SIAM/ASA J. UNCERTAINTY QUANTIFICATION
Vol. 5, pp. 572–597

Robustness of Solutions of the Inverse Problem for Linear Dynamical Systems with Uncertain Data*

S. Stanhope†, J. E. Rubin‡, and D. Swigon‡

Abstract. The problem of estimation of parameters of a dynamical system from discrete data can be formulated as the problem of inverting the map from the parameters of the system to points along a corresponding trajectory. In this work, we focus on linear systems and derive necessary and sufficient conditions for single trajectory data to yield a matrix of parameters with specific dynamical properties. To address the key issue of robustness, we establish conditions that ensure that the desired properties of the solution to the inverse problem are maintained on an open neighborhood of the given data. We then build from these results to find bounds on the uncertainty in the data that can be tolerated without a change in the nature of the inverse problem. In particular, both analytical and numerical estimates are derived for the largest allowable uncertainty in the data for which the qualitative features of the inverse problem solution, such as uniqueness or attractor properties, persist. We also derive the conditions and bounds for the nonexistence of a real parameter matrix corresponding to the given data, which can be utilized in modeling practice to prescribe a level of uncertainty under which the linear model can be rejected as a representation of the data.

Key words. parameter estimation, identifiability, linear systems, inverse problem, uncertain data, matrix logarithm

AMS subject classifications. 15A16, 15A45, 15A60, 34A30, 34A55, 34D30

DOI. 10.1137/16M1062466

1. Introduction.

A fundamental problem in modeling temporally evolving systems is the determination of model parameter values from experimental observations collected at specific time points. Since models can be viewed as forward mappings from sets of parameter values to time-dependent states of model variables, the problem of parameter estimation is often referred to as an inverse problem. Although parameter estimation has received significant attention in the literature, certain fundamental questions about the inverse problem still remain open. Solving the inverse problem becomes even more challenging in the presence of uncertainty in experimental measurements, as may arise due to measurement errors and fluctuations in system components. The overall goal of this work is to derive estimates of the degree of uncertainty in data to which properties of the inverse problem, such as existence

*Received by the editors February 22, 2016; accepted for publication (in revised form) February 8, 2017; published electronically June 27, 2017.

http://www.siam.org/journals/juq/5/M106246.html

Funding: This work was partially supported by National Science Foundation award EMSW21-RTG 0739261. The second author's work was partially supported by NSF awards DMS-1312508 and DMS-1612193. The third author's work was partially supported by National Institutes of Health award GM105728-01A1.

†Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260. Current address: Department of Mathematics, Temple University, Philadelphia, PA 19122 (stanhope@temple.edu).

‡Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260 (jonrubin@pitt.edu, swigon@pitt.edu).

and uniqueness of solutions, are robust.

We start by addressing fundamental issues of existence and uniqueness of solutions to the inverse problem based on a discrete collection of linearly independent data points assumed to be known without uncertainty, before turning to the uncertain case. We focus our analysis on linear models, which are prevalent in the study of many important applications including pharmacokinetics, gene regulation, linear response theory for mechanical and electronic systems, continuous time Markov chain probabilistic models, and near-equilibrium responses of nonlinear systems [5, 12, 9, 2, 16]. In addition to their applicability, linear systems are convenient because in the linear case, there is an explicit structure of the associated forward solution map that can be exploited. Furthermore, we mostly consider data points that are equally spaced in time, as may be obtained from experiments with regimented data collection schedules and for which it may be possible to explicitly solve for the linear system parameter matrix. Despite these advantages, the inverse problem is nontrivial because the solution to a linear dynamical system depends nonlinearly on its parameters.

Specifically, in section 2, we begin by considering data with no uncertainty, where classical results on matrix logarithms yield necessary and sufficient conditions for the data to specify a unique parameter matrix A that solves the inverse problem. Contrary to our expectations, we find that the existence of model parameters corresponding to given data is guaranteed only for a restricted subset of potential data sets and that there is only a limited region in data space that yields a unique set of parameters. Subsequently, we explore how uncertainty in the data impacts the existence and uniqueness of A. In section 3, we provide conditions that ensure that existence or uniqueness of the inverse problem solution is guaranteed to hold in an open neighborhood of data. These steps prepare us for section 4, where we present the main results of the paper, consisting of analytical and numerical estimates of the maximum uncertainty in the data under which the properties of the inverse problem are certain to be preserved. Examples for two-dimensional systems are shown in section 5, where we first define regions in data space for which the solution to the inverse problem has various properties and then illustrate bounds on the maximal permissible uncertainty for those properties. In section 6, we explore an example of a larger-dimensional system by applying our results to a model of gene regulation. In section 7, we briefly remark on the case of nonequally spaced data points, and we address the cases of differing amounts of data in section 8. Finally, we conclude with a discussion in section 9, which includes some comments on open directions and related work in the past literature.

2. Definitions and preliminaries. As in [17], we consider a model defined as a finite-dimensional linear dynamical system, which we denote as

$$\begin{aligned} \dot{x}(t) &= A\,x(t), \\ x(0) &= b. \end{aligned} \tag{2.1}$$

In (2.1), $x(t) \in \mathbb{R}^n$ is the state of the system at time t, the system parameters are the entries of the coefficient matrix $A \in \mathbb{R}^{n\times n}$, and $b \in \mathbb{R}^n$ is the initial condition. For clarity of exposition we will refer to the entire matrix A as the parameter A. We shall define the *parameter space* $\mathcal{P}$ as the set of all parameters A and initial conditions b.

For a fixed A, system (2.1) has a well-defined solution, or trajectory, given by $x(t; A, b) =$

$\Phi(t)b$, where $\Phi(t) = e^{At}$ is the principal matrix solution. The data representing the system are a set of observations of the trajectory values. We assume that data for all of the state variables are available, and we denote by $\mathcal{D}$ the *data space* consisting of a set of $(n+1)$-tuples $d = (x_0, x_1, x_2, \ldots, x_n)$ of points $x_j \in \mathbb{R}^n$; i.e., each such d will be referred to as a *data set*, and each x_j as a *data point*. Sampling the trajectory $x(t; A, b)$ at equally spaced times (without loss of generality $\Delta t = 1$) yields an element of $\mathcal{D}$, namely a data set composed of the specific data points $x_j = x(j;\, A, b) \in \mathbb{R}^n,\ \ j = 0, 1, \ldots, n$; nonuniformly spaced data are discussed in section 7. We use *solution map* to refer to the map $F : \mathcal{P} \to \mathcal{D}$ from parameter space to data space, where $\mathcal{P}$ and $\mathcal{D} \subseteq \mathbb{R}^{n\times(n+1)}$, defined as $F(A, b) = (x_0, x_1, x_2, \ldots, x_n)$ for this choice of $\{x_j\}$ sampled from $x(t; A, b)$. The *inverse problem* is then the problem of inverting the map F to find $F^{-1}(d)$ (i.e., to find A, b such that $F(A, b) = d$) for a given data set d. From the definition of x_j, we find that $b = x_0$, and thus, we focus on the problem of determining the parameter A. If the data set d is obtained from experimental measurements or some other outside source, then this problem may or may not have a solution.

We set out to derive necessary and sufficient conditions that a data set $d \in \mathcal{D}$ must satisfy so that there exists a unique real matrix A for which the dynamical system (2.1) produces data d. These conditions define a subset of the data space on which the inverse map F^{-1} is well defined. Given a uniformly spaced data set $d \in \mathcal{D}$, one can attempt to solve the inverse problem as follows: Denote by X_0 and X_1 the $n \times n$ matrices $[x_0 \,|\, \ldots \,|\, x_{n-1}]$ and $[x_1 \,|\, \ldots \,|\, x_n]$, respectively. The principal matrix solution provides a relation between the data points; letting $\Phi := \Phi(1) = e^A$, we have $x_{j+1} = \Phi x_j$, which implies that $X_1 = \Phi X_0$ and hence $\Phi = X_1 X_0^{-1}$. All that remains is to find A as the matrix logarithm of Φ. Thus, from an operational standpoint, the map F^{-1} is a composition of two nonlinear maps: (i) the map $G : \mathcal{D} \to \mathbb{R}^{n\times n}$ defined by $G(d) = \Phi$, which is defined (and continuous) at all points d such that $x_0, \ldots, x_{n-1}$ are linearly independent (i.e., wherever X_0^{-1} exists), and (ii) the matrix logarithm map, denoted here as $L : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ and defined as $A = L(e^A)$. Hence, $F^{-1} = [L \circ G, I]$, where $L \circ G$ operates on d to produce A and the identity I operates on x_0 to produce b. In view of the above, F^{-1} is well defined if (i) X_0 is invertible, (ii) the matrix logarithm of Φ exists, and (iii) the matrix logarithm of Φ is unique. The case when condition (i) fails was studied extensively in our earlier work [17]; in that case the system (2.1) generating the data d either does not exist or is not identifiable. Conditions (ii) and (iii) can be addressed with the help of two theorems by Culver [8], which characterize the existence and uniqueness of a real matrix logarithm.

Theorem 2.1 (Culver). *Let Φ be a real square matrix. Then there exists a real solution A to the equation $\Phi = e^A$ if and only if Φ is nonsingular and each Jordan block of Φ belonging to a negative real eigenvalue occurs an even number of times.*

Theorem 2.2 (Culver). *Let Φ be a real square matrix. Then there exists a unique real solution A to the equation $\Phi = e^A$ if and only if all of the eigenvalues of Φ are positive real and no Jordan block of Φ appears more than once.*

We note that even when a matrix logarithm exists, there are still issues with how to compute it. Numerical methods for computing the logarithm of a matrix are discussed in [10, 6, 1].

3. Existence and uniqueness of the inverse in n dimensions. As will become clear shortly, Theorems 2.1 and 2.2 identify matrices Φ that are not robust in the sense that they form a set of zero measure. Since we aim to discuss the properties of inverse problem solutions for uncertain data, such as data that are perturbed due to noise, it makes sense to determine conditions that ensure that a given matrix Φ is inside an open set of matrices with particular existence or uniqueness properties.

3.1. Inverse problems on open sets. We now state and prove three corollaries of Theorems 2.1 and 2.2 that are useful for considering uncertain data and, as it turns out, avoid the conditions on Jordan blocks that can become overly cumbersome for practical use in n dimensions. The first corollary characterizes open sets of matrices that have real logarithms, the second corollary characterizes open sets of matrices that have unique real logarithms, and the third corollary characterizes open sets of matrices that do not have real logarithms.

Corollary 3.1 (to Theorem 2.1). *Let Φ^* be an $n \times n$ real matrix. The following statements are equivalent:*

(a) *There exists an open set $U \subset \mathbb{R}^{n\times n}$ containing Φ^* such that for any $\Phi \in U$ the equation $\Phi = e^A$ has an $n \times n$ real solution A.*

(b) *Φ^* has only positive real or complex eigenvalues.*

Proof. Suppose that $\Phi^* \in \mathbb{R}^{n\times n}$ has only positive real or complex eigenvalues. Then condition (a) follows immediately by Theorem 2.1 and the continuous dependence of eigenvalues on matrix entries.

For the converse, suppose that $\Phi^* \in \mathbb{R}^{n\times n}$ has a real matrix logarithm, i.e., $\Phi^* = e^A$, where A is $n \times n$ real matrix. By Theorem 2.1, either Φ^* satisfies (b), or Φ^* has at least one negative eigenvalue and the corresponding Jordan block occurs an even number of times. We now show that the second alternative contradicts the existence of the open set U. To this end, let $\Phi^* = QJQ^{-1}$, where

$$J = \left[\begin{array}{c|c} J_1 & 0 \\ \hline 0 & \begin{array}{cc} J_2 & \\ & \ddots \end{array} \end{array}\right] \tag{3.1}$$

is a Jordan canonical form of Φ^*, with J_1 a Jordan block corresponding to a negative eigenvalue. Let $B = QKQ^{-1}$, where

$$K = \left[\begin{array}{c|c} aI & 0 \\ \hline 0 & 0 \end{array}\right],$$

with $a \in \mathbb{R}$ and I the identity matrix of the same size as J_1. Then for every sufficiently small nonzero a, $\Phi^* + B$ has a negative eigenvalue for which the corresponding Jordan block occurs exactly once and hence there is no real A such that $\Phi^* + B = e^A$. ∎

Corollary 3.2 (to Theorem 2.2). *Let Φ^* be an $n \times n$ real matrix. The following statements are equivalent:*

(a) *There exists an open set $U \subset \mathbb{R}^{n\times n}$ containing Φ^* such that for any Φ in U the equation $\Phi = e^A$ has a unique $n \times n$ real solution A.*

(b) Φ^* *has* n *distinct positive real eigenvalues.*

Proof. The proof is similar to the previous one. Condition (b) implies condition (a) based on Theorem 2.2 and continuity of eigenvalues. For the converse, suppose that $\Phi^* \in \mathbb{R}^{n\times n}$ has a unique real matrix logarithm, i.e., there is a unique $A \in \mathbb{R}^{n\times n}$ such that $\Phi^* = e^A$, and Φ^* is in the interior of an open set of such matrices. By Theorem 2.2, Φ^* has positive real eigenvalues, and no Jordan block appears more than once. There can be more than one Jordan block for the same eigenvalue, but those Jordan blocks must be of different sizes. We can write $\Phi^* = QJQ^{-1}$, where J is the Jordan canonical form of Φ^* as defined in (3.1) and J_i are ordered in size from largest to smallest such that no two J_i are the same.

