Applied Mathematics and Computation 430 (2022) 127260

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Identifiability analysis of linear ordinary differential equation systems with a single trajectory☆

Xing Qiu[a], Tao Xu[b], Babak Soltanalizadeh[b], Hulin Wu, PhD[b,*]

[a] Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY, USA
[b] Department of Biostatistics & Data Science, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA

ARTICLE INFO

Article history:
Received 12 March 2021
Revised 28 February 2022
Accepted 16 May 2022
Available online 23 May 2022

Keywords:
Linear ordinary differential equations
Structural identifiability
Practical identifiability
Inverse problem
Parameter estimation

ABSTRACT

Ordinary differential equations (ODEs) are widely used to model dynamical behavior of systems. It is important to perform identifiability analysis prior to estimating unknown parameters in ODEs (a.k.a. inverse problem), because if a system is unidentifiable, the estimation procedure may fail or produce erroneous and misleading results. Although several qualitative identifiability measures have been proposed, much less effort has been given to developing *quantitative* (continuous) scores that are robust to uncertainties in the data, especially for those cases in which the data are presented as a single trajectory beginning with one initial value. In this paper, we first derived a closed-form representation of linear ODE systems that are not identifiable based on a single trajectory. This representation helps researchers design practical systems and choose the right prior structural information in practice. Next, we proposed several quantitative scores for identifiability analysis in practice. In simulation studies, the proposed measures outperformed the main competing method significantly, especially when noise was presented in the data. We also discussed the asymptotic properties of practical identifiability for high-dimensional ODE systems and conclude that, without additional prior information, many random ODE systems are practically unidentifiable when the dimension approaches infinity.

1. Background and introduction

Ordinary differential equations (ODE) can be used to model complex dynamic systems in a wide variety of disciplines including economics, physics, engineering, chemistry, and biology [1–10]. Such systems are usually represented as

$$\begin{cases} D\mathbf{x}(t) = f(\mathbf{x}(t), \mathbf{u}(t), \theta), \quad t \in (0, T), \\ \mathbf{x}(0) = \mathbf{x}_0. \end{cases} \tag{1}$$

Abbreviations: ODE, ordinary differential equation; GinOE, Ginibre ensemble; ICIS, initial condition-based identifiability score; SCN, smoothed condition number; PIS, practical identifiability score; REE, relative estimation error; AUC, area under the curve.

☆ This work was funded in part by the National Institutes of Health under award number R01 AI087135 and the University of Rochester CTSA award number UL1 TR002001 from the National Center for Advancing Translational Sciences of the National Institutes of Health.

* Corresponding author.
E-mail addresses: xing_qiu@urmc.rochester.edu (X. Qiu), tao.xu@uth.tmc.edu (T. Xu), babak.soltanalizadeh@uth.tmc.edu (B. Soltanalizadeh), hulin.wu@uth.tmc.edu (H. Wu).

https://doi.org/10.1016/j.amc.2022.127260
0096-3003/

$$\mathbf{y}(t) = h(\mathbf{x}(t), \mathbf{u}(t), \theta). \tag{2}$$

Here $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))' \in \mathbb{R}^d$ is the state vector, $D = \frac{\mathrm{d}}{\mathrm{d}t}$ is the first order differential operator[1], $\mathbf{y}(t) \in R^d$ is the output vector, $\mathbf{u}(t)$ is a known system input vector, f, g are known families of linear or nonlinear functions indexed by $\theta \in \mathbb{R}^p$, which is the vector of unknown parameters to be estimated. Equation (1) is called the state equation and Eq. (2) the output or observation equation. In this paper, we focus on an important special case of the above general ODE system: homogeneous linear ODE system with complete observation

$$D\mathbf{x}(t) = A\mathbf{x}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0 = (x_{1,0}, \ldots, x_{d,0})', \qquad \mathbf{y}(t) = \mathbf{x}(t). \tag{3}$$

Here $A \in M_{d\times d}$ is a matrix (called the system matrix) that characterizes the mechanistic relationship between $x_i(t)$; $\mathbf{y}(t) = \mathbf{x}(t)$ means that we can directly observe $\mathbf{x}(t)$ in all dimensions.

