Necessary and Sufficient Conditions for Simultaneous State and Input Recovery of Linear Systems with Sparse Inputs by ℓ_1-Minimization

Kyle Poe, Enrique Mallada, René Vidal

Abstract— **The study of theoretical conditions for recovering sparse signals from compressive measurements has received a lot of attention in the research community. In parallel, there has been a great amount of work characterizing conditions for the recovery both the state and the input to a linear dynamical system (LDS), including a handful of results on recovering sparse inputs. However, existing sufficient conditions for recovering sparse inputs to an LDS are conservative and hard to interpret, while necessary and sufficient conditions have not yet appeared in the literature. In this work, we provide (1) the first characterization of necessary and sufficient conditions for the existence and uniqueness of sparse inputs to an LDS, (2) the first necessary and sufficient conditions for a linear program to recover both an unknown initial state and a sparse input, and (3) simple, interpretable recovery conditions in terms of the LDS parameters. We conclude with a numerical validation of these claims and discuss implications and future directions.**

I. INTRODUCTION

A foundational concept in systems theory is that of *observability*, the condition guaranteeing uniqueness of the initial state of a system given knowledge of the inputs and a sufficient number of measurements for the output [1]. Introduced later was the more stringent notion of *strong observability*, which further guarantees the uniqueness of the initial condition even in the presence of unknown inputs, and is known to be equivalent to the system having no invariant zeros [2]. These conditions have been used to concisely characterize conditions under which either the initial state or inputs to a system, or both, can be recovered, even in the absence of the other [3], [4], [5], [6]. Of particular relevance to time-critical applications is the development of deadbeat or finite-time input reconstructors, which in the discrete setting have been formulated in terms of solutions to a block Toeplitz system [4].

A linear system is, in particular, a compact means of describing a linear relationship $\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{u}$ between a sequence of of inputs $\boldsymbol{u}$ and a sequence of observations $\boldsymbol{y}$, from which even in the most optimistic circumstances generic $\boldsymbol{u}$ can only be reconstructed up to $\ker \boldsymbol{\Psi}$. However, by assuming that $\boldsymbol{u}$ is sparse in an appropriate sense, established results in sparse recovery provide favorable guarantees on exact reconstruction for an appropriate choice of optimization algorithm. The most common such case considered is when $\boldsymbol{u}$ is assumed to have support of size not greater than s, and is termed *regular sparsity*. Other support-based notions of sparsity include block [7], group [8], and tree-based sparsity [9], and are each subsumed by the more general notion of model-based sparsity [10].

For each of these sparsity patterns, the literature has provided tailored optimization problems and recovery guarantees, with varying levels of robustness to noise, ease of checking, and conceptual nuance. For many applications in the noiseless setting, simple ℓ_1-minimization has proven to be the approach of choice due to its relative conceptual ease, implementability as a linear program, and favorable performance even when compared with tailored regularizers [11]. For regular sparsity, the necessary and sufficient condition for successful unique recovery is the satisfaction of the so-called *nullspace property* (NUP), which requires vectors in the nullspace of $\boldsymbol{\Psi}$ to have smaller ℓ_1 norm on s-sparse supports than on the complement of s-sparse supports (see Def. 1). Recent results have even shown that for any support-based notion of sparsity, there exists a straightforward extension of the NUP, termed the *generalized nullspace property*, which provides necessary and sufficient guarantees [12].

In light of the success of this approach to signal reconstruction, recent literature has provided tailored algorithms for sparse recovery in linear dynamical systems (LDSs), where the assumption of sparsity has been variously made on the initial conditions [13], [14], [15], [16], dynamics [17], measurement noise [18], and inputs [19], [20], [21], [22]. Even with all of this prior work, there are few existing guarantees on the performance of these algorithms, and the guarantees that have been produced are typically probabilistic in nature or make restrictive assumptions on the sparsity patterns, such as the state and inputs being simultaneously sparse with respect to an orthogonal dictionary. As a result, many results for the general, noiseless setting, including the establishment of necessary and sufficient guarantees, have not yet appeared in the literature. Our focus in this work is on establishing such guarantees for the basis-pursuit style optimization problem introduced in [20], where the initial state is not sparse, but the inputs are assumed to follow an appropriate generalized support pattern. Existing conditions for even the basic version of this problem are very conservative, and necessary and sufficient conditions have not yet made an appearance.

In this work, we consider the problem of jointly inferring the initial state $\boldsymbol{x}_0$ and sparse inputs $U_N = (\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{N-1})$

K. Poe is with the Department of Biomedical Engineering, Johns Hopkins University, MD 21218 kpoe2@jh.edu

E. Mallada is with the Department of Electrical and Computer Engineering, Johns Hopkins University, MD 21218 mallada@jhu.edu

R. Vidal is with the Department of Electrical and Systems Engineering, University of Pennsylvania, PA 19104 vidalr@seas.upenn.edu

of an LDS $\Sigma = (\boldsymbol{A}, \boldsymbol{\Psi}, \boldsymbol{C})$ without a feedthrough term, i.e.,

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{\Psi}\boldsymbol{u}_k, \quad && \boldsymbol{x}_k \in \mathbb{R}^n, \boldsymbol{u}_k \in \mathbb{R}^m \\ \boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k, && \boldsymbol{y}_k \in \mathbb{R}^p, \end{aligned} \tag{1}$$

from $N+1$ output measurements $Y_N = (\boldsymbol{y}_0, \ldots, \boldsymbol{y}_N)$. In particular, this work makes the following contributions:

1) Necessary and sufficient conditions for uniqueness of $\boldsymbol{x}_0$ and sparse U_N given Y_N.
2) Necessary and sufficient conditions for the ℓ_1-minimization approach of [20] to uniquely recover $\boldsymbol{x}_0$ and sparse U_N given Y_N.
3) Interpretable conditions which are respectively necessary or sufficient for unique solutions to the ℓ_1-minimization approach of [20] in terms of the system parameters, and elaboration on situations where these conditions achieve equality.
4) Illustration of the accuracy of these conditions and provision of intuition for when they are most informative through simulations of random LDSs.

II. PRELIMINARIES

A. Notation

1) Sets and Vector Spaces: Define $\mathbb{N} = \{0, 1, 2, \ldots\}$, $[\![n]\!] := \{0, 1, 2, \ldots, n-1\}$. We use capital script letters to denote vector subspaces $\mathcal{V} \subseteq \mathbb{R}^n$. When $\mathcal{V}, \mathcal{U} \subseteq \mathbb{R}^n$, we denote $\mathcal{V} + \mathcal{U} := \{\boldsymbol{v} + \boldsymbol{u} : \boldsymbol{v} \in \mathcal{V}, \boldsymbol{u} \in \mathcal{U}\} \subseteq \mathbb{R}^n$ and $\mathcal{V}^\perp := \{\boldsymbol{x} \in \mathbb{R}^n : \forall \boldsymbol{v} \in \mathcal{V}, \langle \boldsymbol{x}, \boldsymbol{v} \rangle = 0\}$.

2) Linear Operators: For any matrix $\boldsymbol{A}$ and subspace $\mathcal{U}$, define $\boldsymbol{A}\mathcal{U} := \{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in \mathcal{U}\}$. $\boldsymbol{A}^{-1}$ is defined to be the inverse matrix of $\boldsymbol{A}$ if it exists, and for any affine subspace $\mathcal{V} \subseteq \mathbb{R}^n$, $\boldsymbol{A}^{-1}\mathcal{V} = \{\boldsymbol{x} \in \mathbb{R}^m : \boldsymbol{A}\boldsymbol{x} \in \mathcal{V}\}$. We likewise denote the Moore-Penrose pseudoinverse as $\boldsymbol{A}^+$.

