability is possible if $\Delta \leq 1$. This result is surprising as it says that there is no penalty for not knowing the support region of the spreading function prior to identification.

I. INTRODUCTION

Identification of deterministic linear operators from an input-output measurement is an important problem in many fields of engineering. Concrete examples include system identification in control theory and practice and the measurement of wireless communication channels.

It is natural to ask under which conditions on the operator, identification from an input-output measurement is possible in principle, and how one would go about choosing the probing signal and extracting the operator from the corresponding output signal. This paper addresses these questions in generality by considering the (large) class of linear operators that can be represented as a continuous weighted superposition of time-frequency shift operators, i.e., the operator's response to the signal $x(t)$ is given by

$$y(t) = \int_\tau \int_\nu s_H(\tau, \nu) x(t - \tau) e^{j2\pi\nu t} d\nu d\tau \qquad (1)$$

where s_H denotes the spreading function of the operator H. The representation theorem [1, Thm. 14.3.5] states that the action of any continuous (and hence bounded) linear operator, under additional mild restrictions, can be represented as in (1). In the communications literature [2], [3], [4], operators with input-output relation as in (1) are referred to as linear time-varying channels/systems and s_H is the delay-Doppler spreading function. For the special case of linear time-invariant (LTI) systems, we have $s_H(\tau, \nu) = h(\tau)\delta(\nu)$, so that (1) reduces to

$$y(t) = \int_\tau h(\tau) x(t - \tau) d\tau. \qquad (2)$$

The question of identifiability of LTI systems is readily answered by noting that the system's response to $x(t) = \delta(t)$ is given by its impulse response $h(t)$, which fully characterizes the input-output relation according to (2). LTI systems are therefore always identifiable, provided that the input signal can have infinite bandwidth and we can observe the output signal over an infinite duration. In the general case (i.e., for LTV systems), the situation is fundamentally different. Specifically, Kailath's landmark paper [2] shows that LTV systems with spreading function compactly supported on a rectangle are identifiable if and only if the spreading function's support area satisfies $\Delta \leq 1$. Bello [3] later pointed out that Kailath's identifiability result continues to hold, even if the support region of s_H is scattered across the (τ, ν)-plane, as long as the total support area satisfies $\Delta \leq 1$. Kozek and Pfander [5] and Pfander and Walnut [6] provided functional-analytic proofs of the results in [2], [3]. Common to [2], [3], [5], [6] is the assumption of the support region of s_H being known prior to identification. This is clearly restrictive and often impossible to realize in practice.

Contributions: In this paper, we consider the problem of identifying deterministic linear operators with spreading function compactly supported on an *unknown*, possibly fragmented, region in the (τ, ν)-plane. We do not impose limitations on the total extent of the support region. Our main result shows that an operator with input-output relation (1) is identifiable, without prior knowledge of the spreading function's support region, if and only if the total support area of s_H satisfies $\Delta \leq 1/2$. We then show that this factor of two penalty—compared to the case where the support region is known in advance [2], [3]—can be eliminated, if one asks for identifiability of *almost all* operators only. This result is surprising as it shows that (for almost all operators) there is no price to be paid for not knowing the spreading function's support region in advance. We discuss the design of probing signals and we outline an algorithm which, in the noiseless case, provably recovers all operators with $\Delta \leq 1/2$. Furthermore, we present an algorithm which, again in the noiseless case, provably recovers almost all operators with $\Delta \leq 1$.

Relation to previous work: Recently Bajwa et al. [7] considered the identification of LTV-systems with spreading function s_H supported in a rectangle of area $\Delta \leq 1$ and consisting of a finite number of discrete components with unknown delays and Doppler shifts. In the present paper, we allow general spreading functions that can be supported in the entire (τ, ν)-plane with possibly fragmented support region.

Notation: We use lowercase boldface letters to denote column vectors, e.g., $\mathbf{x}$, and uppercase boldface letters to designate matrices, e.g., $\mathbf{X}$. The superscripts *, ^{H}, and ^{T} stand for complex conjugation, Hermitian transposition, and transposition, respectively. The space spanned by the columns

of $\mathbf{X}$ is designated by $\mathcal{R}(\mathbf{X})$. The entry in the kth row and lth column of $\mathbf{X}$ is denoted by $[\mathbf{X}]_{k,l}$. For the vector $\mathbf{x}$, the Euclidean norm is denoted as $\|\mathbf{x}\|_{\ell_2}$ and the kth entry of $\mathbf{x}$ is $[\mathbf{x}]_k$. $|\Omega|$ stands for the cardinality of the set Ω. For two sets Ω_1 and Ω_2 we define set addition as $\Omega_1 + \Omega_2 = \{\omega : \omega_1 + \omega_2, \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$. $\delta(x)$ denotes the Dirac delta function. For a function $f(\mathbf{x})$, $\mathrm{supp}(f)$ denotes the support set of f. For two functions $f(\mathbf{x}), g(\mathbf{x})$, defined on Ω, we write $\langle f, g \rangle \triangleq \int_\Omega f(\mathbf{x})g^*(\mathbf{x})d\mathbf{x}$ for the inner product; $\|f\| \triangleq \sqrt{\langle f, f \rangle}$ is the norm of f.