Let $\mathbf{x}(t|A, \mathbf{x}_0)$ be the solution curve (a.k.a. the trajectories) of Eq. (3) initiated at $\mathbf{x}_0$ and governed by system matrix A. It is well known that $\mathbf{x}(t|A, \mathbf{x}_0)$ can be represented as a unique matrix exponential

$$\mathbf{x}(t|A, \mathbf{x}_0) = e^{tA}\mathbf{x}_0. \tag{4}$$

As such, the *forward problem* of Eq. (3), defined as solving the ODE system with given A and $\mathbf{x}_0$, has been resolved in the mathematical sense – despite of several known numerical issues in matrix exponentials for high-dimensional data [11].

In practical applications, the parameters that characterize the ODE system, such as A and $\mathbf{x}_0$ in Eq. (3), must be estimated from the real data. This is known as the *inverse problem*. Over the years, many parameter estimation methods have been developed for ODE systems [2,6,7,12–17].

In principle, before performing the parameter estimation, we need to address an important question: are the parameters in a particular ODE model *identifiable* from the data? In this context, "identifiability" loosely means that there is a *unique* mapping between the trajectories and the parameters of a family of ODE systems.

By now, a rich literature on the identifiability of both linear and nonlinear ODE systems is available, see [18] for a thorough review of these methods. Unfortunately, most of them only consider the identifiability of the system matrix (A), and assume that one can choose an arbitrary initial condition $\mathbf{x}_0 \in \mathbb{R}^d$. For example, the *global identifiability* used in some literature on nonlinear ODE identifiability (e.g., [19]) reduces to the following definition for Eq. (3), as pointed out in [20]:

Definition 1.1. Linear ODE system (3) is globally identifiable in a subset $\Omega \subset M_{d\times d}$ iff for all $A, B \in \Omega$, $A \neq B$, there exists $\mathbf{x}_0 \in \mathbb{R}^d$, such that $\mathbf{x}(t|A, \mathbf{x}_0) \neq \mathbf{x}(t|B, \mathbf{x}_0)$.

However, such definition is of little use for linear ODE systems because Stanhope and colleagues proved in [20] that, due to the linearity of Model (3), such ODE system is *always globally identifiable* in the entire parameter space $M_{d\times d}$. Consequently, no computationally intensive symbolic computation on global identifiability is needed for linear ODE systems. The above definition of identifiability is also impractical because in many real world applications, data are only available in the form of one trajectory starting with a single $\mathbf{x}_0$. For example, influenza infection affects the state of transcriptome of a patient, which can be modeled by Eq. (3) [21–24]. However, it is currently impossible for a researcher to select an arbitrary $\mathbf{x}_0$ even in an animal study, because not only we do not have the technology to alter whole transcriptome globally, but also not all transcriptome states are biologically feasible. Furthermore, we cannot repeat the same $\mathbf{x}_0$ for a subject either, because the infection can have long-lasting effects to the immune system of that subject [25].

As a response to this weakness, several researchers developed a concept known as locally strong identifiability [26] or $\mathbf{x}_0$-identifiability [27], that involves data with only one trajectory. For Eq. (3), it can be stated as follows.

Definition 1.2. Eq. (3) is $\mathbf{x}_0$-identifiable w.r.t. a given $\mathbf{x}_0$ iff there exists an open and dense subset $\Omega \subset M_{d\times d}$, such that for all $A, B \in \Omega$, $A \neq B$, we have $\mathbf{x}(t|A, \mathbf{x}_0) \neq \mathbf{x}(t|B, \mathbf{x}_0)$ on $(0, \delta t)$, for some $0 < \delta t < T$.