3) Supports and Norms: For $\boldsymbol{x} \in \mathbb{R}^m$, denote $\operatorname{supp}(\boldsymbol{x}) := \{i \in [\![m]\!] : x_i \neq 0\}$. Denote column i of a matrix $\boldsymbol{A}$ to be $\boldsymbol{A}_i$, and block column i if $\boldsymbol{A}$ is a block matrix. For any subset $S \subseteq [\![m]\!]$, $\boldsymbol{A}_S$ is the submatrix of $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ with columns $(\boldsymbol{A}_S)_i = \boldsymbol{A}_{S_i}$. If $S = (S_k)$ is a tuple of sets and $\boldsymbol{\Gamma}$ is a block matrix, denote $\boldsymbol{\Gamma}_S$ the block matrix with block columns $(\boldsymbol{\Gamma}_S)_k = (\boldsymbol{\Gamma}_k)_{S_k}$. For a block vector U, $(U_S)_k = (U_k)_{S_k}$. Likewise if S, S' are two tuples of sets with the same length, define $S \cup S'$ the tuple of sets s.t. $(S \cup S')_k = S_k \cup S'_k$. We denote the ℓ_0 semi-norm as $\|\boldsymbol{x}\|_0 = |\operatorname{supp}(\boldsymbol{x})|$ and the ℓ_1 and ℓ_2 norms as $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively.

B. The Nullspace Property

Given a matrix $\boldsymbol{\Theta} \in \mathbb{R}^{n \times m}$, where we assume that $\boldsymbol{\Theta}$ has linearly dependent columns ($\operatorname{rank} \boldsymbol{\Theta} < m$), a central problem of sparse recovery is to inquire, under which assumptions on the support of $\boldsymbol{x}$ are there unique solutions to $\boldsymbol{y} = \boldsymbol{\Theta}\boldsymbol{x}$, and what are the algorithms with such unique recovery guarantees? A standard approach is to begin with the optimization problem P_0 that finds the sparsest solution, known to be NP-hard, and proceed to the convex relaxation P_1, known as basis pursuit:

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0, \text{ such that } \boldsymbol{y} = \boldsymbol{\Theta}\boldsymbol{x} \tag{P_0}$$

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1, \text{ such that } \boldsymbol{y} = \boldsymbol{\Theta}\boldsymbol{x} \tag{P_1}$$

Denote $\Delta_s(m) := \{S \subseteq [\![m]\!] : |S| \leq s\}$ to be the set of s-sparse supports for vectors in $\mathbb{R}^m$, or simply Δ_s when m is fixed. A classic result in sparse recovery is that any s-sparse solution to P_1 is the unique solution, if and only if $\boldsymbol{\Theta}$ satisfies the s-NUP:

Definition 1 (Nullspace Property (NUP)): The matrix $\boldsymbol{\Theta} \in \mathbb{R}^{n \times m}$ satisfies the *nullspace property of order s* (s-NUP) if $\forall \boldsymbol{h} \in \ker \boldsymbol{\Theta} \setminus \{0\}$, $\forall S \in \Delta_s$, $\|\boldsymbol{h}_S\|_1 < \|\boldsymbol{h}_{S^c}\|_1$.

C. The Generalized Nullspace Property

The motivating observation of [12] is that sparsity structures tend to satisfy the property that if S is a valid sparse support, so too is $S' \subseteq S$. This relationship describes an abstract simplicial complex:

Definition 2 (Abstract Simplicial Complex (ASC)): Let Δ be a set of sets. Δ is an *abstract simplicial complex* if $\forall S \in \Delta$, $\forall S' \subseteq S$, $S' \in \Delta$. If for some $m \in \mathbb{N}$, $\Delta \subseteq \{S : S \subseteq [\![m]\!]\}$, we say that Δ is an ASC over $[\![m]\!]$.

One can quickly check that $\Delta_s(m)$ is an ASC over $[\![m]\!]$, so ASCs comprise a strict generalization of regular sparsity. We will thus refer to any vector $\boldsymbol{x}$ such that $\operatorname{supp} \boldsymbol{x} \in \Delta$ as Δ-sparse. We additionally make the convenient definition:

$$\mathcal{S}(\Delta) := \{\boldsymbol{x} \in \mathbb{R}^m : \operatorname{supp} \boldsymbol{x} \in \Delta\} \tag{2}$$

which may be geometrically interpreted as the union of subspaces spanned by basis vectors $\{e_k\}_{k \in S}$, where $S \in \Delta$. The associated result of [12] is key:

Definition 3 (Generalized Nullspace Property (GNUP)): Let $\boldsymbol{\Psi} \in \mathbb{R}^{n \times m}$, and let Δ be an ASC over $[\![m]\!]$. We say that $\boldsymbol{\Psi}$ satisfies the *generalized nullspace property with respect to* Δ (Δ-NUP) if $\forall \boldsymbol{h} \in \ker \boldsymbol{\Psi} \setminus \{0\}$, $\forall S \in \Delta$, $\|\boldsymbol{h}_S\|_1 < \|\boldsymbol{h}_{S^c}\|_1$. Equivalently,

$$\operatorname{nsc}(\boldsymbol{\Psi}, \Delta) := \max_{S \in \Delta} \max_{\boldsymbol{h} \in \ker \boldsymbol{\Psi} \setminus \{0\}} \frac{\|\boldsymbol{h}_S\|_1}{\|\boldsymbol{h}\|_1} < \frac{1}{2} \tag{3}$$

where $\operatorname{nsc}(\boldsymbol{\Psi}, \Delta)$ is called the *nullspace constant*.

Proposition 1: Let Δ be a simplicial complex over $[\![m]\!]$. Then any Δ-sparse solution to P_1 is the unique solution, if and only if $\boldsymbol{\Theta}$ satisfies the Δ-NUP.

This result potentially enables necessary and sufficient conditions for much more general types of sparsity patterns than are classically admissible for P_1. As a simple example, one can remark that the Δ_s-NUP is equivalent to the s-NUP, so this is a generalization; but it also encompasses e.g. group sparsity. Since the GNUP is essentially a statement about the kernel of a particular matrix, it is natural to extend this characterization to any subspace with an appropriate choice of basis: for a subspace $\mathcal{U} \subseteq \mathbb{R}^n$ define

$$\operatorname{nsc}(\mathcal{U}, \Delta) := \max_{S \in \Delta} \max_{\boldsymbol{h} \in \mathcal{U} \setminus \{0\}} \frac{\|\boldsymbol{h}_S\|_1}{\|\boldsymbol{h}\|_1}$$

and say that $\mathcal{U}$ satisfies the Δ-NUP if $\operatorname{nsc}(\mathcal{U}, \Delta) < \frac{1}{2}$.

D. Linear Dynamical Systems with Sparse Inputs

In this section, we will state the main problem of the paper. For an LDS $\Sigma = (\boldsymbol{A}, \boldsymbol{\Psi}, \boldsymbol{C})$ whose state-space equations are defined in (1), we define the associated block matrices:

$$\boldsymbol{\mathcal{O}}_N = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{CA} \\ \vdots \\ \boldsymbol{CA}^N \end{bmatrix}, Y_N = \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix}, U_N = \begin{bmatrix} \boldsymbol{u}_0 \\ \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_{N-1} \end{bmatrix}$$
$$\boldsymbol{\Gamma}_N = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \boldsymbol{C\Psi} & 0 & \cdots & 0 \\ \boldsymbol{CA\Psi} & \boldsymbol{C\Psi} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ \boldsymbol{CA}^{N-1}\boldsymbol{\Psi} & \boldsymbol{CA}^{N-2} & \cdots & \boldsymbol{C\Psi} \end{bmatrix}, \tag{4}$$

where we refer to $\boldsymbol{\mathcal{O}}_N$ as the *observability matrix* of Σ and to $\boldsymbol{\Gamma}_N$ as the *input-output matrix*. The reader should note the discrepancy in the number of entries of Y_N and U_N; we have opted to admit this asymmetry for notational convenience.