II. PROBLEM FORMULATION

Given normed linear spaces[1] X, Y, we consider linear operators $H : X \mapsto Y$ that can be represented as a continuous weighted superposition of translation operators T_τ, with $(T_\tau x)(t) \triangleq x(t - \tau)$, and modulation operators M_ν with $(M_\nu x)(t) \triangleq e^{j2\pi\nu t}x(t)$:

$$(Hx)(t) \triangleq \int_\tau \int_\nu s_H(\tau, \nu)(M_\nu T_\tau x)(t)d\nu d\tau. \quad (3)$$

This is a rather general setting, since according to [1, Thm. 14.3.5], any continuous (and hence bounded) linear operator, under additional mild restrictions, can be represented as in (3). In the following, denote the linear space of operators that can be represented according to (3) by $\mathcal{H}$, and define the inner product on this space by

$$\langle H_1, H_2 \rangle_\mathcal{H} \triangleq \langle s_{H_1}, s_{H_2} \rangle = \int_{\mathbb{R}^2} s_{H_1}(\tau, \nu)s_{H_2}^*(\tau, \nu)d(\tau, \nu),$$

$H_1, H_2 \in \mathcal{H}$, with the induced norm $\|H\|_\mathcal{H} \triangleq \sqrt{\langle H, H \rangle_\mathcal{H}}$.

Restrictions on the spreading function: We assume that the support region of s_H has the form

$$M_\Gamma \triangleq \bigcup_{(k,m)\in\Gamma} \left(U + \left(kT, \frac{m}{TL}\right) \right) \subseteq [0, \tau_{\max}) \times [0, \nu_{\max}) \quad (4)$$

where $U \triangleq [0, T) \times [0, 1/TL)$ is a "fundamental cell" in the (τ, ν)-plane and $T \in \mathbb{R}$ is a parameter whose role will become clear shortly. The set of "active cells" is specified by $\Gamma \subseteq \Sigma \triangleq \{(0,0), (0,1), ..., (L-1, L-1)\}$. Consequently, we have $\tau_{\max} = TL$ and $\nu_{\max} = 1/T$. We denote the support area of M_Γ as $\mathcal{A}(M_\Gamma)$. Given a general support region for s_H, possibly fragmented and spread over the entire (τ, ν)-plane, we can choose T and L such that this region can be approximated arbitrarily well by a support region of the form (4) (see Fig. 1).

A. Identifiability

Let us next formally define the notion of identifiability of a set of operators $\mathcal{Q} \subseteq \mathcal{H}$. The set $\mathcal{Q}$ is said to be identifiable from an input-output measurement, if there exists a probing signal $x \in X$ such that for each $H \in \mathcal{Q}$, the action of the operator on the probing signal, Hx, uniquely determines H, i.e., if there exists an $x \in X$ such that

[1]Since our identifiability proof relies on sending Dirac delta impulses, we need to choose X such that it contains generalized functions. To keep the exposition simple, we will not dwell on the resulting functional-analytic subtleties. Instead, we refer the interested reader to [5], [6] for a description of the rigorous mathematical setup required here.
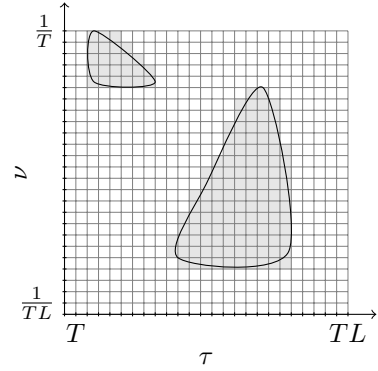

Fig. 1. Approximation of a general support region.

$$H_1 x = H_2 x \implies H_1 = H_2, \quad \forall H_1, H_2 \in \mathcal{Q}. \quad (5)$$

Identifiability is hence equivalent to invertibility of the mapping

$$\mathcal{Q} \to Y : H \mapsto Hx \quad (6)$$

induced by the probing signal x. Invertibility alone is typically not sufficient as one wants to recover H from Hx in a numerically stable fashion, i.e., we want small errors in Hx to result in small errors in the identified operator. This requirement implies that the inverse of the mapping (6) must be continuous (and hence bounded), which finally motivates the following definition.

Definition 1: We say that x *stably* identifies $\mathcal{Q}$ if there exist constants $0 < \alpha \leq \beta < \infty$ such that for all pairs $H_1, H_2 \in \mathcal{Q}$,

$$\alpha\|H_1 - H_2\|_\mathcal{H} \leq \|H_1 x - H_2 x\| \leq \beta\|H_1 - H_2\|_\mathcal{H}. \quad (7)$$

Furthermore, we say that $\mathcal{Q}$ is stably identifiable, if there exists an $x \in X$ such that x stably identifies $\mathcal{Q}$.