Of note, the following natural extension to $\mathbf{x}_0$-identifiability was proposed in [27]:

Definition 1.3. System (3) is structurally identifiable iff there exist open and dense subsets $\Omega \subset M_{d\times d}$, $M^0 \in \mathbb{R}^d$, such that for all $A, B \in \Omega$, $A \neq B$ and all $\mathbf{x}_0 \in M^0$, we have $\mathbf{x}(t|A, \mathbf{x}_0) \neq \mathbf{x}(t|B, \mathbf{x}_0)$ on $(0, \delta t)$, for some $0 < \delta t < T$.

In other words, Definition 1.3 is $\mathbf{x}_0$-identifiability that applies to not one $\mathbf{x}_0$, but an open and dense set $M^0 \in \mathbb{R}^d$. This definition is consistent with the *structural identifiability* [28] and *geometrical identifiability* [26] for nonlinear ODE systems.

As a remark, the *open and dense* condition of Ω was designed to rule out a set of certain "inconvenient" parameters that has zero-measure. For example, it can be shown that if A has repeated eigenvalues, it is not identifiable with a class of other system matrices (see Section S5 and example S7 in Supplementary Text for more details). A workaround is to simply define Ω to be those matrices with no repeated eigenvalues, which is clearly a dense open set in $M_{d\times d}$. However, it is theoretically possible that the said open and dense set Ω may be "small" compared with $M_{d\times d}$ in terms of a measure such as $\lambda_{d\times d}$, the Lebesgue measure. See Example S9 in Supplementary Text, Section S8 for such an example. This can be seen as a weakness because in most real world applications, there is uncertainty in A, so we want to ensure that the identifiability

[1] To avoid confusion, we reserve symbol $'$ (apostrophe) for matrix transpose, not the derivative with respect to t.

applies to *almost every* $A \in M_{d\times d}$, not just a dense set with small measure or probability. As a concrete example, A may be modeled as $M + E$, where M is a deterministic matrix and E a perturbation term sampled from a random matrix distribution such as the real Ginibre ensemble (GinOE, [29]). By definition, if $E \sim$ GinOE, A_{ij} are *i.i.d.* standard normal random variables, therefore the probability measure associated with GinOE and the Lebesgue measure are absolutely continuous with each other. Therefore, the condition that almost every A is identifiable is equivalent to requiring $\lambda_{d\times d}(\Omega^c) = 0$, where $\Omega^c := M_{d\times d} \setminus \Omega$ is the complement of Ω, the collection of all identifiable A.

To the best of our knowledge, the most systematic study of linear ODE identifiability from a single observed trajectory is provided in [20]. In this seminal work, Stanhope and colleagues derived several necessary and sufficient conditions of identifiability that applies to a single trajectory. Specifically, they proposed two concepts, called "identifiability for a single trajectory" and "unconditional identifiability", defined as follows.

Definition 1.4 (identifiability for a single trajectory)**.** System (3) is identifiable for a single trajectory in $\Omega \subset M_{d\times d}$ and a given $\mathbf{x}_0$ iff for all $A, B \in \Omega$, $A \neq B$, we have $\mathbf{x}(t|A, \mathbf{x}_0) \neq \mathbf{x}(t|B, \mathbf{x}_0)$, for some $0 < t < T$.

Definition 1.5 (unconditional identifiability)**.** System (3) is unconditionally identifiable in $\Omega \in M_{d\times d}$ iff for all $A, B \in \Omega$, $A \neq B$ implies that for each nonzero $\mathbf{x}_0 \in \mathbb{R}^d$, $\mathbf{x}(t|A, \mathbf{x}_0) \neq \mathbf{x}(t|B, \mathbf{x}_0)$, for some $0 < t < T$.