It follows that

$$Y_N = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N, \tag{5}$$

Of central interest in this paper is the case where $\forall k$, $\operatorname{supp} \boldsymbol{u}_k \in \Delta$, where Δ is a simplicial complex over $[\![m]\!]$. In this case, we say that the input $\boldsymbol{u}$ and respectively the block vector U_N is *entrywise* Δ*-sparse*; we may equivalently refer to U_N as Δ^N-sparse. We might then ask, under what conditions on what optimization problems can we recover $\boldsymbol{x}_0$ and U_N from Y_N, given the assumption that U_N is Δ^N-sparse?

The optimization problem we focus on is the following, introduced in [20]:

$$\min_{\boldsymbol{x}_0, U_N} \|U_N\|_1 \text{ s.t. } Y_N = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N. \tag{D_1}$$

This optimization problem may be thought of as an implementation of basis pursuit (P_1) for linear systems. Analogously, we would like to characterize the conditions under which this problem is well posed–i.e. no two entrywise Δ-sparse inputs and generic initial conditions produce the same output–and when (D_1) uniquely recovers such Δ-sparse inputs and initial conditions jointly. The former condition may be thought of as an injectivity condition, in the sense that the output Y_N uniquely specifies the initial condition and inputs up to Δ^N-sparsity. Of additional interest are the cases in which the state space is not sufficiently observable to uniquely determine x given Y_N, but we may still recover or uniquely characterize Δ^N-sparse U_N. To these ends, we define the following:

Definition 4: Let $\Sigma = (\boldsymbol{A}, \boldsymbol{\Psi}, \boldsymbol{C})$ be a linear system.

- Σ is *jointly* Δ^N*-injective* if $U, U' \in \mathcal{S}(\Delta^N)$ and $\boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0' + \boldsymbol{\Gamma}_N U_N' = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N \implies (\boldsymbol{x}_0, U_N) = (\boldsymbol{x}_0', U_N')$. We write the set of all such Σ as $\mathcal{R}_0^*(\Delta, N)$.
- Σ is *input* Δ^N*-injective* if $U, U' \in \mathcal{S}(\Delta^N)$ and $\boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0' + \boldsymbol{\Gamma}_N U_N' = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N \implies U_N = U_N'$. We write the set of all such Σ as $\mathcal{R}_0(\Delta, N)$.
- Σ is *jointly* Δ^N*-recoverable with* (D_1) if any solution $(\boldsymbol{x}_0, U_N)$ to (D_1) s.t. $U_N \in \mathcal{S}(\Delta^N)$ is necessarily the unique solution. We write the set of all such Σ as $\mathcal{R}_1^*(\Delta, N)$.
- Σ is *input* Δ^N*-recoverable with* (D_1) if any two solutions $(\boldsymbol{x}_0, U_N), (\boldsymbol{x}_0', U_N')$ to (D_1) s.t. $U_N, U_N' \in \mathcal{S}(\Delta^N)$ satisfy $U_N = U_N'$. We write the set of all such Σ as $\mathcal{R}_1(\Delta, N)$.

The condition established in [20] for when Σ is jointly Δ_s^N-recoverable with (D_1) is based on the coherence $\mu : \mathbb{R}^{n \times m} \to [0, 1]$, defined as

$$\mu(\boldsymbol{\Theta}) = \max_{i \neq j} \frac{|\langle \boldsymbol{\Theta}_i, \boldsymbol{\Theta}_j \rangle|}{\|\boldsymbol{\Theta}_i\|_2 \|\boldsymbol{\Theta}_j\|_2} \tag{6}$$

Henceforth, we define $\boldsymbol{P}_N^\perp := \boldsymbol{I} - \boldsymbol{\mathcal{O}}_N \boldsymbol{\mathcal{O}}_N^+$, the orthogonal projection onto the orthogonal complement of the column space of $\boldsymbol{\mathcal{O}}_N$. The main result of [20], which may be implicitly read as a sufficient condition for $\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$ to satisfy the Ns-NUP, is as follows:

Proposition 2: If $\ker \boldsymbol{\mathcal{O}}_N = 0$ and

$$\mu(\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N) \;<\; \frac{1}{2Ns - 1}$$

then Σ is jointly Δ_s^N-recoverable with (D_1).

As is typical of coherence-based sparse recovery guarantees, it was found that this bound was enormously conservative for even modest N. It is also remarked that the condition in proposition 2 also guarantees that any $\boldsymbol{x}_0$ and U_N such that $|\operatorname{supp}(U_N)| \leq Ns$ is also recovered uniquely by (D_1), so there is reason to suspect this condition could be tightened.

III. NECESSARY AND SUFFICIENT CONDITIONS FOR JOINT STATE AND SPARSE INPUT RECOVERY

Let Δ be an ASC over $[\![m]\!]$ and Σ a linear system as in (1). Our first observation is that the difference between joint and input-only recoverability/injectivity is just a matter of observability:

Lemma 1: $\Sigma \in \mathcal{R}_p^*(\Delta, N)$ if and only if $\ker \boldsymbol{\mathcal{O}}_N = 0$ and $\Sigma \in \mathcal{R}_p(\Delta, N)$.

Proof: By definition, $\mathcal{R}_p^*(\Delta, N) \subseteq \mathcal{R}_p(\Delta, N)$. Suppose $\ker \boldsymbol{\mathcal{O}}_N \neq 0$, then $\boldsymbol{x}_0$ cannot be uniquely determined by a constraint on $\boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0$, a contradiction. Hence $\Sigma \in \mathcal{R}_p^*(\Delta, N)$ implies $\ker \boldsymbol{\mathcal{O}}_N = 0$ and $\Sigma \in \mathcal{R}_p(\Delta, N)$.

Now suppose $\Sigma \in \mathcal{R}_p(\Delta, N)$ and that $\ker \boldsymbol{\mathcal{O}}_N = 0$. Then $\forall (\boldsymbol{x}_0, U_N), (\boldsymbol{x}_0', U_N')$ such that $U_N, U_N' \in \mathcal{S}(\Delta^N)$, $\boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0' + \boldsymbol{\Gamma}_N U_N' = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N \implies U_N = U_N'$, and therefore $\boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0'$. Since $\ker \boldsymbol{\mathcal{O}}_N = 0$, $\boldsymbol{x}_0 = \boldsymbol{x}_0'$. ■

In [20], implicit in the use of coherence for the main result was that when $\ker \boldsymbol{\mathcal{O}}_N = 0$, $\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$ satisfying the Ns-NUP is sufficient Σ to be jointly Δ^N-recoverable with (D_1). Per lemma 1, we may suspect that if we relax the condition of observability, we may still determine a condition on when (D_1) recovers the input, through conditions on $\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$. This is reflected in the fact that the following are equivalent:

$$\exists \boldsymbol{x}_0, \boldsymbol{x}_0', \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N = \boldsymbol{\mathcal{O}}_N \boldsymbol{x}_0' + \boldsymbol{\Gamma}_N U_N' \tag{7}$$
$$\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N U_N = \boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N U_N' \tag{8}$$

For determining sparse input recoverability and injectivity then, we will see it is only necessary to consider properties of $\ker \boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N$.

In the case of regular sparsity, there are several equivalent ways to establish uniqueness of sparse solutions. The typical way is a rank condition on all collections of $2s$ columns of a matrix. Here we generalize this slightly, to Δ-sparsity:

Lemma 2: Let Δ be an ASC over $[\![m]\!]$. $\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{S}(\Delta), \boldsymbol{\Theta x} = \boldsymbol{\Theta x}' \implies \boldsymbol{x} = \boldsymbol{x}'$ if and only if $\forall S, S' \in \Delta$, $\ker \boldsymbol{\Theta}_{S \cup S'} = 0$. When either condition is satisfied, we say $\boldsymbol{\Theta}$ is Δ-injective.