Note that (7) is stated in terms of differences of operators, since $(H_1 - H_2)$ is not necessarily contained in the set $\mathcal{Q}$. The lower bound in (7) guarantees that the inverse of (6) exists and is bounded and hence continuous, as desired. Proving that x stably identifies $\mathcal{Q}$ essentially amounts to proving that $\alpha > 0$ in Definition 1. The ratio β/α provides a measure for the noise sensitivity of the identification process, the closer β/α to one, the better.

III. MAIN RESULTS

Before stating our main results, we define the set of operators with s_H supported on a given area M_Γ:

$$\mathcal{H}_{M_\Gamma} \triangleq \{H \in \mathcal{H} : \mathrm{supp}(s_H) \subseteq M_\Gamma\}. \quad (8)$$

In this notation, Kailath [2] discussed the case where M_Γ is a (single) rectangle, and Bello [3] analyzed the case where M_Γ is an arbitrary region, possibly fragmented and spread over the entire (τ, ν)-plane. We first recall the key results in [3], [6] on the identification of $\mathcal{H}_{M_\Gamma}$ under the assumption of M_Γ known:

Theorem 1 ([3], [6]): Let M_Γ be given. The set of operators $\mathcal{H}_{M_\Gamma}$ is stably identifiable if and only if $\mathcal{A}(M_\Gamma) \leq 1$.

In the following, we consider the set of operators

$$\mathcal{X}(\Delta) \triangleq \bigcup_{\mathcal{A}(M_\Gamma)\leq\Delta} \mathcal{H}_{M_\Gamma} \quad (9)$$

which consists of *all* sets $\mathcal{H}_{M_\Gamma}$ such that $\mathcal{A}(M_\Gamma) \leq \Delta$.

Our main results are as follows.

Theorem 2: The set of operators $\mathcal{X}(\Delta)$ is stably identifiable if and only if $\Delta \leq 1/2$.

Note that Theorem 2 applies to the union of *all* sets of operators with spreading function supported on a region with area no larger than Δ, and, in particular, does not need the support region of s_H to be known in advance. The main implication of Theorem 2 is that the penalty for not knowing the support region of s_H in advance is a factor of two in the support set size of the spreading function. We can eliminate this penalty by relaxing the identification requirement to *almost all $H \in \mathcal{X}(\Delta)$*.

Theorem 3: Almost all $H \in \mathcal{X}(\Delta)$ can be stably identified if $\Delta < 1$.

The factor of two penalty in Theorem 2 has the same roots as the factor of two penalty in spectrum-blind sampling [8], [9], [10], [11], sparse signal recovery [12], and in the recovery of signals that lie in a union of subspaces [13]. It was recognized before—in the context of spectrum-blind sampling—that the factor of two penalty can be eliminated by relaxing the recovery requirement to *almost all* signals [8], [9], [10].

IV. NECESSITY IN THEOREM 2

To prove necessity in Theorem 2 we start with the following lemma which states an equivalent condition on stable identifiability of $\mathcal{X}(\Delta)$. This condition is often easier to verify than that in Definition 1.

Lemma 1: x stably identifies $\mathcal{X}(\Delta)$ if and only if it stably identifies *all* sets

$$\mathcal{H}_{M_\Phi \cup M_\Theta} \triangleq \{H : H = H_1 - H_2, H_1 \in \mathcal{H}_{M_\Phi}, H_2 \in \mathcal{H}_{M_\Theta}\}$$

with $\mathcal{A}(M_\Phi) \leq \Delta$ and $\mathcal{A}(M_\Theta) \leq \Delta$ where $\Phi, \Theta \in \Sigma$.

Proof of Lemma 1: Follows immediately by invoking the definition of stable identifiability and noting that $\alpha\|H_1 - H_2\|_\mathcal{H} \leq \|(H_1 - H_2)x\| \leq \beta\|H_1 - H_2\|_\mathcal{H}, \forall H_1, H_2 \in \mathcal{X}(\Delta)$ if and only if $\alpha\|H\|_\mathcal{H} \leq \|Hx\| \leq \beta\|H\|_\mathcal{H}$ for all $H \in \mathcal{H}_{M_\Phi \cup M_\Theta}$ and for all M_Φ, M_Θ, with $\mathcal{A}(M_\Phi) \leq \Delta$ and $\mathcal{A}(M_\Theta) \leq \Delta$, $\Phi, \Theta \in \Sigma$. ∎