Suppose now that the largest Jordan block, J_1, is not a 1×1 matrix. Let $p_J(t)$ be the characteristic polynomial of J, i.e., $p_J(t) = \prod_{i=1}^{k}(\lambda_i - t)^{d_i}$, where λ_i are all real positive, k is the number of Jordan blocks, d_i is the dimension of block i, and $d_1 \geqslant 2$. Let $B = QKQ^{-1}$, where K is an $n \times n$ matrix with all zero entries except $K_{12} = a$, $K_{21} = -a$. Then, for every $a > 0$, $\hat{J} = J + K$ has entries $\hat{J}_{12} = 1 + a$, $\hat{J}_{21} = -a$, which implies that $\hat{J}$ has characteristic polynomial

$$p_{\hat{J}}(t) = \left[(\lambda_1 - t)^2 + a(1+a)\right](\lambda_1 - t)^{d_1-2}\prod_{i=2}^{k}(\lambda_i - t)^{d_i}$$

with two complex roots. It follows that $\Phi^* + B = e^A$ does not have a unique real solution A for $a > 0$, which contradicts the assumption that A is unique for each Φ in an open set containing Φ^*. Thus, all Jordan blocks J_i are of size 1, and furthermore, since no two J_i can be the same, Φ^* has n distinct eigenvalues. ∎

Statements similar to Corollaries 3.1 and 3.2 can be made to establish the existence of open sets of matrices with other properties. An example follows.

Corollary 3.3 (to Theorem 2.1). *Let* Φ^* *be an* $n \times n$ *real matrix. The following statements are equivalent:*

(a) *There exists an open set* $U \subset \mathbb{R}^{n\times n}$ *containing* Φ^* *such that for each* $\Phi \in U$ *the equation* $\Phi = e^A$ *does not have an* $n \times n$ *real solution* A.

(b) Φ^* *has at least one negative eigenvalue of odd multiplicity.*

Proof. Suppose that $\Phi^* \in \mathbb{R}^{n\times n}$ has at least one negative eigenvalue of odd multiplicity. Then, there is at least one Jordan block associated to it that occurs an odd number of times, and hence, by Theorem 2.1, there is no $A \in \mathbb{R}^{n\times n}$ such that $\Phi^* = e^A$. Moreover, there exists an open neighborhood of Φ^* for which there remains at least one negative eigenvalue of odd multiplicity.[1]

For the converse, suppose that there exists an open set $U \subset \mathbb{R}^{n\times n}$ containing Φ^* such that for each $\Phi \in U$ the equation $\Phi = e^A$ does not have an $n \times n$ real solution A. Since $\Phi^* = e^A$ has no real solution, by Theorem 2.1, Φ^* is singular, or there is a negative eigenvalue of Φ^* which belongs to a Jordan block that appears an odd number of times, or both are true.

[1] A perturbation may split any negative eigenvalue of odd multiplicity into a collection of distinct negative eigenvalues and/or complex pairs, with the sum of all multiplicities equal to that of the original eigenvalue. At least one multiplicity in that sum must be odd for the result to be odd.

If Φ^* is singular with no Jordan block associated to a negative eigenvalue occurring an odd number of times, then in every neighborhood of Φ^* there exists a Φ which is nonsingular with the same condition on Jordan blocks, and $\Phi = e^A$ will have a real solution A. This contradicts the existence of U.

Thus, there is at least one negative eigenvalue of Φ^* for which a Jordan block repeats an odd number of times. Denote these negative eigenvalues by $\lambda_1, \ldots, \lambda_r$, $r \geqslant 1$. Let m_{λ_i} be the multiplicity of λ_i.

Suppose that all m_{λ_i} are even. Let $\Phi^* = QJQ^{-1}$, where J is the Jordan canonical form of Φ^* as defined in (3.1), let $J_1^{\lambda_1}, J_2^{\lambda_1}, \ldots, J_k^{\lambda_1}$ denote all of the Jordan blocks associated to λ_1 (the blocks may have the same size), and denote by J_1 the block-diagonal matrix composed of $J_1^{\lambda_1}, J_2^{\lambda_1}, \ldots, J_k^{\lambda_1}$.

For $i = 1, \ldots, k$ define $W_i^{\lambda_1}$ to be a block-diagonal matrix of the same dimension as $J_i^{\lambda_1}$ composed of repeated blocks K, where

$$K = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$$

for some $a > 0$. If the dimension of $J_i^{\lambda_1}$ is even and equal to $2s$, then $W_i^{\lambda_1}$ is composed of exactly s blocks K, the characteristic polynomial of $J_i^{\lambda_1} + W_i^{\lambda_1}$ is $p(t) = [(\lambda_1 - t)^2 + a(1+a)]^s$, and since $a > 0$, $J_i^{\lambda_1} + W_i^{\lambda_1}$ has only complex eigenvalues.

If the dimension of $J_i^{\lambda_1}$ is odd and equal to $2s+1$, then $W_i^{\lambda_1}$ contains s blocks K and a zero block of size one, the characteristic polynomial of $J_i^{\lambda_1} + W_i^{\lambda_1}$ is $p(t) = [(\lambda_1 - t)^2 + a(1+a)]^s(\lambda_1 - t)$, and since $a > 0$, $J_i^{\lambda_1} + W_i^{\lambda_1}$ has s complex conjugate pairs of eigenvalues and one real negative eigenvalue λ_1.

Define the $m_{\lambda_1} \times m_{\lambda_1}$ matrix,

$$C_1 = \begin{bmatrix} W_1^{\lambda_1} & & \\ & \ddots & \\ & & W_k^{\lambda_1} \end{bmatrix}.$$

Since m_{λ_1} is even, the number of odd sized blocks is even; therefore for every $a > 0$, the matrix $J_1 + C_1$ has an even number of size one Jordan blocks corresponding to eigenvalue λ_1, and all other eigenvalues of $J_1 + C_1$ are complex.

By repeating this process for $\lambda_2, \ldots, \lambda_r$, we can construct the matrix

$$F = \left[\begin{array}{ccc|c} C_1 & & & \\ & \ddots & & \\ & & C_r & \\ \hline & & & 0 \end{array}\right].$$

Let $B = QFQ^{-1}$. Then it follows that for every $a > 0$, all negative eigenvalues of $\Phi^* + B$ will have Jordan blocks that repeat an even number of times. If Φ^* is nonsingular, then $\Phi^* + B = e^A$ has real solution A for every $a > 0$. If Φ^* is singular, then $\Phi^* + B + aI = e^A$ has real solution A for every $a > 0$ sufficiently small. In either case, this implication contradicts

the existence of the open set U. Thus it must be that some m_{λ_i} is odd, and therefore, Φ^* has at least one negative eigenvalue of odd multiplicity. ∎

For a given d such that the associated matrix Φ has n distinct positive eigenvalues, Corollary 3.1 (3.2, or 3.3, respectively) gives an open set $U \subset \mathbb{R}^{n\times n}$ on which the matrix logarithm exists (exists and is unique, or does not exist, respectively). By the continuity of G at d, $G^{-1}(U)$ (the preimage of U) is an open set in $\mathcal{D}$ containing d such that every data set in $G^{-1}(U)$ is generated by a real A (a unique real A, or no real A, respectively). Consequently, we can now summarize the existence and uniqueness results on open sets.

Theorem 3.4. *Suppose that the data $d = (x_0, x_1, x_2)$ are such that X_0 is invertible and $\Phi = X_1 X_0^{-1}$ has*

(a) *only positive real or complex eigenvalues,*
(b) *distinct positive real eigenvalues, or*
(c) *at least one negative eigenvalue of odd multiplicity.*

Then there is an open set $U \subset \mathcal{D}$ containing d such that for every $d^* \in U$, the problem $F(A, b) = d^*$ has, respectively,

(a) *a real solution,*
(b) *a unique real solution, or*
(c) *no real solution.*

4. Analysis of uncertainty in the determination and characterization of the inverse.

Realistic data are never exact but are subject to uncertainty caused by measurement error, fluctuation in experimental conditions, or variability in experimental subjects. A natural question arises as to how large an uncertainty in the data can be tolerated without altering the properties of the solution to the inverse problem. Several scenarios are of interest, such as the following:

- The data imply that the inverse system has a stable node. What is the largest uncertainty in the data that ensures the maintained stability of the equilibrium for the inverse system? What is the largest uncertainty that maintains the node property?
- The data imply that the inverse system has oscillations (damped or undamped). What is the largest uncertainty that maintains the oscillatory property of the system?
- The inverse does not exist for given data. What is the largest uncertainty for which we can still rule out the linear model?

The theory developed in section 3 implies that all of the above criteria can be formulated as conditions on the eigenvalues of the perturbed fundamental matrix. For example, a system with a stable node still has a stable node under perturbation if and only if the eigenvalues of the perturbed fundamental matrix remain real, positive, distinct, and smaller than 1. Thus, in principle, we could construct direct algebraic constraints on the data similar to those given in the supplemental material (M106246_supplement.pdf [local/web 3.70MB]) to define regions of the data space in which data correspond to systems with specific dynamical properties; however, these quickly become prohibitively complex as the problem dimension increases. We will therefore focus instead on the scenario where a specific data set d is given, and derive two types of bounds on the maximal perturbation of d for which a particular property of the system is conserved: lower bounds $\underline{\epsilon}_U$, such that a property associated with d is guaranteed to

hold for all $\hat{d}$ with $\|\hat{d}-d\| < \underline{\epsilon}_{\mathrm{U}}$, and upper bounds $\overline{\epsilon}_{\mathrm{U}}$, such that there is guaranteed to exist a $\hat{d}$ with $\|\hat{d}-d\| = \overline{\epsilon}_{\mathrm{U}}$ such that the property does not hold at $\hat{d}$.

Note that any affine transformation of the data preserves the eigenvalue structure and hence the existence, uniqueness, and stability of the system with respect to the inverse problem. Thus, the data can be varied in a coordinated fashion to an arbitrary extent without affecting qualitative properties of the inverse. Here, however, we focus on finding limits on uncorrelated perturbations of the data. Let

$$C(d,\epsilon) = \bigtimes_{i=0}^{n} c(x_i,\epsilon),$$

where $c(z,\epsilon)$ is a hypercube in $\mathbb{R}^n$ with center $z \in \mathbb{R}^n$ and side length 2ϵ, i.e., where $c(z,\epsilon) = \{\tilde{z} \in \mathbb{R}^n | \max_{1\leqslant j\leqslant n} |\tilde{z}_j - z_j| < \epsilon\}$ for z_i denoting the components of the vector z. The definition of the neighborhood $C(d,\epsilon)$ is chosen so that the parameter ϵ controls the maximum perturbation $\Delta x_0, \Delta x_1, \ldots, \Delta x_n$ in any component of the data $x_0, x_1, \ldots, x_n$; i.e., $\tilde{d} \in C(d,\epsilon)$ if and only if $\max_{i,j} |(\Delta x_i)_j| < \epsilon$, where $|(\Delta x_i)_j| = (\tilde{x}_i)_j - (x_i)_j$. Neighborhood $C(d,\epsilon)$ of the data set $d \in \mathcal{D}$ will be called *permissible* for some qualitative property of the inverse of d (such as existence, uniqueness, stability, and so on) if and only if that qualitative property is shared by inverses of all data sets $\tilde{d} \in C(d,\epsilon)$. The value $\epsilon > 0$ is called the *maximal permissible uncertainty* for some qualitative property of the inverse of d if and only if $C(d,\epsilon)$ is a permissible neighborhood of d for that property and $C(d,\tilde{\epsilon})$ is not a permissible neighborhood for that property for all $\tilde{\epsilon} > \epsilon$.