Proof: Suppose $\exists S, S' \in \Delta$, $\ker \boldsymbol{\Theta}_{S\cup S'} \neq 0$, then let $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^m$ distinct such that $\operatorname{supp} \boldsymbol{x} = S$ and $\operatorname{supp} \boldsymbol{x}' = S'$, and $\boldsymbol{\Theta}(\boldsymbol{x} - \boldsymbol{x}') = 0$. Then $\boldsymbol{\Theta x} = \boldsymbol{\Theta x}'$ but $\boldsymbol{x} \neq \boldsymbol{x}'$, a contradiction. Conversely, suppose $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{S}(\Delta)$ such that $\boldsymbol{\Theta x} = \boldsymbol{\Theta x}'$ but $\boldsymbol{x} \neq \boldsymbol{x}'$, then where $S = \operatorname{supp} \boldsymbol{x}$ and $S' = \operatorname{supp} \boldsymbol{x}'$, $\boldsymbol{\Theta}_{S\cup S'}(\boldsymbol{x} - \boldsymbol{x}')_{S\cup S'} = 0$, so $\ker \boldsymbol{\Theta}_{S\cup S'} \neq 0$. ∎

This notion enables concise necessary and sufficient characterizations of when Δ-sparse solutions are unique, and we will put it to good use in Theorem 1.

Shortly after the publication of [20], Kafashan et al. produced sufficient conditions for both joint Δ_s^N-recovery with (D_1) and a related optimization problem incorporating the assumption of noise [21]. Lemma 1 of [21] gets close to being a necessary condition when $\Delta = \Delta_s$, but it makes use of the restricted isometry constant rather than purely rank-type constraints, and is not stated as necessary and sufficient. We provide a version of the necessary and sufficient condition in the style of [21] to reflect this contribution, making only the change of the restricted isometry condition to Δ-injectivity:

Theorem 1 (Δ^N-Injectivity of Σ): Let Δ be an ASC over $[\![m]\!]$ and $\Sigma = (\boldsymbol{A}, \boldsymbol{\Psi}, \boldsymbol{C})$ a linear system with state space of dimension n. The following are equivalent:

1) Σ is jointly Δ^N-injective
2) $\forall S, S' \in \Delta^N$, $\operatorname{rank}[\boldsymbol{\mathcal{O}}_N\ (\boldsymbol{\Gamma}_N)_{S\cup S'}] = n + \operatorname{rank}((\boldsymbol{\Gamma}_N)_{S\cup S'})$ and $\boldsymbol{C\Psi}$ is Δ-injective
3) $\ker \boldsymbol{\mathcal{O}}_N = 0$ and $\boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N$ is Δ^N-injective.

Proof: Suppose Σ is jointly Δ^N-injective, then if $U_N, U_N' \in \mathcal{S}(\Delta)$, $\boldsymbol{\mathcal{O}}_N\boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N = \boldsymbol{\mathcal{O}}_N\boldsymbol{x}_0' + \boldsymbol{\Gamma}_N U_N' \implies \boldsymbol{x}_0' = \boldsymbol{x}_0$ and $U_N = U_N'$. It follows that for every $S, S' \in \Delta^N$, $[\boldsymbol{\mathcal{O}}_N\ (\boldsymbol{\Gamma}_N)_{S\cup S'}]$ is full column rank, so $\operatorname{rank}[\boldsymbol{\mathcal{O}}_N\ (\boldsymbol{\Gamma}_N)_{S\cup S'}] = n + \operatorname{rank}((\boldsymbol{\Gamma}_N)_{S\cup S'})$, and $\ker(\boldsymbol{\Gamma}_N)_{S\cup S'} = 0$. It follows that every $((\boldsymbol{\Gamma}_N)_k)_{S_k\cup S_k'}$ is full rank, so $(\boldsymbol{\mathcal{O}}_{N-k}\boldsymbol{\Psi})_{S_k\cup S_k'}$ is full rank $\forall k \in [\![N]\!]$. In particular we may take $k = 0$, so $\forall S, S' \in \Delta$, $(\boldsymbol{\mathcal{O}}_0\boldsymbol{\Psi})_{S\cup S'} = (\boldsymbol{C\Psi})_{S\cup S'}$ is full rank, and therefore $\boldsymbol{C\Psi}$ is Δ-injective.

Likewise, if $(\boldsymbol{C\Psi})_{S\cup S'}$ is full rank, since $\boldsymbol{\Gamma}_N$ is block triangular with $\boldsymbol{C\Psi}$ on the diagonal, we have that $(\boldsymbol{\Gamma}_N)_{S\cup S'}$ is full rank for every $S, S' \in \Delta^N$. It follows that if $\operatorname{rank}[\boldsymbol{\mathcal{O}}_N\ (\boldsymbol{\Gamma}_N)_{S\cup S'}] = n + \operatorname{rank}((\boldsymbol{\Gamma}_N)_{S\cup S'})$, then $[\boldsymbol{\mathcal{O}}_N\ (\boldsymbol{\Gamma}_N)_{S\cup S'}]$ is full rank, so if U_N is Δ^N-sparse and $\boldsymbol{x}_0$ is generic such that $Y_N = \boldsymbol{\mathcal{O}}_N\boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N$, they are unique. Therefore, Σ is jointly Δ^N-injective. We have thus shown that (1) $\iff$ (2).

Note that from the above, we have (2) if and only if $\forall S, S' \in \Delta^N$, $\ker[\boldsymbol{\mathcal{O}}_N\ (\boldsymbol{\Gamma}_N)_{S\cup S'}] = 0$, which is the case if and only if $\ker \boldsymbol{\mathcal{O}}_N = 0$, $\ker(\boldsymbol{\Gamma}_N)_{S\cup S'} = 0$, and $\operatorname{im}\boldsymbol{\mathcal{O}}_N \cap \operatorname{im}(\boldsymbol{\Gamma}_N)_{S\cup S'} = 0$. The third point holds iff $\operatorname{rank}\boldsymbol{P}_N^{\perp}(\boldsymbol{\Gamma}_N)_{S\cup S'} = \operatorname{rank}(\boldsymbol{\Gamma}_N)_{S\cup S'}$, so this is again equivalent to $\ker \boldsymbol{\mathcal{O}}_N = 0$ and $\ker \boldsymbol{P}_N^{\perp}(\boldsymbol{\Gamma}_N)_{S\cup S'} = \ker(\boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N)_{S\cup S'} = 0$. We conclude by noting that this is equivalent to $\ker \boldsymbol{\mathcal{O}}_N = 0$ and $\boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N$ being Δ^N-injective. ∎

Having shown the result for uniqueness of solutions/sparse injectivity, we proceed to the problem of recoverability. Recalling that we are only interested in Δ^N-sparse solutions U_N, we could obtain a necessary and sufficient condition from the GNUP if this support pattern is an ASC. Technically, as Δ^N is a set of tuples of sets, it cannot be a simplicial complex, but if one instead interprets these tuples as disjoint unions, we uncover an ASC structure:

Lemma 3: Δ^N is a simplicial complex up to bijection.

Proof: It is clear that every $(S_k)_{k\in[\![N]\!]} \in \Delta^N$ can be identified with the set $\tilde{S} = \bigcup_{k\in[\![N]\!]}\{k\} \times S_k$. Let $S \in \Delta^N$ and suppose $\tilde{S}' \subseteq \tilde{S}$. Then for all k, $\{k\} \times S_k' \subseteq \{k\} \times S_k \implies S_k' \subseteq S_k$, so $S_k' \in \Delta$. Therefore $S' \in \Delta^N$. ∎

It is then clear that we can apply the GNUP in this context, to obtain necessary and sufficient conditions on Δ^N-recovery.