Before formally proving necessity in Theorem 2, we comment on an important aspect of the difference between known and unknown spreading function support region. If the support region is known and given, by say M_Γ, and we consider the set $\mathcal{H}_{M_\Gamma}$, it follows that the difference $H_1 - H_2$ for $H_1, H_2 \in \mathcal{H}_{M_\Gamma}$ satisfies $(H_1 - H_2) \in \mathcal{H}_{M_\Gamma}$. If we, however, consider the set $\mathcal{X}(\Delta)$, we have that $H_1, H_2 \in \mathcal{X}(\Delta)$ does not imply $(H_1 - H_2) \in \mathcal{X}(\Delta)$ in general. To see this, simply take $H_1, H_2 \in \mathcal{X}(\Delta)$ such that the support regions of H_1 and H_2 have area Δ and are disjoint. We do, however, have that $H_1 - H_2 \in \mathcal{X}(2\Delta)$, $\forall H_1, H_2 \in \mathcal{X}(\Delta)$. This observation lies at the heart of the factor of two penalty in Δ as quantified by Theorem 2.

Necessity in Theorem 2: Follows by taking $\Delta > 1/2$ and noting that we can find M_Φ, M_Θ with $\mathcal{A}(M_\Phi) = \mathcal{A}(M_\Theta) = \Delta$

and $M_\Phi \cap M_\Theta = \emptyset$. This implies $\mathcal{A}(M_\Phi \cup M_\Theta) > 1$ and hence application of Theorem 1 to the set $\mathcal{H}_{M_\Phi \cup M_\Theta}$ establishes that $\mathcal{H}_{M_\Phi \cup M_\Theta}$ is not stably identifiable. By Lemma 1 this then implies that $\mathcal{X}(\Delta)$ is not stably identifiable. ∎

V. SUFFICIENCY IN THEOREM 2

We prove sufficiency in Theorem 2 by finding a probing signal that stably identifies $\mathcal{X}(\Delta)$. Concretely, we choose a weighted TL-periodic train of Dirac impulses

$$x(t) = \sum_{k \in \mathbb{Z}} c_k \delta(t + kT), \quad c_k = c_{k+L}, \ \forall k \in \mathbb{Z} \qquad (10)$$

as probing signal. The specific choice of the coefficients $\mathbf{c} = \{c_0, ..., c_{L-1}\}$ will turn out to be crucial and will be discussed later. The main idea of our proof is to 1) reduce the identification problem to that of solving a linear system of L equations with L^2 unknowns, and 2) to apply Lemma 1 to show that a unique solution of this underdetermined system of equations exists whenever $\Delta \leq 1/2$ (and $\mathbf{c}$ is chosen appropriately).

We start by computing the response of H to $x(t)$ in (10):

$$y(t) = (Hx)(t) = \sum_{k \in \mathbb{Z}} c_k \int_\nu s_H(t + kT, \nu)e^{j2\pi\nu t}d\nu. \quad (11)$$

Next, define the Zak transform [14] (with parameter TL) of the signal $u(t)$ as

$$\mathcal{Z}_u(t, f) \triangleq \sum_{m \in \mathbb{Z}} u(t - mTL)e^{j2\pi mTLf}$$

for $(t, f) \in [0, TL) \times [0, 1/TL)$. The Zak transform of $y(t)$ in (11) is given by

$$\mathcal{Z}_y(t, f) =$$
$$= \sum_{k,m \in \mathbb{Z}} c_k \int_\nu s_H(t - mTL + kT, \nu)e^{j2\pi\nu(t-mTL)}d\nu\, e^{j2\pi mTLf}$$
$$= \sum_{k',m \in \mathbb{Z}} c_{k'} \int_\nu s_H(t + k'T, \nu)e^{j2\pi\nu t}e^{-j2\pi\nu mTL}d\nu\, e^{j2\pi mTLf}$$
$$= \sum_{k \in \mathbb{Z}} c_k \frac{1}{TL} \sum_{m \in \mathbb{Z}} s_H\left(t + kT, f + \frac{m}{TL}\right)e^{j2\pi t\left(f + \frac{m}{TL}\right)}$$

where we used the substitution $k' = k - mL$ and the last step follows from the Poisson summation formula. Next, we substitute $t = t' + pT$ with $p \in \{0, ..., L - 1\}$ and $t' \in [0, T)$. This amounts to splitting the fundamental rectangle $[0, TL) \times [0, 1/TL)$ of the Zak transform into L "cells" U, where U was defined in Section II, and yields

$$z_p(t', f) \triangleq \mathcal{Z}_y(t' + pT, f), \quad (t', f) \in U, \quad p = 0, ..., L - 1$$
$$= \sum_{k,m \in \mathbb{Z}} \frac{c_k}{TL} s_H\left(t' + pT + kT, f + \frac{m}{TL}\right)e^{j2\pi(t'+pT)\left(f+\frac{m}{TL}\right)}$$
$$= \sum_{k' \in \mathbb{Z}} \frac{c_{k'-p}}{TL} \sum_{m \in \mathbb{Z}} s_H\left(t' + k'T, f + \frac{m}{TL}\right)e^{j2\pi(t'+pT)\left(f+\frac{m}{TL}\right)}$$
$$= \sum_{k=0}^{L-1} \frac{c_{k-p}}{TL} \sum_{m=0}^{L-1} s_H\left(t' + kT, f + \frac{m}{TL}\right)e^{j2\pi(t'+pT)\left(f+\frac{m}{TL}\right)}$$