We begin with an analytical and numerical description of the maximal permissible uncertainty for existence and uniqueness of the inverse $F^{-1}(d)$ of d. The extension to other properties will be described in section 4.4. We derive both lower and upper analytical bounds on maximal permissible uncertainty, describe a numerical procedure for computing the bounds, and then compare the estimates with direct numerical results for several examples.

4.1. Analytical lower bound. Consider a fixed data set $d = (x_0, x_1, x_2, \ldots, x_n) \in \mathcal{D}$ such that the associated matrix $\Phi = X_1 X_0^{-1} = [x_1 \,|\, \ldots \,|\, x_n][x_0 \,|\, \ldots \,|\, x_{n-1}]^{-1}$ has n distinct positive eigenvalues. Thus, $[A,b] = F^{-1}(d)$ is unique and there is a neighborhood of d for which uniqueness persists. For any perturbed data set $\tilde{d} = (\tilde{x}_0, \tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n) \in \mathcal{D}$ with $\tilde{x}_i = x_i + \Delta x_i$, define the perturbed data matrices $\tilde{X}_0 = X_0 + \Delta X_0 = [x_0 \,|\, \ldots \,|\, x_{n-1}] + [\Delta x_0 \,|\, \ldots \,|\, \Delta x_{n-1}]$ and $\tilde{X}_1 = X_1 + \Delta X_1$ (analogously). Let $\tilde{\Phi} = \tilde{X}_1 \tilde{X}_0^{-1}$ be the fundamental matrix of the perturbed data.

Let ϵ_{U} be the maximal permissible uncertainty in the data d to ensure the existence of a unique inverse. By definition, for any perturbation of the data with $\max_{i,j} |(\Delta x_i)_j| < \epsilon_{\mathrm{U}}$ the matrix $\tilde{\Phi}$ has a unique logarithm A, and for any $\hat{\epsilon} > \epsilon_{\mathrm{U}}$ there exists a perturbation of the data with $\epsilon_{\mathrm{U}} < \max_{i,j} |(\Delta x_i)_j| < \hat{\epsilon}$ such that $\tilde{\Phi}$ does not have a unique logarithm.

A lower bound $\underline{\epsilon}_{\mathrm{U}}$ on ϵ_{U} can be obtained by the following result, where $\|\cdot\|$ denotes a matrix norm that is either the maximum row sum norm $\|\cdot\|_\infty$ or the maximum column sum norm $\|\cdot\|_1$, defined as

$$\|A\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|, \qquad \|A\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

Theorem 4.1. *Let $d \in \mathcal{D}$ be such that Φ has n distinct positive eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $m_1 = \frac{1}{2}\min_{i\neq j}|\lambda_i - \lambda_j|$, $m_2 = \min_{1\leqslant i\leqslant n}\{\lambda_i\} > 0$, and $\delta_{\mathrm{U}} = \min\{m_1, m_2\}$. If $\epsilon > 0$ is such that $\epsilon \leqslant \underline{\epsilon}_{\mathrm{U}} := f(\delta_{\mathrm{U}}, d)$, where*

$$f(\delta, d) = \frac{\delta}{n\,(\delta + 1 + \|\Lambda\|)\,\|S^{-1}\|\|X_0^{-1}S\|}, \tag{4.1}$$

$\Phi = S\Lambda S^{-1}$, and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, then for any $\tilde{d} \in C(d, \epsilon)$, $\tilde{\Phi}$ has n distinct positive eigenvalues, and hence the equation $e^{\tilde{A}} = \tilde{\Phi}$ has a unique solution $\tilde{A}$.

The proof of Theorem 4.1 utilizes several preliminary results that we now present. The first result and its proof make use of Theorems 6.1.1 (Gershgorin disc theorem) and 6.3.2 in [13] and the proofs presented therein.

Lemma 4.2. *Let $\Phi \in \mathbb{R}^{n\times n}$ be diagonalizable with $\Phi = S\Lambda S^{-1}$ and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Let $E \in \mathbb{R}^{n\times n}$. If $\tilde{\lambda}$ is an eigenvalue of $\Phi + E$, then $\tilde{\lambda} \in D$, where*

$$D = \bigcup_{i=1}^{n} D_i, \qquad D_i = \{z \in \mathbb{C} : |z - \lambda_i| \leqslant \|S^{-1}ES\|\}.$$

Furthermore, if λ_i are all distinct and the discs D_i are pairwise disjoint, then each D_i contains exactly one eigenvalue of $\Phi + E$.

Proof. By similarity, $\Phi + E$ has the same eigenvalues as $\Lambda + S^{-1}ES$. Denote by $\tilde{e}_{ij}$ the elements of $S^{-1}ES$. Then, by the Gershgorin disc theorem, the eigenvalues of $\Lambda + S^{-1}ES$ are contained in the union of the discs

$$Q_i = \left\{ z \in \mathbb{C} : |z - (\lambda_i + \tilde{e}_{ii})| \leqslant \sum_{\substack{j=1 \\ j\neq i}}^{n} |\tilde{e}_{ij}| \right\}.$$

Clearly, each disc Q_i is contained in the disc

$$P_i = \left\{ z \in \mathbb{C} : |z - \lambda_i| \leqslant \sum_{j=1}^{n} |\tilde{e}_{ij}| \right\}.$$

Furthermore, in view of

$$\sum_{j=1}^{n} |\tilde{e}_{ij}| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |\tilde{e}_{ij}| = \|S^{-1}ES\|_\infty,$$

each disc P_i is contained in the disc

$$D_i = \{z \in \mathbb{C} : |z - \lambda_i| \leqslant \|S^{-1}ES\|_\infty\}.$$

Thus, if $\tilde{\lambda}$ is an eigenvalue of $\Phi + E$, then $\tilde{\lambda} \in Q_i \subseteq P_i \subseteq D_i$ for some i, and therefore $\tilde{\lambda} \in D$. The argument for the norm $\|\cdot\|_1$ is constructed in a similar fashion by replacing row sums with column sums in the relations above.

If λ_i are all distinct and the sets D_i are pairwise disjoint, then the discs Q_i are pairwise disjoint, and the Gershgorin disc theorem implies that there is exactly one eigenvalue of $\Phi + E$ in each Q_i and hence each D_i. ∎

Lemma 4.3. *Suppose the eigenvalues $\lambda_1, \dots, \lambda_n$ of Φ are real, positive, and distinct. Let $\Phi = S\Lambda S^{-1}$ with $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, and let $m_1 = \frac{1}{2}\min_{i \neq j}|\lambda_i - \lambda_j|$, $m_2 = \min_{1 \leqslant i \leqslant n}\{\lambda_i\}$, and $\delta = \min\{m_1, m_2\}$. If $\|S^{-1}ES\| < \delta$, then the eigenvalues of $\Phi + E$ are real, positive, and distinct.*

Proof. Suppose the eigenvalues $\lambda_1, \dots, \lambda_n$ of Φ are real, positive, and distinct. Then Φ is diagonalizable as $\Phi = S\Lambda S^{-1}$ with $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Let R_i be the disc

$$R_i = \{z \in \mathbb{C} : |z - \lambda_i| \leqslant m_1\},$$

and let D_i be defined as in the statement of Lemma 4.2. Since $\|S^{-1}ES\| < \delta \leqslant m_1$, it follows that $D_i \subseteq R_i$. The sets R_i are pairwise disjoint by the definition of m_1, so the sets D_i are pairwise disjoint, and by Lemma 4.2, each D_i contains exactly one eigenvalue of $\Phi + E$. The center of D_i is $\lambda_i \in \mathbb{R}$, so if D_i were to contain a complex eigenvalue of $\Phi + E$, it would also contain its conjugate, which is a contradiction. Thus, the eigenvalues of $\Phi + E$ are real and distinct. Furthermore, the inequality $\|S^{-1}ES\| < \delta \leqslant m_2$ implies that $D_i \subseteq \{z \in \mathbb{C} : |z - \lambda_i| \leqslant m_2\} \subseteq \{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}$, and hence the eigenvalues of $\Phi + E$ are all positive. ∎

Thus, to guarantee that the eigenvalues $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$ of $\tilde{\Phi} = \Phi + E$ are real, positive, and distinct, it suffices to choose the perturbation matrix E such that $\|S^{-1}ES\| < \delta$. Theorem 4.1 provides a lower bound on the largest allowable perturbation of the data points such that this condition holds.

Proof of Theorem 4.1. Let $\tilde{d} \in C(d, \epsilon)$, and let $\tilde{\Phi}$ be the associated fundamental matrix as previously defined. Let $E = \tilde{X}_1\tilde{X}_0^{-1} - X_1X_0^{-1}$. Applying the definitions of $\tilde{X}_0, \tilde{X}_1$, and the relation $X_1 = \Phi X_0$ and multiplying on the right by $(X_0 + \Delta X_0)$ yields

$$E(X_0 + \Delta X_0) = (X_1 + \Delta X_1) - \Phi(X_0 + \Delta X_0).$$

Next, we distribute the E on the left-hand side, rearrange, multiply both sides by X_0^{-1}, and use the definition of Φ to obtain

$$E = (\Delta X_1 - \Phi\Delta X_0)X_0^{-1} - E\Delta X_0X_0^{-1}$$

and hence

$$S^{-1}ES = S^{-1}(\Delta X_1 - \Phi\Delta X_0)X_0^{-1}S - S^{-1}ESS^{-1}\Delta X_0X_0^{-1}S.$$

By taking the norm of both sides of the last equation, using the matrix norm property on the right-hand side, and reordering the terms, we can find the following bound on $\|S^{-1}ES\|$, provided that $\|S^{-1}\Delta X_0X_0^{-1}S\| < 1$:

$$\|S^{-1}ES\| \leqslant \frac{\|S^{-1}(\Delta X_1 - \Phi\Delta X_0)X_0^{-1}S\|}{1 - \|S^{-1}\Delta X_0X_0^{-1}S\|}.$$

After introducing $\|S^{-1}\Delta X\| := \max\{\|S^{-1}\Delta X_0\|, \|S^{-1}\Delta X_1\|\}$ and using the decomposition $\Phi = S\Lambda S^{-1}$, we obtain the following bound on $\|S^{-1}ES\|$ in terms of $\|S^{-1}\Delta X\|$ (provided that $\|S^{-1}\Delta X\|\|X_0^{-1}S\| < 1$):

$$\|S^{-1}ES\| \leqslant \frac{\|S^{-1}\Delta X\|(1 + \|\Lambda\|)\|X_0^{-1}S\|}{1 - \|S^{-1}\Delta X\|\|X_0^{-1}S\|}. \tag{4.2}$$

Now, let

$$q = \frac{\delta_{\mathrm{U}}}{(1 + \|\Lambda\|)} + 1, \qquad \hat{\epsilon} = n\|S^{-1}\|\epsilon,$$

and let ϵ be defined as in the statement of the theorem. It follows from $\epsilon \leqslant \underline{\epsilon}_{\mathrm{U}}$ that $\hat{\epsilon} \leqslant \frac{q-1}{q\|X_0^{-1}S\|}$ or, equivalently, $\frac{1}{q} \leqslant 1 - \hat{\epsilon}\|X_0^{-1}S\|$. The condition $\tilde{d} \in C(d, \epsilon)$ implies that $|(\Delta x_i)_j| < \epsilon$ for all $i \in \{1, \ldots, n\}$, $j \in \{0, \ldots, n\}$, and so $\|S^{-1}\Delta X\| \leqslant \|S^{-1}\|\|\Delta X\| \leqslant \hat{\epsilon}$ (which holds for both $\|\cdot\|_1$ and $\|\cdot\|_\infty$). Thus,

$$\frac{1}{q} < 1 - \|S^{-1}\Delta X\|\|X_0^{-1}S\|. \tag{4.3}$$

Substitution of (4.3) into inequality (4.2) can be used to conclude that $\|S^{-1}ES\| < \delta_{\mathrm{U}}$. By Lemma 4.3, $\tilde{\Phi}$ has n distinct positive eigenvalues. ∎

In the special case where the first n data points $x_0, \ldots, x_{n-1}$ are fixed, we can obtain a tighter bound on the size of Δx_n. We look for the largest uncertainty ϵ of the final data point x_n so that for any $\tilde{x}_n \in c(x_n, \epsilon)$, $\tilde{d} = (x_0, \ldots, x_{n-1}, \tilde{x}_n)$ gives an associated $\tilde{\Phi}$ with n distinct positive eigenvalues. Fixing the first n data points implies that $\Delta X_0 = 0$ and hence $E = \Delta X_1 X_0^{-1}$.