Theorem 2 (Δ^N-Recoverability with (D_1)): Let Δ be an ASC over $[\![m]\!]$ and $\Sigma = (\boldsymbol{A}, \boldsymbol{\Psi}, \boldsymbol{C})$ a linear system. The following are equivalent:

1) Σ is jointly Δ^N-recoverable with (D_1).
2) $\ker \boldsymbol{\mathcal{O}}_N = 0$ and $\boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N$ satisfies the Δ^N-NUP.

Proof: Recall that Σ is jointly Δ^N-recoverable if and only if $\ker \boldsymbol{\mathcal{O}}_N = 0$ and for any two solutions $(\boldsymbol{x}_0, U_N), (\boldsymbol{x}_0, U_N')$ to $D_1(\Sigma, N)$ such that $U_N, U_N' \in \mathcal{S}(\Delta^N)$ satisfy $U_N = U_N'$. This is equivalent to any Δ^N-sparse solution to the optimization problem $\min_{U_N}\|U_N\|_1$ s.t. $\exists \boldsymbol{x}_0, Y_N = \boldsymbol{\mathcal{O}}_N\boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N$ being the unique solution, and we have that $\exists \boldsymbol{x}_0, Y_N = \boldsymbol{\mathcal{O}}_N\boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N$ if and only if $\boldsymbol{P}_N^{\perp}Y_N = \boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N U_N$. Therefore, any Δ^N-sparse solution to $\min_{U_N}\|U_N\|_1$ s.t. $\exists \boldsymbol{x}_0, Y_N = \boldsymbol{\mathcal{O}}_N\boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N$ is necessarily unique if and only if $\boldsymbol{P}_N^{\perp}\boldsymbol{\Gamma}_N$ satisfies the Δ^N-NUP, and we may conclude. ∎

As system-theoretic statements, these conditions establish a deadbeat unknown-input state estimator and input reconstructor for linear systems with sparse inputs, analogous to the generic input version in [23]. From the point of view of sparse recovery, these conditions mirror the role of the spark and the standard nullspace property for P_0, P_1. Like these conditions, they are clearly NP-hard to verify; however, it is clear that in the case of regular sparsity, the number of supports to check to verify the Δ_s^N-NUP is far smaller than for the Ns-NUP. We will see this idea reflected in the next section: in many cases, it indeed suffices to check much easier conditions.

IV. Interpretable Conditions for Δ^N-Recoverability with (D_1)

In the last section, we saw that by casting Σ as a block matrix-vector system, standard sparse recovery arguments yield a spark-like condition for Δ^N-injectivity, and subsequent application of the generalized nullspace property yields an analogous condition for Δ^N-recoverability. Once we have $\ker \mathcal{O}_N = 0$, each of these conditions are entirely determined by attributes of $\ker P_N^\perp \boldsymbol{\Gamma}_N$. This space is precisely the subspace of inputs U_N for which there exists an initial condition $\boldsymbol{x}_0$, $\boldsymbol{\Gamma}_N U_N = \mathcal{O}_N \boldsymbol{x}_0$. By constructing lower and upper bounds on $\operatorname{nsc}(P_N^\perp \boldsymbol{\Gamma}_N, \Delta^N)$ via $\operatorname{nsc}(\boldsymbol{C\Psi}, \Delta)$ and $\operatorname{nsc}((\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C}, \Delta)$, we obtain conditions that are respectively necessary and sufficient for Σ to be input Δ^N-recoverable.

A. A Necessary Condition

In [20], it was empirically observed that $\mu(\boldsymbol{C\Psi}) \leq \mu(\boldsymbol{\Gamma}_N) \leq \mu(\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N)$. The intuition one is tempted to derive from this is that a more incoherent $\boldsymbol{C\Psi}$ indicates a greater chance of (D_1) succeeding. To concretize this idea, one need look no further than the question of recovering the very last input to a system, left untouched by the system dynamics. Specifically, it is useful to consider the one-to-one correspondence between elements of $\ker \boldsymbol{C\Psi}$ and inputs $U_N \in \ker \boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$ with all zero entries except for the last:

Lemma 4: Suppose $\boldsymbol{u}$ is an input such that $\forall k < N, \boldsymbol{u}_k = 0$ and $\boldsymbol{u}_{N-1} \in \ker \boldsymbol{C\Psi}$. Then $U_N \in \ker \boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$.

Proof: Suppose U_N is as described, then with $\boldsymbol{x}_0 = 0$, $k < N \implies \boldsymbol{x}_k = 0 \implies \boldsymbol{y}_k = 0$. We also have that $\boldsymbol{y}_N = \boldsymbol{C}\boldsymbol{x}_N = \boldsymbol{C\Psi}\boldsymbol{u}_{N-1} = 0$. Therefore, $U_N \in \ker \boldsymbol{\Gamma}_N \subseteq \ker \boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$. ■

This suggests the following natural necessary condition:

Proposition 3: Σ can only be input Δ^N-recoverable with (D_1) (resp. injective) if $\boldsymbol{C\Psi}$ satisfies the Δ-NUP (resp. is Δ-injective)

Proof: The case of injectivity is implied by characterization 2 of Theorem 1. Suppose Σ is Δ^N-recoverable. Then by lemma 4, for any input $\boldsymbol{u}$ such that $\forall k < N$, $\boldsymbol{u}_k = 0$ and $\boldsymbol{u}_{N-1} \in \ker \boldsymbol{C\Psi}$, U_N is the unique input component of the solution to (D_1) if U_N is Δ^N-sparse, so $\min_U \|U_N\|_1 = \min_{\boldsymbol{v}} \|\boldsymbol{v}\|_1$ s.t. $\boldsymbol{C\Psi}\boldsymbol{u}_{N-1} = \boldsymbol{C\Psi}\boldsymbol{v}$ always recovers Δ-sparse $\boldsymbol{u}_{N-1}$, hence $\boldsymbol{C\Psi}$ satisfies the Δ-NUP. ■

Another way to put this is in terms of the nullspace constant:

Corollary 1: Σ can only be input Δ^N-recoverable if $\operatorname{nsc}(\boldsymbol{C\Psi}, \Delta) < 0.5$.

It is clear that if $\boldsymbol{C\Psi}$ does not satisfy Δ-NUP, there will always exist entrywise Δ-sparse inputs which cannot be recovered by the problem (D_1). As this will be the case for any N, this is in a sense the tightest necessary condition for this class of problems.

B. A Sufficient Condition

We will now show that whenever the output of a system is uniformly zero, at each time point k *regardless* of initial condition, one must have $\boldsymbol{u}_k \in (\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C}$:

Lemma 5: Let $\boldsymbol{u}$ be an input such that $U_N \in \ker \boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$. Then $\forall k < N$, $\boldsymbol{u}_k \in (\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C}$.