where we used the substitution $k' = p + k$ and the last step is a consequence of $s_H(\tau, \nu) = 0$ for $(\tau, \nu) \notin [0, TL) \times [0, 1/T)$, by definition. We can now rewrite the last equation in vector-matrix form by defining the column vectors $\mathbf{z}(t, f)$ and $\mathbf{s}(t, f)$:

$$[\mathbf{z}(t, f)]_p \triangleq z_p(t, f) e^{-j2\pi pTf}, \quad p = 0, ..., L-1$$

and $\quad \mathbf{s}(t, f) \quad \triangleq \quad [s_{0,0}(t, f), s_{0,1}(t, f), ..., s_{0,L-1}(t, f),$ $s_{1,0}(t, f), ..., s_{L-1,L-1}(t, f)]^T$ with

$$s_{k,m}(t, f) \triangleq s_H\left(t + kT, f + \frac{m}{TL}\right) e^{j2\pi \left(f + \frac{m}{TL}\right)t}. \quad (12)$$

It is easily seen that the vector $\mathbf{s}(t, f)$, $(t, f) \in U$, fully characterizes the spreading function. With all definitions in place, we finally obtain

$$\mathbf{z}(t, f) = \mathbf{A_c}\mathbf{s}(t, f), \quad (t, f) \in U \quad (13)$$

with the $L \times L^2$ matrix

$$\mathbf{A_c} \triangleq [\mathbf{A_{c,0}}| ... |\mathbf{A_{c,L-1}}], \quad \mathbf{A_{c,k}} \triangleq \frac{1}{TL}\mathbf{C_{c,k}}\mathbf{F}^H$$

where $[\mathbf{F}]_{p,m} = e^{-j2\pi \frac{pm}{L}}$, $p, m = 0, ..., L-1$, and $\mathbf{C_{c,k}}$ is the diagonal matrix with diagonal entries $\{c_k, c_{k-1}, ..., c_{k+1}\}$.

The proof will be effected by applying Lemma 1. Concretely we will prove stable identifiability of $\mathcal{H}_{M_\Phi \cup M_\Theta}$ for all pairs M_Φ, M_Θ with $\mathcal{A}(M_\Phi) \leq 1/2$ and $\mathcal{A}(M_\Theta) \leq 1/2$. By setting $M_\Gamma = M_\Phi \cup M_\Theta$, this is equivalent to proving stable identifiability of $\mathcal{H}_{M_\Gamma}$ for all M_Γ with $\mathcal{A}(M_\Gamma) \leq 1$. We therefore consider $H \in \mathcal{H}_{M_\Gamma}$, and note that, by definition $s_{k,m}(t, f) = 0$, $\forall (k, m) \notin \Gamma$. Denote the vector obtained from $\mathbf{s}(t, f)$ by selecting the entries corresponding to the active cells Γ by $\mathbf{s}_\Gamma(t, f)$ and let $\mathbf{A}_\Gamma$ be the matrix containing the columns of $\mathbf{A_c}$ that correspond to these cells. Then (13) becomes[2]

$$\mathbf{z}(t, f) = \mathbf{A}_\Gamma \mathbf{s}_\Gamma(t, f), \quad (t, f) \in U. \quad (14)$$

Next, we formally relate (14) to the definition of stable identifiability through the following lemma, whose proof is given in the appendix.

Lemma 2: Let x be given by (10). Then, the bounds α, β in (7) for the set of operators $\mathcal{H}_{M_\Gamma}$ are given as

$$\alpha_\Gamma = \sqrt{TL} \inf_{\|\mathbf{v}\|_{\ell_2}=1} (\mathbf{A}_\Gamma \mathbf{v}), \ \beta_\Gamma = \sqrt{TL} \sup_{\|\mathbf{v}\|_{\ell_2}=1} (\mathbf{A}_\Gamma \mathbf{v}). \ (15)$$

The proof of sufficiency in Theorem 2 is now completed by showing that for all M_Γ with $\mathcal{A}(M_\Gamma) \leq 1$, $\mathcal{H}_{M_\Gamma}$ is stably identifiable, i.e., $\alpha_\Gamma > 0$. This amounts to proving that $\mathbf{A}_\Gamma$ has full rank for all M_Γ such that $\mathcal{A}(M_\Gamma) \leq 1$, i.e., for all $\Gamma \in \Sigma$ such that $|\Gamma| \leq L$. What comes to our rescue here is a result in [15] which states that for almost all $\mathbf{c}$, each $L \times L$ submatrix of $\mathbf{A_c}$ has full rank. Hence, there exists a $\mathbf{c}$ such that $\alpha_\Gamma > 0$ for all M_Γ with $\mathcal{A}(M_\Gamma) \leq 1$. In the remainder of the paper, $\mathbf{c}$ is chosen such that each $L \times L$ submatrix of $\mathbf{A_c}$ has full rank.