Theorem 4.4. *Let $d \in \mathcal{D}$ be such that Φ has n distinct positive eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $m_1 = \frac{1}{2}\min_{i \neq j}|\lambda_i - \lambda_j|$, $m_2 = \min_{1 \leqslant i \leqslant n}\{\lambda_i\} > 0$, and $\delta_{\mathrm{U}} = \min\{m_1, m_2\}$. If $\underline{\epsilon} > 0$ is such that*

$$\underline{\epsilon} < \max\left\{\frac{\delta_{\mathrm{U}}}{\|S^{-1}\|_\infty\|X_0^{-1}S\|_\infty}, \frac{\delta_{\mathrm{U}}}{n\|S^{-1}\|_1\|X_0^{-1}S\|_1}\right\},$$

then for any $\tilde{d} = (x_0, \ldots, x_{n-1}, \tilde{x}_n)$, where $\tilde{x}_n \in c(x_n, \underline{\epsilon})$, the associated matrix $\tilde{\Phi}$ has n distinct positive eigenvalues and hence the equation $e^{\tilde{A}} = \tilde{\Phi}$ has a unique solution $\tilde{A}$.

Proof. Given fixed $x_0, \ldots, x_{n-1}$ and $\tilde{x}_n \in c(x_n, \epsilon)$, we have $\|\Delta X_1\|_\infty = \|[0 \cdots 0\ \Delta x_n]\|_\infty = \max_{1 \leqslant j \leqslant n}|(\Delta x_n)_j| < \epsilon$ and $\|\Delta X_1\|_1 < n\epsilon$. If $\|X_0^{-1}S\|_\infty < n\|X_0^{-1}S\|_1$, then

$$\|S^{-1}ES\|_\infty \leqslant \|S^{-1}\Delta X_1\|_\infty\|X_0^{-1}S\|_\infty < \epsilon\|S^{-1}\|_\infty\|X_0^{-1}S\|_\infty < \delta_{\mathrm{U}}.$$

In the opposite case,

$$\|S^{-1}ES\|_1 \leqslant \|S^{-1}\Delta X_1\|_1\|X_0^{-1}S\|_1 < n\epsilon\|S^{-1}\|_1\|X_0^{-1}S\|_1 < \delta_{\mathrm{U}}.$$

In both cases, by Lemma 4.3, $\tilde{\Phi} = \Phi + E$ has n distinct positive eigenvalues. ∎

4.2. Analytical upper bound. To construct an upper bound on ϵ_{U}, we need only provide a technique for constructing a perturbation $\Delta x_0, \Delta x_1, \ldots, \Delta x_n$ for which the corresponding matrix $\tilde{\Phi}$ does not have n distinct real positive eigenvalues. Naively, for any specified diagonal form $\tilde{\Lambda}$ of $\tilde{\Phi}$, we can choose an arbitrary set of eigenvectors $\tilde{S}$, compute $\tilde{\Phi} = \tilde{S}\tilde{\Lambda}\tilde{S}^{-1}$, choose the first data point $\tilde{x}_0$, compute the remaining data points as $\tilde{x}_k = \tilde{\Phi}^k\tilde{x}_0$, and compute the error $\epsilon = \max_{i,j}|(\Delta x_j)_i|$. This procedure can be used to provide a valid upper bound on ϵ_{U}, but the resulting bound will be too large to be useful. Instead, we suggest the following approach.

Let $\hat{\Phi}$ be the matrix similar to Φ that has the form of a companion matrix, i.e.,

$$\hat{\Phi} = P\Phi P^{-1} = \begin{bmatrix} 0 & 0 & \dots & 0 & y_1 \\ 1 & 0 & \dots & 0 & y_2 \\ 0 & 1 & \dots & 0 & y_3 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & y_n \end{bmatrix}.$$

One can think of $\hat{\Phi}$ as the fundamental matrix for a trajectory for which the data set $\hat{d}$ is composed of the vectors of the standard basis of $\mathbb{R}^n$, together with the data vector $y = (y_1, y_2, \dots, y_n)^T$; i.e., $\hat{\Phi} = G(\hat{d})$ for $\hat{d} = (e_1, e_2, e_3, \dots, e_n, y)$. The matrix P defines the affine transformation that takes the data d into the standard (normalized) data $\hat{d}$, i.e., $P : \mathbb{R}^n \to \mathbb{R}^n$, where $Px_j = e_{j+1}$, $Px_n = y$. This implies that $P = X_0^{-1}$ and $\hat{\Phi} = X_0^{-1}X_1$.[2]

The vector y (the last column of $\hat{\Phi}$) is uniquely determined by the eigenvalues of Φ. Likewise, we can specify $\tilde{y}$ of the companion matrix $\hat{\tilde{\Phi}} = (\tilde{X}_0)^{-1}\tilde{\Phi}\tilde{X}_0$ by prescribing the eigenvalues of $\tilde{\Phi}$. The vector $\tilde{y}$ satisfies the relation

$$\tilde{y} = (\tilde{X}_0)^{-1}\tilde{x}_n = (X_0 + \Delta X_0)^{-1}(x_n + \Delta x_n),$$

which, using $x_n = X_0 y$, implies a formula for the perturbation of x_n in terms of $\tilde{y}$ and the perturbation of all other data points:

$$\Delta x_n = \Delta X_0 \tilde{y} + X_0(\tilde{y} - y). \tag{4.4}$$

This formula provides a linear constraint on the perturbation of the data in terms of the imposed eigenvalue properties (as represented by vector $\tilde{y}$) that does not require the knowledge of the eigenvector matrix S. Let $C(d, \tilde{\epsilon})$ be the smallest neighborhood of d that contains a data set $\tilde{d}$ corresponding to a companion matrix defined by $\tilde{y}$. In view of (4.4), the problem of finding $\tilde{\epsilon}$ can be reformulated as a linear programming problem of minimizing ϵ while satisfying the constraints

$$w_i = \sum_{j=0}^{n-1} (\Delta x_j)_i \tilde{y}_{j+1} - (\Delta x_n)_i, \qquad 1 \leqslant i \leqslant n, \tag{4.5}$$

$$-\epsilon \leqslant (\Delta x_j)_i \leqslant \epsilon, \qquad 1 \leqslant i \leqslant n,\ 0 \leqslant j \leqslant n, \tag{4.6}$$

where $w = X_0(y - \tilde{y})$. Note that the constraints (4.5) and inequalities (4.6) separate into n distinct problems for each individual i. For each i, the domain defined by inequalities (4.6) is a hypercube, and the constraint in (4.5) is a hyperplane in ΔX. The smallest ϵ_i corresponds to the situation in which the surface of the hypercube (but not its interior) and the hyperplane will have one or more points in common. One such point will occur at a vertex

[2]The supplemental material (M106246_supplement.pdf [local/web 3.70MB]) includes a discussion on how the companion matrix formulation can also be used to define a region of data on which the inverse map F^{-1} is well defined.

of the hypercube due to its geometry. At this vertex, $|(\Delta x_j)_i| = \epsilon_i$, and hence (4.5) reduces to

$$w_i = \epsilon_i \left(\sum_{j=0}^{n-1} (\delta_j)_i \tilde{y}_{j+1} - (\delta_n)_i \right),$$

where $|(\delta_j)_i| = |(\Delta x_j)_i/\epsilon_i| = 1$. It follows that

$$|\epsilon_i| \geqslant \frac{|w_i|}{\|\tilde{y}\|_1 + 1},$$

where the equality is attained (and hence ϵ_i is minimized) at the vertex for which $(\delta_j)_i = \operatorname{sgn}(y_{j+1} w_i)$ and $(\delta_n)_i = -\operatorname{sgn} w_i$. Now that we found the minimum ϵ_i for each i, we can define the solution of the linear programming problem (4.5)–(4.6) as $\epsilon = \max_i \epsilon_i$:

$$\tilde{\epsilon} = \frac{\max_{1 \leqslant i \leqslant n} |w_i|}{\|\tilde{y}\|_1 + 1} = \frac{\|X_0(y - \tilde{y})\|_\infty}{\|\tilde{y}\|_1 + 1}. \tag{4.7}$$

Equation (4.7) provides an upper bound $\bar{\epsilon}_{\mathrm{U}}$ on ϵ_{U} for any appropriate choice of $\tilde{y}$; this approach does not provide an explicit formula for the minimizer of the linear programming problem, however.

If one desires an analytical upper bound together with an explicit formula for the perturbations Δx_i that realize this bound, one can use the inequality

$$\epsilon_{\mathrm{U}} \leqslant \min_{\Delta X_0} \max_{0 \leqslant j \leqslant n} \{\|\Delta x_j\|_\infty\} \leqslant \min_{\Delta X_0} \max\{\|\Delta X_0\|, \|\Delta X_0 \tilde{y} + X_0(\tilde{y} - y)\|\}, \tag{4.8}$$

where the matrix norm $\|\cdot\|$ can be either $\|\cdot\|_\infty$ or $\|\cdot\|_1$. Two crude estimates of (4.8) can be obtained by putting $\Delta X_0 = 0$, which implies $\epsilon_{\mathrm{U}} \leqslant \|X_0(\tilde{y} - y)\|$, or by choosing ΔX_0 such that $\Delta X_0 \tilde{y} + X_0(\tilde{y} - y) = 0$ (for example, as $\Delta X_0 = X_0(y - \tilde{y})\tilde{y}^T/\|\tilde{y}\|_2^2$), which yields $\epsilon_{\mathrm{U}} \leqslant \|X_0(y - \tilde{y})\tilde{y}^T\|/\|\tilde{y}\|_2^2$. A more refined approximation is then provided by the following convex interpolation of the two crude estimates: $\Delta X_0 = \alpha w \tilde{y}^T$, with $0 \leqslant \alpha \leqslant \|\tilde{y}\|_2^{-2}$, and $w = X_0(y - \tilde{y})$. An optimum in (4.8) is reached when

$$\|\Delta X_0\| = \|\Delta X_0 \tilde{y} + X_0(\tilde{y} - y)\|,$$

which implies

$$\alpha \|w \tilde{y}^T\| = \|w\|(1 - \alpha \|\tilde{y}\|_2^2)$$

and hence

$$\alpha = \frac{\|w\|}{\|w \tilde{y}^T\| + \|w\| \|\tilde{y}\|_2^2}.$$

An upper bound on ϵ_{U} (distinct from (4.7)) is therefore provided by the quantity

$$\bar{\epsilon}_{\mathrm{U}} = \frac{\|X_0(y - \tilde{y})\| \|X_0(y - \tilde{y})\tilde{y}^T\|}{\|X_0(y - \tilde{y})\tilde{y}^T\| + \|X_0(y - \tilde{y})\| \|\tilde{y}\|_2^2}. \tag{4.9}$$

The upper bound estimates given above, whether they are obtained as a solution of the linear programming problem (4.7) or using (4.9), depend on the choice of $\tilde{y}$, i.e., the choice of eigenvalues of the perturbed matrix $\tilde{\Phi}$. One can refine these bounds by further optimization over all appropriate values of those eigenvalues.

4.3. Numerical bound. In addition to finding analytical upper and lower bounds on the uncertainty of the data using the techniques described above, one can also take a numerical approach to estimating ϵ_{U}. For simplicity in representing the set $C(d,\epsilon)$ graphically, we will focus our discussion on the case of a two-dimensional system; however, the approach can be extended to n dimensions.

Fix $d = (x_0, x_1, x_2) \in (R^2)^3$. To estimate ϵ_{U}, we will discretize the surface of $C(d,\epsilon)$ and examine whether each grid point yields a unique inverse. By gradually increasing ϵ we can find the bound as the largest value of ϵ for which a grid point fails to give a unique inverse. In practice, we surround each data point x_j with a collection M_j of points equally spaced along the edge of a square with center point x_j and side length 2ϵ (see Figure 1). Depending on the desired precision, we choose M_j to consist of 8, 16, or 32 grid points. Then, we pair any point $p_0 \in M_0$ with any points $p_1 \in M_1$ and $p_2 \in M_2$ to define the matrix $\Phi = [p_1 \,|\, p_2][p_0 \,|\, p_1]^{-1}$. In accordance with Theorem 2.2, the eigenvalues of Φ will determine whether the solution to $\Phi = e^A$ is unique.