Proof: Let $\boldsymbol{u}$ be as above, and let $\boldsymbol{x}_0$ be such that $\mathcal{O}_N \boldsymbol{x}_0 + \boldsymbol{\Gamma}_N U_N = Y_N = 0$. Suppose the claim is false. Then $\exists k \leq N-2$, $\boldsymbol{C\Psi}\boldsymbol{u}_k \notin \boldsymbol{CA}\ker \boldsymbol{C}$, so if $\boldsymbol{x}_k \in \ker \boldsymbol{C}$, $\boldsymbol{C}(\boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{\Psi}\boldsymbol{u}_k) = \boldsymbol{y}_{k+1} \neq 0$. Since we assumed $Y_N = 0 \implies \boldsymbol{y}_{k+1} = 0$, this cannot be, so $\boldsymbol{x}_k \notin \ker \boldsymbol{C}$, which gives our contradiction as $\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k \neq 0$. ■

A fact not immediately visible is that this subspace arises as the kernel of $\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1$:

Lemma 6:

$$\ker \boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1 = (\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C}$$

Proof:

$$\begin{aligned} \ker \boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1 &= \{\boldsymbol{u} : \boldsymbol{\Gamma}_1 \boldsymbol{u} \in \operatorname{im} \mathcal{O}_1\} \\ &= \{\boldsymbol{u} : \exists \boldsymbol{x}_0, \boldsymbol{C}\boldsymbol{x}_0 = 0 \text{ and } \boldsymbol{C\Psi}\boldsymbol{u} = \boldsymbol{CA}\boldsymbol{x}_0\} \\ &= \{\boldsymbol{u} : \boldsymbol{C\Psi}\boldsymbol{u} \in \boldsymbol{CA}\ker \boldsymbol{C}\} \\ &= (\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C} \end{aligned}$$

■

This leads us to the following:

Proposition 4: Σ is input Δ^N-recoverable for every N if and only if $\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1$ satisfies the Δ-NUP.

Proof: Σ is input Δ^N-recoverable for every N only if Σ is input Δ-recoverable; by Theorem 2, we have the forward implication. Since lemma 5 ensures that every entry $\boldsymbol{u}_k$ of an input $U_N \in \ker \boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$ satisfies $\boldsymbol{u}_k \in (\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C} \overset{\text{lemma 6}}{=} \ker \boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1$, for any $S \in \Delta^N$,

$$\begin{aligned} \|(U_N)_S\|_1 &= \sum_{k \in [\![N]\!]} \|(\boldsymbol{u}_k)_{S_k}\|_1 \\ &< \sum_{k \in [\![N]\!]} \operatorname{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta) \|\boldsymbol{u}_k\|_1 \\ &= \operatorname{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta) \|U_N\|_1 \\ \implies \frac{\|(U_N)_S\|_1}{\|U_N\|_1} &< \operatorname{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta) < \frac{1}{2} \\ &\implies \operatorname{nsc}(\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N, \Delta^N) < \frac{1}{2} \end{aligned}$$

therefore $\boldsymbol{P}_N^\perp \boldsymbol{\Gamma}_N$ satisfies the Δ^N-NUP, and we conclude by Theorem 2. ■

This may be also be phrased as a sufficient condition for any N, in terms of the nullspace constant:

Corollary 2: Σ is input Δ^N recoverable if $\operatorname{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta) < 0.5$.

Since this condition is clearly also necessary for $N = 1$, we might expect this to be closer to necessity for smaller values of N.

C. Tight Cases

While these conditions only form the tip of the iceberg in terms of possible guarantees for Δ^N-recovery, they are unlike existing conditions in the literature as they each achieve necessity and sufficiency for certain classes of systems. Implicit in proposition 4 is the necessary and sufficient condition for input Δ-recoverability, though this can only indicate joint Δ-recoverability when $\operatorname{rank} \mathcal{O}_1 = n$. The major condition leading to these conditions being tight is when $\ker \boldsymbol{C\Psi} = (\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker \boldsymbol{C}$, which is equivalent to $\boldsymbol{A}\ker \boldsymbol{C} \subseteq \ker \boldsymbol{C}$–that is, the system is unobservable. This hints at an intuition that systems that are more observable, e.g. systems that quickly divulge information about their initial conditions, may actually result in nullspace constants of these spaces that are further apart. We also note that $(\boldsymbol{C\Psi})^{-1}\boldsymbol{CA}\ker\boldsymbol{C} = (\boldsymbol{C\Psi})^{+}\boldsymbol{CA}\ker\boldsymbol{C}+\ker\boldsymbol{C\Psi}$, so characteristics of the space $(\boldsymbol{C\Psi})^{+}\boldsymbol{CA}\ker\boldsymbol{C}$ directly mediate how close the conditions established in this section are to one another. We might then expect that the gap between them to scale with $\dim\ker\boldsymbol{C}$, and therefore to be small when p is not much less than n.

V. NUMERICAL VALIDATION

To validate and provide intuition for our results, we perform two types of numerical experiments on ensembles of random LDSs with s-sparse inputs, i.e. Δ-sparse with $\Delta \in \{\Delta_s : s \in \mathbb{N}\}$:

1) We evaluate the ability of (D_1) to jointly recover the initial state and sparse inputs for different system parameters and sparsity levels.
2) We explore the relationship between $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$ as a function of system parameters and sparsity levels.

Throughout, we employ the 1-step TSA branch-and-bound algorithm of [24] to compute nullspace constants to a tolerance of ± 0.05. This level of precision was chosen to permit the estimation of nullspace constants for much larger sparsity levels with a statistically significant amount of systems than would be permitted if we were to compute them exactly.

A. System Generation and Computing Information

For a given n, m, p, systems $\Sigma = (\boldsymbol{A}, \boldsymbol{\Psi}, \boldsymbol{C})$ were generated randomly, mirroring the strategy of [20]:

- $\boldsymbol{A}$ is i.i.d. Gaussian in each entry with variance $1/n$, s.t. each eigenvalue λ of $\boldsymbol{A}$ satisfies $|\lambda| < 0.9$.
- $\boldsymbol{\Psi}$ is i.i.d Gaussian in each entry with variance $1/n$.
- $\boldsymbol{C}$ is i.i.d. Gaussian with unit variance.

For all systems considered, we choose $m = 20$, to ensure tractability of Δ_s-nullspace constant estimates. We also restrict our attention to $n \le m$ and $p \le m$, focusing on the usual assumption of "overcompleteness" of $\boldsymbol{\Psi}$. We analyze a total of 25173 systems across all parameter ranges. We conducted all experiments and analysis in Python, using CVXPY with the GUROBI solver for implementation of linear programming.

B. Success of recovery

The first analysis we present provides numerical support for the claims of necessity and sufficiency of the conditions established in corollaries 1 and 2, that $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) < 0.5$ is necessary and $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) < 0.5$ is sufficient for Σ to be input Δ_s^N-recoverable.

For each system, we choose a random $s \le 10$, and simulate 30 random combinations of initial conditions and entrywise s-sparse inputs of length $N = n$. As in [20], we sample each entry of the initial conditions uniformly from $[-5, 5]$, choose every support $S_k \subseteq [\![m]\!]$ uniformly without replacement s.t. $|S_k| = s$, and sample each nonzero input entry uniformly on $[-5, 5]$. In Table I, we provide the fraction of systems for which there was at least one imperfect joint recovery of the 30 trials conducted as a function of the nullspace constants $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$, which were determined up to being less than 0.5, greater than 0.5, or in $[0.45, 0.55]$. From the table, we see our conditions behave as expected: all systems with $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) < 0.5$ exhibit perfect recovery; a majority of systems with $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) > 0.5$ exhibit imperfect recovery; and $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) \ge \mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$ for all systems. In particular, there were many systems such that, for the sparsity level considered, $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) < 0.5$ was not sufficient for perfect input recovery, as $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) > 0.5$.

TABLE I: Number of systems with at least one imperfect joint recovery across 30 trials out of total occurrences of systems, with given ranges of $\sigma_{\boldsymbol{C\Psi}} := \mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\sigma_{\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1} := \mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$. $n \in \{12, 16, 20\}$, $1 \le p \le 20$, $1 \le s \le 5$ and $m = 20$. $\sigma \sim 0.5$ if $|\sigma - 0.5| \le 0.05$ (experimental tolerance).