---

[2]Pfander and Walnut [6] used the probing signal (10) to prove that, for known spreading function support region, $\Delta \leq 1$ is sufficient for stable identifiability. The crucial difference between [6] and our setup is that we need *each* submatrix of $\mathbf{A_c}$ of L columns to have full rank, as we do not assume knowledge of the support region.

VI. RECOVERING THE SPREADING FUNCTION

We next discuss an algorithm for the recovery of operators $H \in \mathcal{X}(\Delta)$ from the operator's response Hx to the probing signal $x(t)$ in (10). We start by noting that recovering H amounts to recovering s_H which by (12) is equivalent to recovering $\mathbf{s}(t, f)$ from (13). This will be accomplished by first identifying the support set of s_H, (i.e., the active cells of $\mathbf{s}(t, f)$), and then solving the resulting linear system of equations (14).

1) Support set recovery: If we assume that $\mathcal{A}(M_\Gamma) \leq 1/2$, then the support set Γ can be recovered from $\mathbf{z}(t, f)$ by solving:

$$(\text{P0}) \begin{cases} \text{minimize} & |\Gamma| \\ \text{subject to} & \mathbf{z}(t, f) = \mathbf{A}_\Gamma \mathbf{s}_\Gamma(t, f), \quad (t, f) \in U \end{cases}$$

where the constraint is over all $\mathbf{s}_\Gamma(t, f)$ with $|\Gamma| \leq L/2$. This is a standard problem, and solutions have been proposed in the context of spectrum-blind sampling [8], [9], [11], all involving a correlation matrix, which in our setup becomes $\mathbf{Z} \triangleq \int_U \mathbf{z}(t, f)\mathbf{z}^H(t, f)d(t, f)$. The main difference to signal recovery in the context of spectrum-blind sampling [8], [9], [11] is that here a function of two variables has to be recovered rather than a function of one variable. Using (14) we can express $\mathbf{Z}$ according to

$$\mathbf{Z} = \mathbf{A}_\Gamma \mathbf{S}_\Gamma \mathbf{A}_\Gamma^H \quad (16)$$

where $\mathbf{S}_\Gamma = \int_U \mathbf{s}_\Gamma(t, f)\mathbf{s}_\Gamma^H(t, f)d(t, f)$. Using similar arguments as in [10], it can be shown that (P0) is equivalent to

$$(\overline{\text{P0}}) \begin{cases} \text{minimize} & |\Gamma| \\ \text{subject to} & \mathbf{Z} = \mathbf{A}_\Gamma \mathbf{S}_\Gamma \mathbf{A}_\Gamma^H \end{cases}$$

where the constraint is over all Hermitian matrices $\mathbf{S}_\Gamma \in \mathbb{C}^{|\Gamma| \times |\Gamma|}$. Since $\mathbf{Z}$ is normal, it can be decomposed as $\mathbf{Z} = \mathbf{Q}\mathbf{Q}^H$, where the $R = \text{rank}(\mathbf{Z})$ columns of the matrix $\mathbf{Q} \in \mathbb{C}^{L \times R}$ are orthogonal. As shown by Feng [10], $(\overline{\text{P0}})$ (and hence (P0)) is equivalent to

$$(\text{P0}') \begin{cases} \text{minimize} & |\Gamma| \\ \text{subject to} & \mathbf{Q} = \mathbf{A}_\Gamma \mathbf{G}_\Gamma \end{cases} \quad (17)$$

where the constraint is over all $\mathbf{G}_\Gamma \in \mathbb{C}^{|\Gamma| \times R}$. (P0$'$) is known in the literature as the finite multiple-measurement vector (MMV) problem[3] [16]. Application of [16, Thm. 2.4] ensures that (P0$'$) provably recovers the correct support set Γ as long as $|\Gamma| \leq L/2$, i.e., as long as the area of the unknown support set of the spreading function satisfies $\Delta \leq 1/2$.

2) Recovery for known support set: Once the support set Γ has been identified, we solve (14) for $\mathbf{s}_\Gamma(t, f)$, which based on (12) yields s_H and hence H. Note that (14) has to be solved over the continuum of values $(t, f) \in U$. We can expand all quantities in (14) into two-dimensional Fourier series over U, which results in a system of countably many linear equations to be solved.