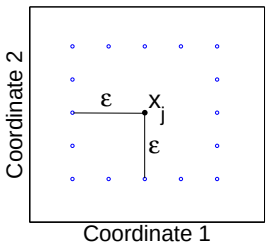

Figure 1. *Grid M_j surrounding a sample data point.*

4.4. Analytical bounds for additional properties. Using the results we have obtained so far, we can derive upper and lower bounds on the uncertainty in data that preserves additional qualitative properties of the solution to the inverse problem, as long as these properties can be defined as conditions on the eigenvalues of the matrix Φ. For example, let $d \in \mathcal{D}$ be such that Φ has n distinct real eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $0 < \lambda_j < 1$, $j = 1, \ldots, n$. Then, the associated matrix A (with $\Phi = e^A$) has n distinct negative eigenvalues, and hence the equilibrium is a stable node. Let ϵ_{SN} define the maximal permissible uncertainty in the data d under which the equilibrium remains a stable node. A lower bound on ϵ_{SN} can be obtained using the same argument as in Theorem 4.1, except with the bound on maximum perturbation of eigenvalues, δ_{U}, replaced by the quantity δ_{SN} that guarantees that the perturbed eigenvalues remain real, distinct, and between 0 and 1. Below we present, without proofs, analytical lower bound statements analogous to Theorem 4.1 for the cases of a stable node and a stable system, as well as for the case in which we require no solution to exist. We leave it to the reader to derive upper bounds using the line of reasoning presented in section 4.2.

Theorem 4.5. *Let $d \in \mathcal{D}$ be such that Φ has n distinct positive eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $0 < \lambda_j < 1$, $j = 1, \ldots, n$. Let $m_1 = \frac{1}{2}\min_{i \neq j}|\lambda_i - \lambda_j|$, $m_2 = \min_{1 \leqslant j \leqslant n}\{\lambda_j\}$, $m_3 = \min_{1 \leqslant j \leqslant n}\{1 - \lambda_j\}$, and $\delta_{\text{SN}} = \min\{m_1, m_2, m_3\}$. If $\epsilon > 0$ is such that $\epsilon \leqslant \underline{\epsilon}_{\text{SN}} = f(\delta_{\text{SN}}, d)$ with f defined as in (4.1), then for any $\tilde{d} \in C(d,\epsilon)$, $\tilde{\Phi}$ has n distinct positive eigenvalues $\tilde{\lambda}_j$*

with $0 < \tilde{\lambda}_j < 1,\ j = 1, \ldots, n$, *and hence the equation* $e^{\tilde{A}} = \tilde{\Phi}$ *has a unique matrix solution* $\tilde{A}$ *for which the origin of* (2.1) *is a stable node.*

To guarantee stability of the equilibrium with respect to all inverse problem solutions without demanding uniqueness of a solution, we require that $d \in \mathcal{D}$ be such that the eigenvalues of Φ are not real negative and satisfy $0 < |\lambda_j| < 1,\ j = 1, \ldots, n$. Then, the associated matrix A (with $\Phi = e^A$) has (possibly complex) eigenvalues with negative real part, and thus the equilibrium at the origin is stable. Let ϵ_{S} define the maximal permissible uncertainty in the data d such that the equilibrium remains stable. A lower bound on ϵ_{S} is obtained in the following result.

Theorem 4.6. *Let* $d \in \mathcal{D}$ *be such that the eigenvalues of* Φ *are not real negative and satisfy* $0 < |\lambda_j| < 1,\ j = 1, \ldots, n$. *Let* $m_1 = \min_{1 \leqslant j \leqslant n}\{1 - |\lambda_j|\}$, $m_2 = \min_j |\lambda_j|$ *for all* j *such that* $\mathrm{Re}(\lambda_j) > 0$, $m_3 = \min_j |\mathrm{Im}(\lambda_j)|$ *for all* j *such that* $\mathrm{Re}(\lambda_j) < 0$, *and* $\delta_{\mathrm{S}} = \min\{m_1, m_2, m_3\}$. *If* $\epsilon > 0$ *is such that* $\epsilon \leqslant \underline{\epsilon}_{\mathrm{S}} = f(\delta_{\mathrm{S}}, d)$ *with* f *defined as in* (4.1), *then for any* $\tilde{d} \in C(d, \epsilon)$, *the eigenvalues of* $\tilde{\Phi}$ *satisfy* $0 < |\tilde{\lambda}_j| < 1,\ j = 1, \ldots, n$, *and are not real negative. Hence the equation* $e^{\tilde{A}} = \tilde{\Phi}$ *has a solution* $\tilde{A}$, *and every such solution has all eigenvalues with negative real part.*

Finally, it is interesting to consider the case of nonexistence of an inverse. In particular, given data d for which a real inverse A does not exist, i.e., data that do not represent the trajectory of any real linear system, what is the greatest amount of uncertainty for which nonexistence of a real solution persists, and hence a linear model should not be considered as a possible mechanism underlying the observed uncertain data? Let $d \in \mathcal{D}$ be such that Φ has at least one negative real eigenvalue of odd multiplicity, which implies that there is no real matrix A such that $\Phi = e^A$. Let ϵ_{DNE} define the maximal permissible uncertainty in the data d under which the inverse problem will be guaranteed to remain without a real solution. A lower bound on ϵ_{DNE} is obtained in the following result.

Theorem 4.7. *Let* $d \in \mathcal{D}$ *be such that* Φ *has at least one negative real eigenvalue of odd multiplicity. Let the collection of such eigenvalues be denoted by* $\lambda_1, \ldots, \lambda_k$, *where* λ_k *is the closest to zero. Let* $m_1 = \max_{1 \leqslant i \leqslant k}(\min_{j(j \neq i)} \frac{1}{2}|\lambda_j - \lambda_i|)$, $m_2 = |\lambda_k|$, *and* $\delta_{\mathrm{DNE}} = \min\{m_1, m_2\}$. *If* $\epsilon > 0$ *is such that* $\epsilon \leqslant \underline{\epsilon}_{\mathrm{DNE}} = f(\delta_{\mathrm{DNE}}, d)$ *with* f *defined as in* (4.1), *then for any* $\tilde{d} \in C(d, \epsilon)$, $\tilde{\Phi}$ *has at least one negative eigenvalue of odd multiplicity, and hence the equation* $e^{\tilde{A}} = \tilde{\Phi}$ *has no real solution.*

5. Examples for two-dimensional systems.

In the case of two-dimensional linear systems, one can represent several of the previous results in a more explicit fashion and extend the stability results to encompass various classifications of the equilibrium. We present these extensions here, along with several numerical examples that can be conveniently depicted in the phase plane.

5.1. Regions of existence and uniqueness of the inverse.

The criteria in Theorems 2.1 and 2.2 are based on the Jordan structure of Φ, which for real 2×2 matrices can only take a few different forms and hence can be analyzed completely. Utilizing the relationship between the eigenvalues of a matrix and its trace and determinant, the criteria for existence and uniqueness of the matrix logarithm of Φ can be fully characterized by conditions on the

trace and determinant of Φ which, in turn, can be expressed as conditions on the data from which Φ is constructed. For notational simplicity, let $D = \det \Phi$ and $T = \operatorname{tr} \Phi$.

The analysis is based on the following straightforward corollaries of Culver's theorems.

Corollary 5.1. *Let $\Phi \in \mathbb{R}^{2\times 2}$. There exists $A \in \mathbb{R}^{2\times 2}$ such that $\Phi = e^A$ if and only if $D \neq 0$ and any of the following holds:*

(a) $D > 0$, $T > 0$, *and* $T^2 \geqslant 4D$,
(b) $T^2 < 4D$,
(c) $\Phi = \lambda I$, *with* $\lambda < 0$.

Corollary 5.2. *Let $\Phi \in \mathbb{R}^{2\times 2}$. There exists a unique $A \in \mathbb{R}^{2\times 2}$ such that $\Phi = e^A$ if and only if $T > 0$, $T^2 \geqslant 4D > 0$, and $\Phi \neq \lambda I$ for all $\lambda \in \mathbb{R}$.*

In two dimensions, the matrix Φ is constructed from the uniformly spaced data $d = (x_0, x_1, x_2) \in \mathcal{D}$, as $\Phi = X_1 X_0^{-1}$. The determinant and trace of Φ are then given by $D = \det X_1 / \det X_0$ and $T = \det[x_0|x_2] / \det X_0$. The conditions defined in Corollaries 5.1 and 5.2 thus define regions in six-dimensional data space in which the inverse A exists or exists and is unique.

Once we know that a real matrix A exists, additional conditions on the data may be found that define certain important properties of the system. The classification of the equilibrium point at the origin is easily determined by eigenvalues of A (as found in any standard textbook on theory of systems of linear ODEs) and hence can also be transformed into conditions on T and D.

Theorem 5.3. *The equilibrium point $x = 0$ of system (2.1) with matrix $A \in \mathbb{R}^{2\times 2}$ is*

(a) *a stable node if $T > 0$, $T^2 > 4D > 0$, and $1 > D > T - 1$;*
(b) *an unstable node if $T > 0$ and $T^2 > 4D > 4(T - 1)$;*
(c) *a saddle if $T - 1 > D > 0$;*
(d) *a stable spiral if $T^2 < 4D < 4$;*
(e) *an unstable spiral if $T^2 < 4D$ and $D > 1$;*
(f) *a center if $D = 1$ and $T^2 < 4$.*

Figure 2(a) graphically summarizes the results of Corollaries 5.1 and 5.2 and Theorem 5.3 in the T, D plane. The region where A is unique contains all of the systems where the equilibrium point is a saddle or node (marked in blue), and the region corresponding to a nonunique A contains all of the systems where the equilibrium point is a spiral or a center (marked in black). The label DNE indicates that in this region, the matrix logarithm results in complex matrices A, whereas we are only interested in data resulting from systems with real parameters ($A \in \mathbb{R}^{2\times 2}$). (Note that the diagram differs from that presented in standard textbooks for 2×2 linear systems because here T and D refer to the trace and determinant of $\Phi = e^A$ and not the trace and determinant of A.)

Just like Corollaries 5.1 and 5.2, Theorem 5.3 defines regions in six-dimensional data space in which the inverse solution has various stability properties. By fixing two of the data points, one can visualize two-dimensional cross sections of these regions defined by conditions on the third data point. Figure 2(b) shows the outcomes associated with different regions where x_2 can be located, given example locations of x_0 and x_1.

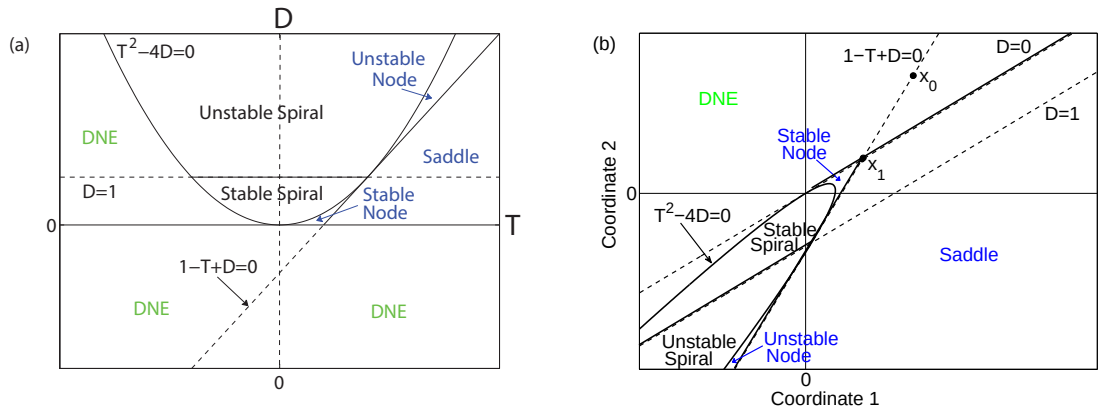

Figure 2. (a) *Properties of the inverse problem solution depicted in terms of conditions on* $D = \det \Phi$ *and* $T = \operatorname{tr} \Phi$. *Uniqueness of the solution to the inverse problem holds in the regions labeled in blue, and existence holds in the regions labeled in black.* (b) *Conditions on the position of* x_2 *to give various properties when* x_0 *and* x_1 *are fixed. In both figures, solid lines form boundaries between regions; dashed lines do not.*

5.2. Bounds on maximal permissible uncertainty. In this section we will illustrate the dependence of the maximal permissible uncertainty on the data and the property being maintained.