	$\sigma_{\boldsymbol{C\Psi}} < 0.5$	$\sigma_{\boldsymbol{C\Psi}} \sim 0.5$	$\sigma_{\boldsymbol{C\Psi}} > 0.5$
$\sigma_{\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1} > 0.5$	153/800	34/289	9068/9866
$\sigma_{\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1} \sim 0.5$	0/116	60/336	2/10
$\sigma_{\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1} < 0.5$	0/2100	0/0	0/0

Figure 1 visualizes this data in a different way, presenting a joint recovery phase transition plot over p and s for $n \in \{12, 16, 20\}$ and $N = n$. The intensity of each pixel indicates the empirical probability of a system with dimensions p, n, m and sparsity level s exhibiting imperfect joint recovery. The colormap normalization is in log-scale to better illustrate the case of perfect recovery for all trials, which we plot as white. The red and blue outlined regions indicate (p, n, m, s) such that every system simulated satisfied $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) > 0.5$ (red, failure of necessary condition) or $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) < 0.5$ (blue, satisfaction of sufficient condition). We observe that there are no white pixels contained in the red outlined regions, supporting the claim of $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) > 0.5$ being necessary for joint Δ_s^n-recoverability. The red outlined region also becomes strictly smaller as n increases, as expected as this indicates $\boldsymbol{C\Psi}$ with generically higher rank. Likewise, we observe that there are no colored pixels contained in the blue outlined regions, supporting the claim that $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) < 0.5$ is sufficient for joint Δ_s^n-recoverability. The blue regions appear to shrink to the right as n increases, widening the gap in s between the red and blue outlined regions for smaller p. This reflects the fact that increased n will generically result in an increase in the dimension of $\ker\boldsymbol{C}$, and thus an increase in the dimension of $\ker\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1$.

C. Analysis of NSC Relationship

We conclude our analysis by taking a closer look at the relationship between the two main quantities motivated in this work, $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$. Here we do not terminate once determining whether the constant is above or below 0.5, as was done for the previous section, opting instead to compute bounds on these constants for each system

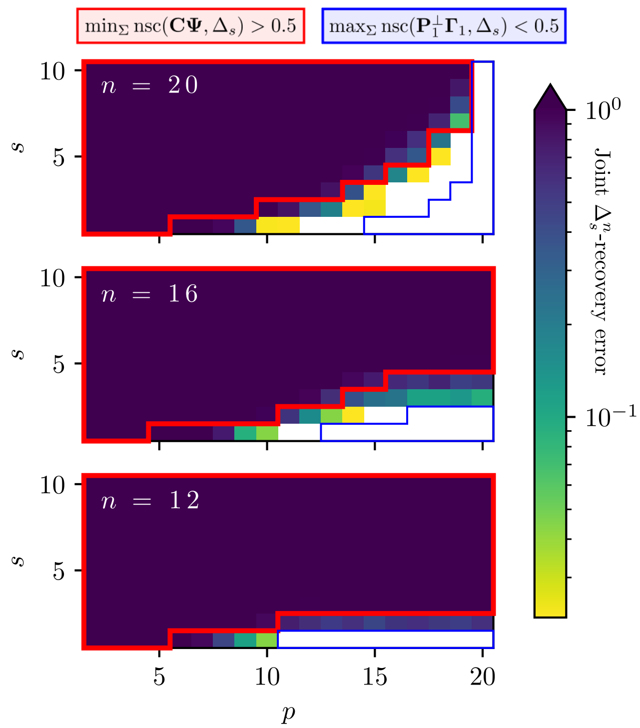


Fig. 1: Empirical probability of imperfect joint recovery of entrywise Δ_s-sparse signals and generic initial conditions as a function of s and p with $m = 20$, $n \in \{12, 16, 20\}$, and $N = n$; white indicates all signals perfectly recovered. Within the red outlines are (s, p) such that all simulated systems satisfied $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) > 0.5$. Within the blue outline are (s, p) such that all simulated systems satisfied $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s) < 0.5$.

up to a tolerance of ± 0.05. In figure 2, for various n, m, p, s we plot $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ against $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$. Each plotted point represents the midpoint of the computed bounds on these constants for a given Σ and sparsity level s. We see that for every set of parameters considered, the nullspace constants tend to increase with the sparsity level, and that $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) \le \mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$ as expected from the fact that $\ker \boldsymbol{C\Psi} \subseteq (\boldsymbol{C\Psi})^+ \boldsymbol{CA} \ker \boldsymbol{C} + \ker \boldsymbol{C\Psi} = \ker \boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1$.

In the top set of figures we fix $n = 19$ and $m = 20$, and illustrate the trend of $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$ as p increases. For $p = 11$, $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$ is quite pessimistic, echoing the wide gap between the red and blue regions of figure 1 for $n = 20$ for moderate p. As p increases, points approach the diagonal and tend to decrease in magnitude, reflecting the intuitively more favorable recovery properties of a larger number of measurements. When $p = n$, we see that the two constants become equal; this reflects the fact that $\boldsymbol{C} \in \mathbb{R}^{p \times n}$ will be generically full rank and thus $\ker \boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1 = (\boldsymbol{C\Psi})^+ \boldsymbol{CA} \ker \boldsymbol{C} + \ker \boldsymbol{C\Psi} = \ker \boldsymbol{C\Psi}$.

In the bottom set of figures we fix $p = 11$ and $m = 20$, and illustrate the trend of $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$ as n increases. Up to our tolerance, we see again that for $n = 11$, $n = p$ and so the constants are equal. With the increase in n, we observe that $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ stays roughly fixed, while $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$ increases. This is qualitatively similar to the effect of decreasing p, but in that case, *both* constants were affected. This comes as no surprise, as for $p \le n \le m$ we expect the distribution of $\boldsymbol{C\Psi}$ to be relatively unaffected, but $\dim \ker \boldsymbol{C}$ to increase, and thus for $\dim \ker \boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1$ to increase. Another, more intuitive interpretation is that by increasing n with p fixed, there is an increased amount of interference from the initial condition mixed in with the sparse inputs, and so guarantees on the recovery of even the sparse inputs alone will naturally worsen for recovery with a given number of time steps N. The fact that recovery empirically does better across the board for larger n in figure 1 is explained by the fact that with a larger state space, a larger number of time steps will continue to yield insight about previous inputs and the initial condition.

Overall, these plots support the point that propositions 3 and 4 can actually be fairly tight, provided that p is not much less than n. Larger p is in general seen to be better across the board, while increasing n relative to p results in less favorable sufficient guarantees. Importantly, they also reinforce the fact that in most, but not all, cases, both $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ and $\mathrm{nsc}(\boldsymbol{P}_1^\perp \boldsymbol{\Gamma}_1, \Delta_s)$ will be either less than or greater than 0.5. The cases for which they disagree are few, indicating potential practical utility for these quantities for determining joint Δ^N-recovery with (D_1) in applications.