---

[3]The MMV problem is usually formulated as follows: Minimize $\|\mathbf{G}\|_{\text{row-0}}$ subject to $\mathbf{Q} = \mathbf{A_c}\mathbf{G}$, where the constraint is over all $\mathbf{G} \in \mathbb{C}^{L^2 \times R}$ and $\|\mathbf{G}\|_{\text{row-0}}$ is the number of rows of $\mathbf{G}$ that contain at least one non-zero entry.

VII. Identification for Almost All $H \in \mathcal{X}(\Delta)$

We next prove Theorem 3. The proof is inspired by [9, Thm. 1], and is constructive as it specifies the recovery algorithm (for almost all $H \in \mathcal{X}(\Delta), \Delta < 1$). The basic idea is to use a MUSIC-like [17] algorithm based on (16), which allows us to recover Γ under the following two conditions:

1) $\Delta \leq 1 - 1/L$. The penalty of $1/L$ is technical and can be made arbitrarily small by choosing L large enough.
2) The functions $s_{k,m}(t,f), (k,m) \in \Gamma$, are *linearly independent* on U, i.e., there is no vector $\mathbf{a} \in \mathbb{C}^N, \mathbf{a} \neq \mathbf{0}, N = |\Gamma|$, such that $\mathbf{a}^H \mathbf{s}_\Gamma(t,f) = 0, \forall (t,f) \in U$.

We recognize that almost all $H \in \mathcal{X}(\Delta)$ satisfy Condition 2.

Proof of sufficiency in Theorem 3: The proof is effected by establishing that, under Conditions 1 and 2 above, the support set Γ is uniquely specified by the indices of the columns of $\mathbf{U}_n^H \mathbf{A}_\mathbf{c}$ that are equal to $\mathbf{0}$. Here, $\mathbf{U}_n$ is the matrix of eigenvectors of $\mathbf{Z}$ corresponding to zero eigenvalues. To see this, we perform an eigenvalue decomposition of $\mathbf{Z}$ in (16) to obtain

$$\mathbf{Z} = \begin{bmatrix} \mathbf{U}_z & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_z & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_z^H \\ \mathbf{U}_n^H \end{bmatrix} = \mathbf{A}_\Gamma \mathbf{S}_\Gamma \mathbf{A}_\Gamma^H \quad (18)$$

where $\mathbf{U}_z$ contains the eigenvectors of $\mathbf{Z}$, corresponding to the non-zero eigenvalues of $\mathbf{Z}$. Lemma 3 in the appendix establishes that, thanks to Condition 2 above, for almost all $H \in \mathcal{X}(\Delta)$, $\mathbf{S}_\Gamma$ has full rank. As discussed in Section V, each set of L or fewer columns of $\mathbf{A}_\mathbf{c}$ is linearly independent. Condition 1 ensures that $|\Gamma| \leq L - 1$, which implies that $\mathbf{A}_\Gamma$ has full rank for all sets Γ in question. Hence we get from (18) that

$$\mathcal{R}(\mathbf{A}_\Gamma) = \mathcal{R}(\mathbf{A}_\Gamma \mathbf{S}_\Gamma \mathbf{A}_\Gamma^H) = \mathcal{R}(\mathbf{U}_z \mathbf{\Lambda}_z \mathbf{U}_z^H) = \mathcal{R}(\mathbf{U}_z). \quad (19)$$

$\mathcal{R}(\mathbf{U}_n)$ is the orthogonal complement of $\mathcal{R}(\mathbf{U}_z)$ in $\mathbb{C}^L$. It therefore follows from (19) that $\mathbf{U}_n^H \mathbf{A}_\Gamma = \mathbf{0}$. Therefore, the columns of $\mathbf{U}_n^H \mathbf{A}_\mathbf{c}$ that correspond to indices $(k,m) \in \Gamma$ are equal to $\mathbf{0}$. To conclude the proof, we show, by contradiction, that no other columns of $\mathbf{U}_n^H \mathbf{A}_\mathbf{c}$ are equal to $\mathbf{0}$. Suppose that $\mathbf{U}_n^H \mathbf{a} = \mathbf{0}$ where $\mathbf{a}$ is any column of $\mathbf{A}_\mathbf{c}$ corresponding to an index pair $(k',m') \notin \Gamma$. Then $\mathbf{a} \in \mathcal{R}(\mathbf{U}_z) = \mathcal{R}(\mathbf{A}_\Gamma)$. This would mean that the L or fewer columns of $\mathbf{A}_\mathbf{c}$ corresponding to the indices $(k,m) \in \{\Gamma \cup (k',m')\}$ would be linearly dependent. The proof is completed by noting that this stands in contradiction to the fact that—since we assume that $\mathbf{c}$ is chosen accordingly—each set of L or fewer columns of $\mathbf{A}_\mathbf{c}$ is linearly independent. ∎