Example 5.4 (differences between analytical bounds and numerical estimates of maximal permissible uncertainty). Let $d = (x_0, x_1, x_2) = ((10, 2)^T, (6.065, -4.44)^T, (7, -10)^T)$, which are points that are equally spaced in time on a trajectory of the system (2.1) with

$$A = \begin{bmatrix} -0.6724 & -0.7201 \\ -0.8610 & -0.0244 \end{bmatrix}.$$

Since $\Phi = e^A$ has two distinct positive eigenvalues, A is the unique matrix that yields the data d. For this data set, the direct numerical estimate of the maximum permissible uncertainty for uniqueness is $\tilde{\epsilon}_{\mathrm{U}} = 1.075$. The analytical lower bounds on ϵ_{U} are substantially smaller than the numerical bound, namely $\underline{\epsilon}_{\mathrm{U}}^{\infty} = 0.083$ and $\underline{\epsilon}_{\mathrm{U}}^{1} = 0.149$ (where $\underline{\epsilon}_{\mathrm{U}}^{\infty}$ is found by applying the norm $\|\cdot\|_\infty$ and $\underline{\epsilon}_{\mathrm{U}}^{1}$ by using the norm $\|\cdot\|_1$ in Theorem 4.1).

The analytical upper bound depends on the choice of eigenvalues for the perturbed matrix $\tilde{\Phi}$. Using (4.9) and choosing a perturbed matrix with one zero eigenvalue and the second eigenvalue equal to the average of eigenvalues of Φ yields the upper bound $\overline{\epsilon}_{\mathrm{U}} = 1.245$. Optimization over the value of the second eigenvalue yields a better estimate $\overline{\epsilon}_{\mathrm{U}} = 1.078$, which is essentially identical to the numerical upper bound. Choosing a perturbed matrix with identical eigenvalues equal to the average of eigenvalues of Φ leads to $\overline{\epsilon}_{\mathrm{U}} = 1.841$, and optimization over the position of the double eigenvalue gives the upper bound $\overline{\epsilon}_{\mathrm{U}} = 1.570$. Using the linear programming estimate of (4.7) gives the same bounds as the zero eigenvalue choice of perturbation. Choosing a perturbed matrix with identical eigenvalues equal to the average of eigenvalues of Φ and solving (4.9) yields $\overline{\epsilon}_{\mathrm{U}} = 1.224$, and optimization over the position of the double eigenvalue gives $\overline{\epsilon}_{\mathrm{U}} = 1.140$. A summary of the optimal bounds is given in the second column of Table 1.

Table 1
Best estimates of ϵ_{X}, $\mathrm{X} \in \{\mathrm{SN}, \mathrm{U}, \mathrm{S}, \mathrm{DNE}\}$, *for Example* 5.5.

Property (X)	Stable node (SN)	Unique (U)	Stable (S)	Nonexistence (DNE)
x_2	$(3.68, -3.46)$	$(7, -10)$	$(3.1, -5.5)$	$(3, 6)$
$\tilde{\epsilon}_{\mathrm{X}}$ numerical estimate	0.216	1.075	0.519	3.055
$\underline{\epsilon}_{\mathrm{X}}$ by Theorem 4.1	0.059	0.149	0.207	0.609
$\bar{\epsilon}_{\mathrm{X}}$ by (4.7)	0.217^{*}	$1.078^{\dagger}$	$0.519^{\ddagger}$	$3.059^{\#}$
$\bar{\epsilon}_{\mathrm{X}}$ by (4.9)	0.251^{*}	$1.078^{\dagger}$	$0.665^{\ddagger}$	$3.683^{\dagger}$

$^{*}\lambda_1 = \lambda_2$, $^{\dagger}\lambda_1 = 0$, $^{\ddagger}\lambda_2 = 1$, $^{\#}\lambda_2 \to \infty$.

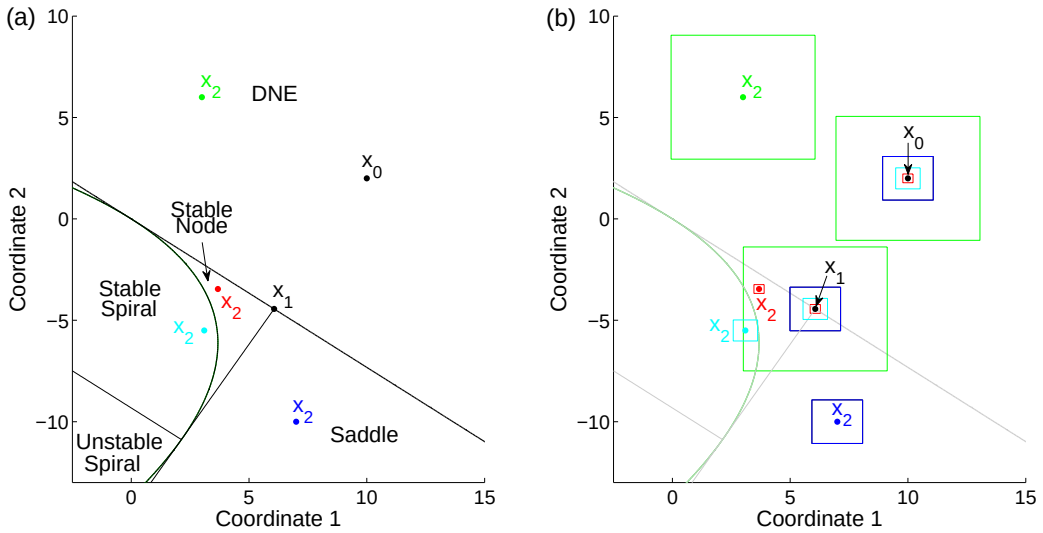

Figure 3. (a) *Regions where particular properties hold when* x_0 *and* x_1 *are fixed, based on Theorem* 5.3. *Various* x_2 *are selected, and estimates of the maximal permissible uncertainty* ϵ_{X} *are then computed.* (b) *Squares depict the numerically obtained bounds on the uncertainty allowed to preserve stable node (red), stability (cyan), uniqueness (blue), and nonexistence (green) properties. Coordinates of* x_2 *and numerical values of the bounds are listed in Table* 1.

Example 5.5 (dependence of maximal permissible uncertainty on x_2). As in Example 5.4, let $x_0 = (10, 2)^T$, $x_1 = (6.065, -4.44)^T$. Theorem 5.3 defines regions in $\mathbb{R}^2$ that specify the nature of the equilibrium based on the location of the last data point x_2. The boundaries of the regions are shown in Figure 3(a). We select a sample point x_2 from each labeled region, and in each case, we compute various estimates of the maximal permissible uncertainty ϵ to preserve the corresponding property. We depict each uncertainty by outlining in Figure 3(b) the square-shaped sets $c(x_0, \epsilon)$, $c(x_1, \epsilon)$, and $c(x_2, \epsilon)$ in the phase plane $\mathbb{R}^2$, which can be interpreted as follows: Given any $\tilde{x}_0 \in c(x_0, \epsilon)$, $\tilde{x}_1 \in c(x_1, \epsilon)$, and $\tilde{x}_2 \in c(x_2, \epsilon)$, the matrix A that yields the data $d = (\tilde{x}_0, \tilde{x}_1, \tilde{x}_2)$ has the appropriate property. Table 1 summarizes the location of x_2, the property being preserved, and the best estimate of ϵ in each case. As can be seen in Figure 3, it appears that the proximity of x_2 to the boundary of the region in which the desired property holds impacts the size of the resulting ϵ.

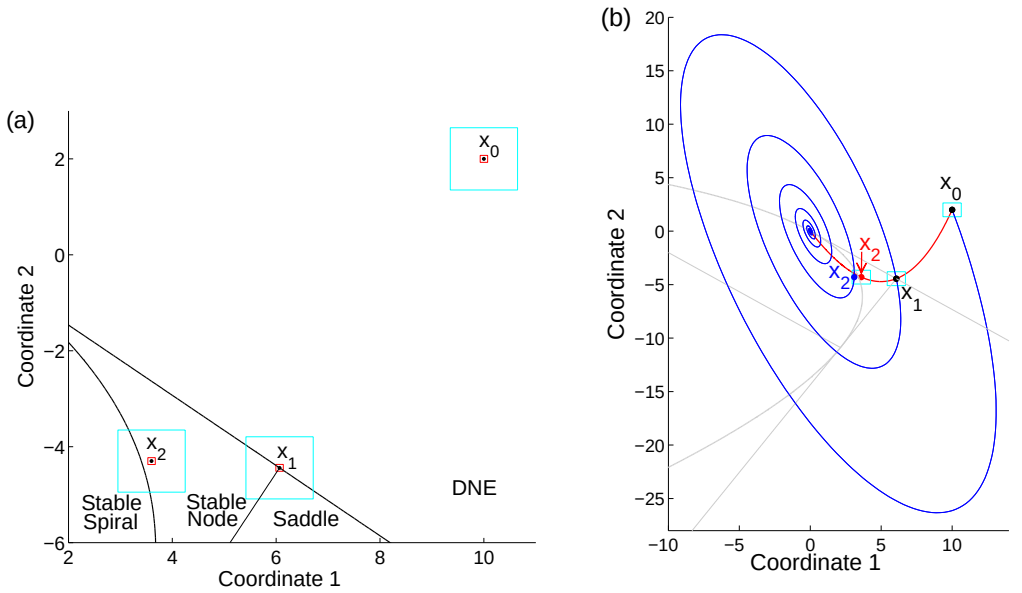

Figure 4. (a) *Numerical estimates* $\tilde{\epsilon}_{\mathrm{U}} = \tilde{\epsilon}_{\mathrm{SN}}$ *(red) and* $\tilde{\epsilon}_{\mathrm{S}}$ *(cyan) for data depicted.* (b) *Stable node trajectory (red) that passes through the data* d *at equally spaced time points, and stable spiral trajectory (blue) through perturbed data* $\tilde{d} = (x_0, x_1, (3.1, -4.3)^T)$*, where* $\tilde{d} \in C(d, \tilde{\epsilon}_{\mathrm{S}})$ *but* $\tilde{d} \notin C(d, \tilde{\epsilon}_{\mathrm{U}})$*.*

Example 5.6 (dependence of maximal permissible uncertainty on the choice of solution property). Consider the data set d with x_0, x_1 as in Examples 5.4 and 5.5 and with fixed $x_2 = (3.6, -4.3)^T$, corresponding to a dynamical system with a stable node at the origin. The maximal permissible uncertainty of the data depends on which property we require to be preserved. In this case we can choose among uniqueness, the stable node property, and stability. If we want to guarantee a unique solution to the inverse problem, we find that $\tilde{\epsilon}_{\mathrm{U}} = 0.072$. This is quite small due to the proximity of x_2 to the border of the stable spiral region where A is nonunique, as seen in Figure 4(a). Note that $c(x_2, \tilde{\epsilon}_{\mathrm{U}})$ does not extend all the way out to the boundary of the stable spiral region, because these boundary lines are derived with x_0 and x_1 fixed, but we allow uncertainty in all three data points. For preservation of the stable node property, we observe that in this case $\tilde{\epsilon}_{\mathrm{SN}} = \tilde{\epsilon}_{\mathrm{U}}$; however, this relation does not hold universally. It would not be true, for example, if x_2 were located within the stable node region but closer to the saddle region where A is unique. If we want to ensure that the data lie on a trajectory that converges to the origin (i.e., preserve the stability of the system), we find that the maximal permissible uncertainty is $\tilde{\epsilon}_{\mathrm{S}} = 0.647$, which is significantly larger than $\tilde{\epsilon}_{\mathrm{U}}$. Thus we can guarantee stability for larger uncertainty in the data than what is needed to preserve the uniqueness of solutions. Figure 4(a) depicts the data in the phase plane and the numerically computed maximal permissible uncertainties for the uniqueness, stable node, and stability properties. Within the set $C(d, \tilde{\epsilon}_{\mathrm{S}})$, any choice of data will lie on a stable trajectory; however, the uniqueness of the inverse problem may not be preserved. In Figure 4(b) we illustrate two different data sets contained in $C(d, \tilde{\epsilon}_{\mathrm{S}})$: one data set on a stable node trajec-

tory (red), corresponding to a unique A, and a second data set belonging to a stable spiral trajectory (blue), where A is not unique.