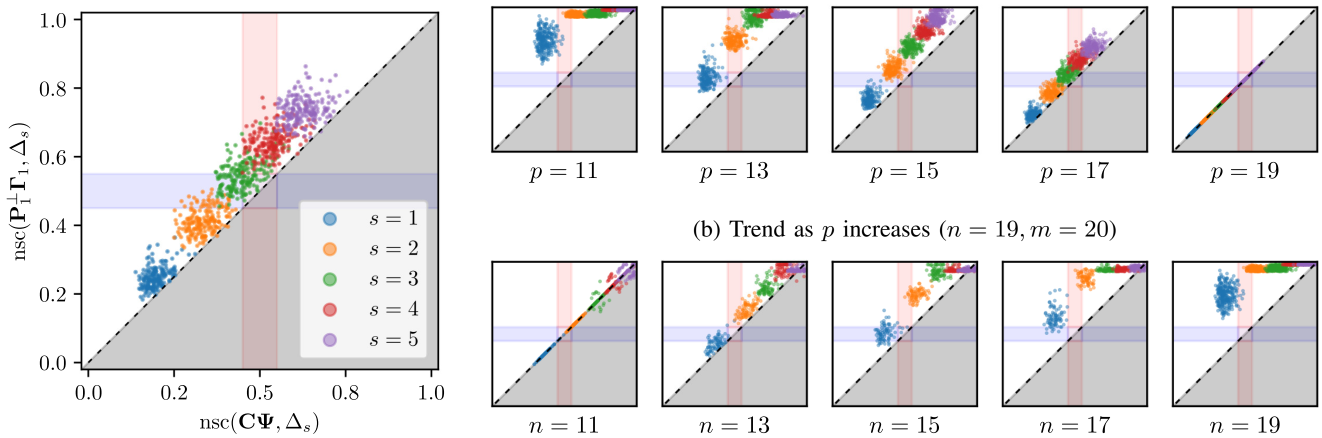
VI. Conclusions and Future Work

We have presented the first necessary and sufficient conditions under which the joint recovery of generic initial conditions and sparse inputs for a linear dynamical system is well-posed and may be carried out via ℓ_1-minimization. Leveraging this characterization, we further provide two simple necessary, and sufficient conditions for joint Δ^N-recoverability. In contrast to previous work, these conditions have intuitive justifications and can be computationally verified for most systems. Through ℓ_1-recovery experiments on random systems, we showed that these conditions are useful indicators of recovery performance.

We have identified several exciting avenues for future research building on the necessary and sufficient conditions introduced here. One direction is to extend these results to general systems with feedthrough terms (nonzero $\mathbf{D}$ matrix) and broaden the scope of recovery results to incorporate recovery with delay. More generally, we aim to explore notions of strong observability for linear systems with sparse inputs. Indeed, there are many significant results in linear systems theory that we believe admit direct extensions to the case of sparse inputs, and for which necessary and sufficient conditions using ℓ_1-minimization and related characterizations may be achieved by building on the generalized nullspace property, just as was illustrated here.

Acknowledgements

The authors would like to thank Dr. Adam Charles for invaluable discussions during the preparation of this manuscript.



(a) Example: $n = 19, m = 20, p = 17$

(b) Trend as p increases ($n = 19, m = 20$)

(c) Trend as n increases ($p = 11$, $m = 20$)

Fig. 2: Distribution of Δ_s-nullspace constants for random systems up to a tolerance of 0.05 (11656 total systems). (2a) shows in detail the case of $n = 19, m = 20, p = 17$. The red region is where $|\,\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) - 0.5| < 0.05$, and the blue region indicates $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s) - 0.5| < 0.05$. White regions indicate that $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s), \mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$ are each greater/less than 0.5 with certainty. The shaded region under the diagonal corresponds to the impossible case $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s) > \mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$. (2b) illustrates that as p increases, both $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s), \mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$ tend to decrease, with $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$ converging to $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$. (2c) illustrates that as n increases, $\mathrm{nsc}(\boldsymbol{P}_1^{\perp}\boldsymbol{\Gamma}_1, \Delta_s)$ increases while $\mathrm{nsc}(\boldsymbol{C\Psi}, \Delta_s)$ is relatively unaffected.

References

[1] R. E. Kalman, "Mathematical Description of Linear Dynamical Systems," in *Journal of the Society for Industrial and Applied Mathematics Series A Control*, vol. 1, no. 2, 1963, pp. 152–192.

[2] H. Payne and L. Silverman, "On the discrete time algebraic Riccati equation," in *IEEE Transactions on Automatic Control*, vol. 18, no. 3, 1973, pp. 226–234.

[3] M. Hou and R. J. Patton, "Input Observability and Input Reconstruction," in *Automatica*, vol. 34, no. 6, 1998, pp. 789–794.

[4] A. Ansari and D. S. Bernstein, "Deadbeat unknown-input state estimation and input reconstruction for linear discrete-time systems," in *Automatica*, vol. 103, no. C, 2019, pp. 11–19.

[5] J. Kurek, "The state vector reconstruction for linear systems with unknown inputs," in *IEEE Transactions on Automatic Control*, vol. 28, no. 12, 1983, pp. 1120–1122.

[6] B. Molinari, "A strong controllability and observability in linear multivariable control," in *IEEE Transactions on Automatic Control*, vol. 21, no. 5, 1976, pp. 761–764.

[7] Y. C. Eldar and H. Bolcskei, "Block-sparsity: Coherence and efficient recovery," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 2885–2888.

[8] R. Jenatton, J.-Y. Audibert, and F. Bach, "Structured variable selection with sparsity-inducing norms," *The Journal of Machine Learning Research*, vol. 12, pp. 2777–2824, 2011.

[9] L. He and L. Carin, "Exploiting Structure in Wavelet-Based Bayesian Compressive Sensing," in *IEEE Transactions on Signal Processing*, vol. 57, no. 9, 2009, pp. 3488–3497.

[10] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-Based Compressive Sensing," in *IEEE Transactions on Information Theory*, vol. 56, no. 4, 2010, pp. 1982–2001.

[11] E. Elhamifar and R. Vidal, "Block-Sparse Recovery via Convex Optimization," in *IEEE Transactions on Signal Processing*, vol. 60, no. 8, 2012, pp. 4094–4107.

[12] M. D. Kaba, M. Zhao, R. Vidal, D. P. Robinson, and E. Mallada, "What is the largest sparsity pattern that can be recovered by 1-norm minimization?" *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 3060–3074, 2021.

[13] M. B. Wakin, B. M. Sanandaji, and T. L. Vincent, "On the observability of linear systems from random, compressive measurements," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 4447–4454.

[14] G. Joseph and C. R. Murthy, "On the Observability of a Linear System With a Sparse Initial State," in *IEEE Signal Processing Letters*, vol. 25, no. 7, 2018, pp. 994–998.

[15] ——, "Measurement Bounds for Observability of Linear Dynamical Systems Under Sparsity Constraints," in *IEEE Transactions on Signal Processing*, vol. 67, no. 8, 2019, pp. 1992–2006.

[16] W. Dai and S. Yuksel, "Observability of a Linear System Under Sparsity Constraints," in *IEEE Transactions on Automatic Control*, vol. 58, no. 9, 2013, pp. 2372–2376.

[17] A. Charles, A. Balavoine, and C. Rozell, "Dynamic Filtering of Time-Varying Sparse Signals via l1 Minimization," in *IEEE Transactions on Signal Processing*, vol. 64, no. 21, 2016, pp. 5644–5656.

[18] D. Kahl, A. Weber, and M. Kschischo, "Localization of Invariable Sparse Errors in Dynamic Systems," in *IEEE Transactions on Control of Network Systems*, vol. 8, no. 4, 2021, pp. 1649–1658.

[19] M. S. Asif, A. Charles, J. Romberg, and C. Rozell, "Estimation and dynamic updating of time-varying signals with sparse variations," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 3908–3911.

[20] S. Sefati, N. J. Cowan, and R. Vidal, "Linear systems with sparse inputs: Observability and input recovery," in *2015 American Control Conference (ACC)*. IEEE, pp. 5251–5257.

[21] M. Kafashan, A. Nandi, and S. Ching, "Relating observability and compressed sensing of time-varying signals in recurrent linear networks," in *Neural Networks*, vol. 83, 2016, pp. 11–20.

[22] G. Joseph and C. R. Murthy, "A Noniterative Online Bayesian Algorithm for the Recovery of Temporally Correlated Sparse Vectors," in *IEEE Transactions on Signal Processing*, vol. 65, no. 20, 2017, pp. 5510–5525.

[23] A. Ansari and D. S. Bernstein, "Deadbeat Input Reconstruction and State Estimation for Discrete-Time Linear Time-Varying Systems," in *2019 American Control Conference (ACC)*, pp. 5123–5127.

[24] M. Cho, K. Vijay Mishra, and W. Xu, "Computable performance guarantees for compressed sensing matrices," in *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, 2018, p. 16.