APPENDIX

Proof of Lemma 2: Starting from (14), we get for fixed values of $(t,f) \in U$

$$\frac{\alpha_\Gamma}{\sqrt{TL}} \|\mathbf{s}_\Gamma(t,f)\|_{\ell_2} \leq \|\mathbf{z}(t,f)\|_{\ell_2} \leq \frac{\beta_\Gamma}{\sqrt{TL}} \|\mathbf{s}_\Gamma(t,f)\|_{\ell_2} \quad (20)$$

with $\alpha_\Gamma, \beta_\Gamma$ defined in (15). Squaring and integrating (20) over U yields

$$\int_U \|\mathbf{z}(t,f)\|_{\ell_2}^2 d(t,f) = \sum_{p=0}^{L-1} \int_U |z_p(t,f)|^2 d(t,f)$$

$$= \int_{[0,TL] \times [0,1/TL]} |\mathcal{Z}_y(t,f)|^2 d(t,f) = \frac{1}{TL} \|Hx\|^2 \quad (21)$$

where the last equality follows from the unitarity of the Zak transform [14]. Similarly, we have

$$\int_U \|\mathbf{s}_\Gamma(t,f)\|_{\ell_2}^2 d(t,f) = \|s_H\|^2 = \|H\|_{\mathcal{H}}^2. \quad (22)$$

Inserting (22) and (21) into (20) completes the proof. ∎

Lemma 3: $\mathbf{S}_\Gamma$ has full rank if and only if the functions $s_{k,m}(t,f), (k,m) \in \Gamma$, are linearly independent on U.

Proof: Assume that $\mathbf{S}_\Gamma$ does not have full rank. Then there exists an $\mathbf{a} \neq \mathbf{0}, \mathbf{a} \in \mathbb{C}^N$ such that $\mathbf{a}^H \mathbf{S}_\Gamma = \mathbf{0}$ and hence $\mathbf{a}^H \mathbf{S}_\Gamma \mathbf{a} = \int_U \mathbf{a}^H \mathbf{s}_\Gamma(t,f)(\mathbf{a}^H \mathbf{s}_\Gamma(t,f))^H d(t,f) = 0$. Since $\mathbf{a}^H \mathbf{s}_\Gamma(t,f)(\mathbf{a}^H \mathbf{s}_\Gamma(t,f))^H \geq 0, \forall (t,f) \in U$, we must have $\mathbf{a}^H \mathbf{s}_\Gamma(t,f) = 0$ a.e. on U, which implies that the set $s_{k,m}(t,f), (k,m) \in \Gamma$, is linearly dependent on U. Now assume that the set $s_{k,m}(t,f), (k,m) \in \Gamma$, is linearly dependent on U. Then, there exists an $\mathbf{a} \neq \mathbf{0}$ such that $\int_U \mathbf{a}^H \mathbf{s}_\Gamma(t,f) \mathbf{s}_\Gamma^H(t,f) d(t,f) = \mathbf{0}$, and hence, using $\mathbf{S}_\Gamma = \int_U \mathbf{s}_\Gamma(t,f) \mathbf{s}_\Gamma^H(t,f) d(t,f)$, by definition, we get $\mathbf{a}^H \mathbf{S}_\Gamma = \mathbf{0}$ which proves that $\mathbf{S}_\Gamma$ does not have full rank. ∎

References

[1] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Birkhäuser, 2001.

[2] T. Kailath, "Measurements on time-variant communication channels," *IRE Trans. Inf. Theory*, vol. 8, no. 5, pp. 229–236, Sep. 1962.

[3] P. Bello, "Measurement of random time-variant linear channels," *IEEE Trans. Inf. Theory*, vol. 15, no. 4, pp. 469–475, Jul. 1969.

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

[5] W. Kozek and G. E. Pfander, "Identification of operators with bandlimited symbols," *SIAM J. Math. Anal.*, vol. 37, no. 3, pp. 867–888, 2005.

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

[7] W. U. Bajwa, K. Gedalyahu, and Y. C. Eldar, "Identification of underspread linear systems with application to super-resolution radar," *arXiv:1008.0851v2*, Dec. 2010.

[8] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. of IEEE Int. Conf. Acoust. Speech Sig. Proc. (ICASSP)*, vol. 3, Atlanta, GA, USA, May 1996.

[9] Y. Bresler, "Spectrum-blind sampling and compressive sensing for continuous-index signals," in *Proc. of Information Theory and Applications Workshop (ITA)*, San Diego, CA, Jan. 2008, pp. 547–554.

[10] P. Feng, "Universal minimum-rate sampling and spectrum-blind reconstruction for multiband signals," Ph.D. dissertation, Univ. Illinois, Urbana-Champaign, IL, 1997.

[11] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[12] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Natl. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.

[13] Y. Lu and M. Do, "A theory for sampling signals from a union of subspaces," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2334–2345, Jun. 2008.

[14] A. J. E. M. Janssen, "The Zak transform: A signal transform for sampled time-continuous signals." *Philips J. Res.*, vol. 43, no. 1, pp. 23–69, 1988.

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

[16] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, 2006.

[17] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.