6. Application to a model of gene regulation. Holter et al. [12] use linear dynamical models to describe the time evolution of gene expression levels in various biological systems. They first deduce characteristic modes of the gene expression data using the singular value decomposition and then derive a discrete time transition matrix to model the time evolution of the system. One system they present describes the dynamics of six groups of genes involved in sporulation. The original measurements of expression levels of 1116 genes, collected in irregular intervals over a period of 12 hours, were ordered to obtain similar expression patterns and divided into six groups to reflect the time of first induction or repression [7]. The translation matrix N reported in [12] describes the transition between levels $C(t)$ and $C(t + \Delta t)$, where $\Delta t = 1$.

For the purpose of illustrating the implications of our results on the interpretation of the sporulation model of Holter et al. [12], we have transformed the irregularly spaced data into equally spaced data (with an interval of 2 hours) and performed uncertainty analysis. The transformed data are shown in Figure 5, and the trajectories of the model are shown in Figure 6, upper left panel (unstable spiral).

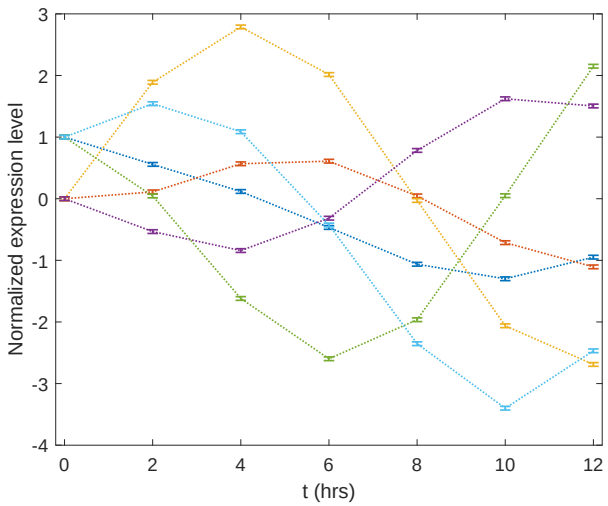

Figure 5. *Data for the sporulation gene dynamics model of* [12]. *Error bars indicate the upper bound on maximum permissible uncertainty for spiral behavior. Data points are connected by dotted lines to guide the eye.*

The matrix A corresponding to the dynamical system has eigenvalues $-0.1840 \pm 0.5117i$, $0.0500 \pm 0.4420i$, $0.0264 \pm 0.1954i$ and hence has a four-dimensional unstable manifold with spiral dynamics and a two-dimensional stable manifold with spiral dynamics. We have used the methods presented in this work to characterize the maximum permissible uncertainty that preserves either the unstable spiral behavior of the model, the spiral property of the model, or the existence of the inverse, labeled as ϵ_{UnSp}, ϵ_{Sp}, or ϵ_{E}, respectively. The results are shown in Table 2. The upper bound estimates are not sharp since they have not been optimized over all possible choices of $\tilde{y}$.

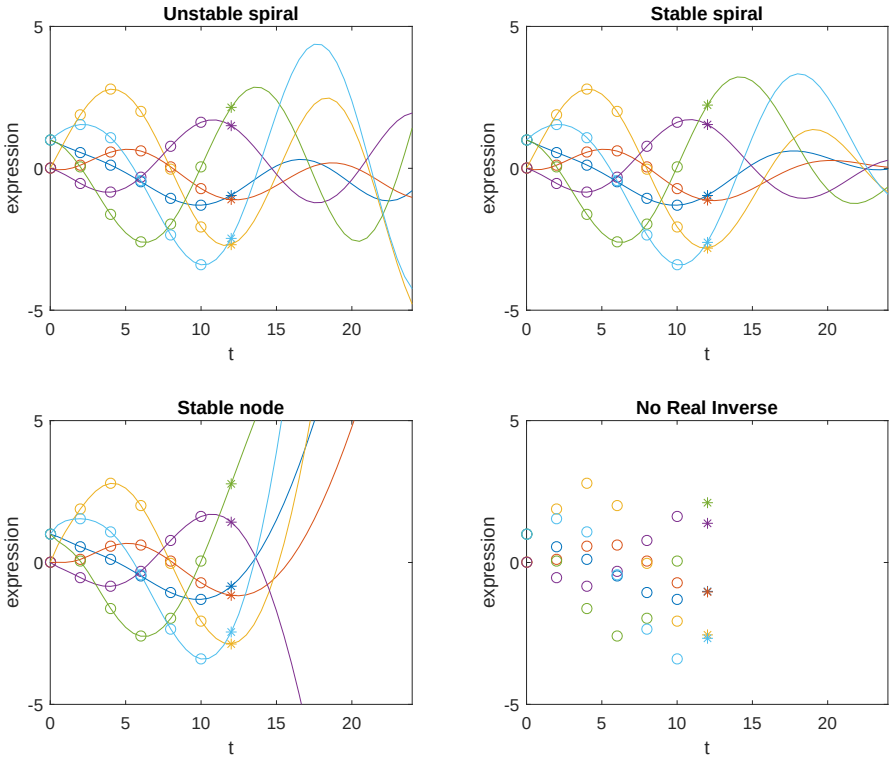

Figure 6. *Trajectories of models obtained by perturbing the last data point of the sporulation gene dynamics model and computing the corresponding parameter matrix A. The five unperturbed data points (identical in each panel) are marked with circles. The perturbed data point is labeled with a star; these points differ across panels and lead to different properties of the inverse problem solution, including nonexistence of a real parameter matrix for the data in the lower right panel. Within each panel, each color represents a different variable of the system, while the same variables share the same color across panels. See the supplemental material* (M106246_supplement.pdf [local/web 3.70MB]) *for an expanded view of the trajectories.*

Table 2
Best estimates of ϵ_X for sporulation model.

Property (X)	Unstable spiral (UnSp)	Spiral (Sp)	Existence (E)
$\tilde{\epsilon}_X$ numerical estimate	9.5×10^{-4}	0.0017	0.0015
$\underline{\epsilon}_X$ by Theorem 4.1	2.52×10^{-5}	1.53×10^{-4}	1.31×10^{-4}
$\bar{\epsilon}_X$ by (4.7)	0.0037	0.0301	0.0111
$\bar{\epsilon}_X$ by (4.9)	0.0145	0.0654	0.0329

The uncertainty ϵ_{UnSp} is, not surprisingly, rather small, since the unstable eigenvalues of A are close to zero, and any perturbation of the data that results in those eigenvalues crossing the imaginary axis will stabilize the system. Surprisingly, ϵ_{Sp} is also small, as indicated by the error bars in Figure 5. Once the upper bound is reached, there exist data for which the corresponding matrix has eigenvalues with positive real part and no imaginary component. Finally, ϵ_{E} appears to be larger than ϵ_{UnSp} but potentially smaller than ϵ_{Sp}, which indicates that

the loss of existence of a real inverse can be induced by a smaller perturbation than can the loss of oscillatory behavior. The effects of selected perturbations of the data on the trajectory of the system can be seen in Figure 6. Note that when the equilibrium is a stable node, although the trajectories appear to diverge, they do approach 0 after an excursion to no more than 350 on the vertical axis (see Figure 3.1 of the supplemental material M106246_supplement.pdf [local/web 3.70MB]).

7. Remark on nonuniform spacing of data. In the results presented so far we have assumed that the data available about the system (2.1) are spaced uniformly in time, i.e., that d contains the data points $x_0, x_1, x_2, \ldots, x_n \in \mathbb{R}^n$, where $x_j = x(j; A, b) = e^{Aj}b$. In this section we shall make a few remarks on how the results can be extended to the case in which data are spaced nonuniformly, with the restriction that the sampling times are still integer multiples of some Δt, assumed without loss of generality to be equal to 1. Such situations occur frequently for medical data, which are usually collected more frequently during the early course of a disease and less frequently during recovery.

The following lemma relates the eigenvalues of $\Phi = e^A$ to the data $d = (x_{j_0}, x_{j_1}, x_{j_2}, \ldots, x_{j_n})$.

Lemma 7.1. *Assume that for $i = 0, \ldots, n$, $x_{j_i} = \Phi^{j_i} b$, where j_i are integers such that $0 = j_0 < j_1 < \cdots < j_n$. Let $X_0 = [x_{j_0} \mid \ldots \mid x_{j_{n-1}}]$, and assume that X_0 is invertible and Φ is nonsingular. Let $y = X_0^{-1} x_{j_n}$ be a vector with entries $y_1, y_2, \ldots, y_n$. Then λ is an eigenvalue of Φ only if λ is a root of the polynomial*

$$\lambda^{j_n} - y_n \lambda^{j_{n-1}} - \cdots - y_2 \lambda^{j_1} - y_1 = 0. \tag{7.1}$$

Proof. Since $y = X_0^{-1} x_{j_n}$, it follows that $x_{j_n} = X_0 y$, i.e.,

$$\Phi^{j_n} b = \sum_{i=0}^{n-1} y_{i+1} \Phi^{j_i} b. \tag{7.2}$$

The assumption that X_0 is invertible is equivalent to the statement that the vectors $\{x_{j_i}\}_{i=0}^{n-1}$ are not confined to a proper subspace of $\mathbb{R}^n$, and hence, if $x_{j_i} = \Phi^{j_i} b = e^{Aj_i} b$, b is not confined to a proper Φ-invariant subspace of $\mathbb{R}^n$. Thus b, when decomposed in the basis of ordinary and generalized eigenvectors of Φ, has a nonzero component along the ordinary or generalized eigenvector for every eigenvalue of Φ. Let v be the eigenvector corresponding to the eigenvalue λ of Φ. If b has a nonzero component along v, then (7.2) implies that

$$\lambda^{j_n} v = \sum_{i=0}^{n-1} y_{i+1} \lambda^{j_i} v,$$

which proves the statement. If the component of b along v is zero, but b has a nonzero component along u such that $(A - \lambda I)^{k-1} u \neq 0$ and $(A - \lambda I)^k u = 0$, then (7.2) implies that

$$\lambda^{j_n + k} u = \sum_{i=0}^{n-1} y_{i+1} \lambda^{j_i + k} u,$$

which for nonzero λ reduces to the previous case. ∎

Lemma 7.1 can be employed in solving the inverse problem as follows: First, the vector y is computed from the data $d = (x_{j_0}, x_{j_1}, x_{j_2}, \ldots, x_{j_n})$. Second, all roots of the polynomial (7.1) are found using numerical techniques. Third, a combination of n distinct roots is chosen from the collection of roots. Fourth, the companion matrix $\hat{\Phi}$ of Φ is formed from the chosen roots, and a set of vectors z_{j_i}, $i = 0, 1, \ldots, n-1$ is computed by taking $z_0 = e_1$ and $z_{j_{i+1}} = \hat{\Phi}^{j_{i+1}-j_i} z_{j_i}$. Finally, the matrix Φ is found as $\Phi = P^{-1}\hat{\Phi}P$, where P is the (unique) $n \times n$ matrix such that $Px_{j_i} = z_{j_i}$ for $i = 0, 1, \ldots, n-1$.

The procedure outlined above can find any matrix Φ that is robust, in the sense of Corollaries (3.1)–(3.3). The condition that X_0 be invertible (i.e., b not be confined to a proper Φ-invariant subspace) is essential for identifiability of A, as we have already observed in [17] for the special case of $j_i = i$. If the indices j_i differ by more than 1, then the polynomial (7.1) has more roots than the matrix Φ has eigenvalues. Arbitrary combinations of such roots will lead to different alternative matrices Φ and hence to nonuniqueness of solutions of the inverse problem. By an appropriate choice of the roots that make up the eigenvalues of Φ, one may be able to control the properties of the matrix Φ and, in turn, the existence, uniqueness, and stability properties of the parameter matrix A of the system (2.1).

8. Remark on data size. Since in this paper we are primarily concerned with the invertibility of the solution map, we have naturally assumed that the parameter and data spaces are of the same dimension. This is a limitation from a practical point of view since in realistic scenarios the relation between the number of parameters and the size of the data may not be under control. We will briefly address here three such scenarios.

Scenario 1: If the number of time points at which data are collected is smaller than $n+1$, then $\dim(\mathcal{D}) < \dim(\mathcal{P})$, and it is clear that the parameter matrix cannot be completely determined from the data. Surprisingly, just one missing time point results in a complete loss of information about the eigenvalues of the solution matrix Φ. This follows from the companion formulation $\hat{\Phi} = X_0^{-1}X_1$, introduced in section 4.2. Since $\hat{\Phi}$ is determined by the eigenvalues of Φ, for any choice of such eigenvalues, i.e., for any choice of $\hat{\Phi}$, there exists x_n such that the data set $d = (x_0, x_1, x_2, \ldots, x_n)$ is compatible with $\hat{\Phi}$, namely $x_n = X_0 y$. In other words, the knowledge of X_0, i.e., the knowledge of the first n data points $x_0, x_1, \ldots, x_{n-1}$ for an n-dimensional linear dynamical system, does not provide any information about the eigenvalues of that system. Thus, we cannot deduce from just n or fewer data points whether the observed n-dimensional linear system is stable or unstable, whether its fixed point is a node or spiral or saddle, or even whether the data are generated by a system with a real parameter matrix. Figure 6 illustrates this point by showing trajectories of a linear dynamical system (2.1) from Holter et al. [12] that all share an identical matrix X_0 but differ widely in dynamical properties. The same observation can be made about similar cases in which any one of the $n+1$ data points $x_0, x_1, x_2, \ldots, x_n$ is missing.

Scenario 2: If the data is collected at $m \geqslant n+1$ equally spaced time points, say $t = 0, 1, 2, \ldots, m$, but not all the variables are observed, then one can also encounter situations with $\dim(\mathcal{D}) \neq \dim(\mathcal{P})$. In this case one may be able recover from the data the eigenvalues of the matrix Φ, but not the trajectories of the unobserved variables. To this end, note that, just like above, $x_n = X_0 y = [x_0|x_1|\ldots|x_{n-1}]y$ and hence the data obey for any j and any

variable i the linear difference equation

$$(x_{j+n})_i = y_n(x_{j+n-1})_i + \cdots + y_2(x_{j+1})_i + y_1(x_j)_i. \tag{8.1}$$

Furthermore, recall that the kth data point is related to the initial condition as $x_k = \Phi^k b$, and suppose further that Φ is diagonalizable with distinct eigenvalues, $\Phi = S\Lambda S^{-1}$, where the eigenvectors s_i are scaled so that $b = \sum_j s_j$ (we here implicitly assume that the initial condition has components along every eigenvector, and hence the system is identifiable from the full trajectory). Then, for all i, $(x_k)_i = \sum_j (s_j)_i \lambda_j^k$. Standard techniques for solving linear difference equations imply that by collecting $2n+1$ consecutive data $(x_0)_i, \ldots, (x_2n)_i$ for any single variable i, one can recover the entries of the vector y from which one can find the eigenvalues $\lambda_1, \ldots, \lambda_n$ and, subsequently, solve for $(s_1)_i, \ldots, (s_n)_i$. If data for other variables are known, then more entries of the matrix S can be recovered. Incidentally, one can use (8.1) to predict the future trajectory of the ith variable without the knowledge of the behavior of any other variables.

Scenario 3: If the data are collected at $m > n+1$ time points, say $t = 0, 1, 2, \ldots, m$, and all variables are observed, then $\dim(\mathcal{D}) > \dim(\mathcal{P})$. If all of the data lie on a single trajectory, then one can use any collection of consecutive $n+1$ data points to recover Φ and subsequently A. If the data do not all lie on a single trajectory, possibly due to a measurement error, then the equation $[x_1|x_2|\ldots|x_m] = \Phi[x_0|x_1|\ldots|x_{m-1}]$ is not satisfied. The usual approach is then to find Φ and b that "best" represent the available data. There are many ways to do this, depending on the information one has about the measurement error: (i) Find Φ that minimizes $\|X_1 - \Phi X_0\|$ for some norm of choice; (ii) find y that minimizes $\|x_{j+n} - X_j y\|$; (iii) find Φ, b that minimize $\sum_k \|x_k - \Phi^k b\|$, $\sqrt{\sum_k \|x_k - \Phi^k b\|^2}$, or $\max_k \|x_k - \Phi^k b\|$; or (iv) define Φ as the average of $\Phi_j = X_{j+1} X_j^{-1}$. For example, the quantity $\|X_1 - \Phi X_0\|_F$ (where $\|\cdot\|_F$ denotes the Frobenius norm) is minimized by $\Phi = X_1 X_0^+$, where X_0^+ denotes the Moore–Penrose pseudoinverse of the matrix X_0 [14, 4, 15]. The study of the uncertainty of inverse problem solutions in this case requires a separate treatment in which all methods described above would be analyzed and compared.

9. Discussion.

We have analyzed the inverse problem for linear dynamical systems, i.e., the problem of finding the value of the parameter matrix for which a linear system generates a given discrete data set consisting of points equally spaced in time on a single trajectory. Our results establish regions in data space that give solutions with particular properties, such as uniqueness or stability, and give bounds on the maximal allowable uncertainty in the data set that can be tolerated while maintaining these characteristics.

Three types of bounds on uncertainties are presented: analytical lower bounds, below which properties are guaranteed to hold for all perturbations of data; analytical upper bounds, which provide proven perturbations of data for which properties are guaranteed to be lost; and numerical bounds, derived from direct sampling of data points. Our results indicate that the upper bounds, when optimized over all potential eigenvalues, provide excellent agreement with the numerical estimates. The numerical methods are hypothetically applicable to systems of arbitrary size; however, the combinatorial problem of pairing together all possible data points along grid points can pose a challenge as the dimension of the system increases. Similarly, the computation of the analytical upper bound (via (4.7) or (4.9)) requires optimization that

becomes computationally expensive for larger systems. Although the analytical lower bound significantly underestimates the maximal permissible uncertainty, it provides a bound that is immediately accessible for systems of higher dimension, without increased computation. Since we focused on the derivation of these bounds, the question of how these bounds scale with system size remains open for future investigation. Furthermore, in this work, we have considered only random perturbations of the data matrix Φ. Due to the special construction of the matrix Φ, it may be possible to improve the analytical lower bound by considering structured matrix perturbations. Many results have been established concerning the bounds on eigenvalues for structured perturbations of matrices [13, 11] that may prove useful in this effort.

This paper can be considered as an extension of our earlier work on identifiability of linear and linear-in-parameters dynamical systems from a single trajectory [17]. In that study, we assumed that a trajectory, or a collection of data representing the trajectory, of a dynamical model was given and asked whether there was a unique choice of parameter matrix for which that model could generate the given data. We showed that for a linear initial value problem with coefficient matrix A and initial condition b, uniqueness requires that $\{b, Ab, \ldots, A^{n-1}b\}$ are linearly independent, and this condition can be translated into a condition on the geometric structure of the observed trajectory. For linear-in-parameters systems $\dot{x} = Af(x)$ we found a similar geometrical condition that guarantees identifiability of A.

A variety of earlier works considered identification of linear systems or parameter matrices from discrete data. Allen and Pruess proposed a method for approximating A in system (2.1) from a discrete collection of data points [3]. Their approach begins by defining an approximating function for the data (e.g., a cubic spline approximation), and they use equally spaced points along this curve to compute a matrix $\hat{A}$ that approximates the true parameter matrix A. A key distinction between their work and the initial analysis presented here is that they use points on an approximation of the trajectory, while we assume that the data represent exact points on the actual trajectory; their results also do not treat uncertainty in data.

In other past work, Singer and Spilerman investigated the problem of identifying the matrix Q in the Markov model $P' = QP$, where P and Q are $n \times n$ matrices [16]. They derive conditions for $P = e^Q$ to have a unique solution. Their results are consistent with the findings of Culver, but with additional constraints to account for the requirement that the model be a continuous time Markov structure. They additionally comment on the case of identifying Q from noisy observations and suggest exploring in a neighborhood of P to detect nonuniqueness of the matrix logarithm through observations on the eigenvalues of the matrices in this neighborhood.

It is of interest to note that much work has been done in determining the maximal allowable uncertainty in the parameter matrix A such that the solution to (2.1) remains stable [11]. This well-known bound is called the stability radius. In our investigation of the inverse problem, ϵ_S has an analogous meaning, but we quantify the uncertainty in the data space rather than in parameter space.

Our results also include bounds on regions of data space where the inverse problem cannot be solved. The utility of such results is that they can provide an approach for model rejection. That is, suppose we have a data set d acquired from measurements of some physical

phenomenon that we believe can be modeled with a linear system of differential equations. Perhaps it is known that the measurement error for any data point x_i is approximately given by ϵ. If we find that there is no real matrix A that yields the collected data d and further find that $\epsilon_{\text{DNE}} > \epsilon$, then we can conclude with certainty that the data cannot come from a system that can be modeled with a linear system of differential equations, and thus we can reject the linear model.

Our work is also related to the important problem of determining identifiability in parameter estimation, which seeks a way to explicitly define sets in parameter space on which a model is identifiable. The connection to identifiability is apparent if we consider the set in parameter space defined by $F^{-1}(C(d, \epsilon_{\text{U}}))$. On this set we have that $F^{-1}(d_1) \neq F^{-1}(d_2)$ implies $d_1 \neq d_2 \in C(d, \epsilon_{\text{U}})$, which is to say that two distinct parameter sets must yield distinct data. So, $F^{-1}(C(d, \epsilon_{\text{U}}))$ defines a set in parameter space on which the model is identifiable.

REFERENCES

[1] A. H. Al-Mohy and N. J. Higham, *Improved inverse scaling and squaring algorithms for the matrix logarithm*, SIAM J. Sci. Comput., 34 (2012), pp. C153–C169, https://doi.org/10.1137/110852553.

[2] L. J. S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*, 2nd ed., CRC Press, Boca Raton, FL, 2011.

[3] R. C. Allen and S. A. Pruess, *An analysis of an inverse problem in ordinary differential equations*, SIAM J. Sci. Statist. Comput., 2 (1981), pp. 176–185, https://doi.org/10.1137/0902015.

[4] A. Bjerhammar, *Application of calculus of matrices to method of least squares: With special reference to geodetic calculations*, Trans. Roy. Inst. Tech. Stockholm, no. 49 (1951), pp. 1–86.

[5] P. L. Bonate, *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*, Springer, New York, 2006.

[6] S. H. Cheng, N. J. Higham, C. S. Kenney, and A. J. Laub, *Approximating the logarithm of a matrix to specified accuracy*, SIAM J. Matrix Anal. Appl., 22 (2001), pp. 1112–1125, https://doi.org/10.1137/S0895479899364015.

[7] S. Chu, J. DeRisi, M. Eisen, J. Mulholland, D. Botstein, P. O. Brown, and I. Herskowitz, *The transcriptional program of sporulation in budding yeast*, Science, 282 (1998), pp. 699–705.

[8] W. J. Culver, *On the existence and uniqueness of the real logarithm of a matrix*, Proc. Amer. Math. Soc., 17 (1966), pp. 1146–1151.

[9] K. Godfrey, *Compartmental Models and Their Application*, Academic Press, London, 1983.

[10] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008, https://doi.org/10.1137/1.9780898717778.

[11] D. Hinrichsen and A. J. Pritchard, *Mathematical Systems Theory* I*, Modelling, State Space Analysis, Stability and Robustness*, Springer-Verlag, Berlin, 2005.

[12] N. S. Holter, A. Maritan, M. Cieplak, N. V. Fedoroff, and J. R. Banavar, *Dynamic modeling of gene expression data*, Proc. Natl. Acad. Sci. USA, 98 (2001), pp. 1693–1698.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1990.

[14] E. H. Moore, *On the reciprocal of the general algebraic matrix*, Bull. Amer. Math. Soc., 26 (1920), pp. 394–395.

[15] R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406–413.

[16] B. Singer and S. Spilerman, *The representation of social processes by Markov models*, Amer. J. Soc., 82 (1976), pp. 1–54.

[17] S. Stanhope, J. E. Rubin, and D. Swigon, *Identifiability of linear and linear-in-parameters dynamical systems from a single trajectory*, SIAM J. Appl. Dyn. Syst., 13 (2014), pp. 1792–1815, https://doi.org/10.1137/130937913.