Between these two definitions, Definition 1.5 adheres more to the traditional definition of structural identifiability for nonlinear ODE systems. Roughly speaking, it means that System (3) is identifiable from a single trajectory initiated from *every* $\mathbf{x}_0 \in \mathbb{R}^d$. Unfortunately, it is of little practical use for linear ODE system because no system satisfies this condition for an unconstrained parameter estimation problem, namely, $\Omega = M_{d\times d}$. In fact, we showed that (Supplementary Text, Section S1): (a) when the dimension d is odd, unconditional identifiability is not attainable for all $\Omega \subseteq M_{d\times d}$, and (b) when d is even, unconditional identifiability is not attainable for all $\Omega \subseteq M_{d\times d}$ such that $\lambda_{d\times d}\big(M_{d\times d} \setminus \Omega\big) = 0$. In summary, a large body of prior work in identifiability analysis are geared towards nonlinear ODEs with arbitrarily many observed trajectories, which is of little utility to linear ODEs, and this issue cannot be fixed by simply removing a zero-measure set from their definitions.

One major contribution of Stanhope and colleagues is that they established a beautiful connection between the algebraic and geometric aspects of linear ODE systems in [20, Theorem (3.4)]. We find it easier to state this important result by first define the following minimalist definition of identifiability.

Definition 1.6 (. $(A, \mathbf{x}_0)$-identifiability) For system (3), we call A is identifiable at $\mathbf{x}_0$ if for all $B \in M_{d\times d}$, $\mathbf{x}(t|A, \mathbf{x}_0) \neq \mathbf{x}(t|B, \mathbf{x}_0)$, for some $0 < t < T$.

Remarks 1. Definition 1.6 is not equivalent to Definition 1.4 applied to $\Omega := M_{d\times d}$ because in Definition 1.6, A is fixed and B is an arbitrary matrix in Ω, while in Definition 1.4, both A and B are arbitrary matrices in Ω. In short, Definition 1.6 is an intrinsic property of a *single* system, not a *collective* property of a set of system matrices.

Using this definition, [20, Theorem (3.4)] can be restated as follows: the $(A, \mathbf{x}_0)$-identifiability holds if and only if the solution curve $\mathbf{x}(t|A)$ is not contained in a *proper invariant subspace* of A. Based on this powerful theoretical result, they proposed to use $\kappa(X_1)$, the condition number of the matrix of a subset of discrete observations (see Section 3.3.1 for more details), to test the identifiability for discrete data with noise in practice.

However, their study is not without shortcomings. First, they did not derive the explicit structure of the largest subset $\Omega \subseteq M_{d\times d}$ for a give $\mathbf{x}_0$ in identifiability analysis for a single trajectory, nor the equivalent class of all $B \in M_{d\times d}$ such that $\mathbf{x}(t|B, t) = \mathbf{x}(t|A, t)$ when the system A is deemed unidentifiable at a given $\mathbf{x}_0$. Secondly, while using $\kappa(X_1)$ to check the practical identifiability of an ODE system is a clever heuristic, it has much room for improvement because: a) not all data are used in $\kappa(X_1)$, therefore it does not utilize data efficiently; b) measurement errors are not directly reflected in this score and there is no analysis of the asymptotic properties of $\kappa(X_1)$ from the statistical perspective; and c) by definition, $\kappa(X_1)$ depends on the availability of data at multiple time points, so it requires solving the ODE numerically in simulation studies, which can be time consuming for high-dimensional systems and/or when a large set of systems are considered.

In this study, we first derive a closed-form representation of $(A, \mathbf{x}_0)$-unidentifiable class, which is defined in Definition 2.1 as the collection of system matrices that are not identifiable for a given pair of A and $\mathbf{x}_0$. We also provide explicit structures of the equivalent class of unidentifiable systems due to repeated eigenvalues in A in Supplementary Text, Section S5. We believe these results will be valuable for future studies that combine *a priori* topological constraints (e.g., knowing which entries in A are zero in advance) and identifiability. In light this, we give a brief discussion of the best practice of using prior information to resolve the identifiability issues in Supplementary Text, Section S6. More systematic studies in this direction warrant a future study.

Secondly, we specify explicit, computable principles of $(A, \mathbf{x}_0)$-identifiability based on either $\mathbf{x}_0$ or the entire solution trajectory. These results are presented in our Theorems 2.4 and 3.2. To assist practical identifiability analyses, we propose three continuous scores: the initial condition-based identifiability score (ICIS, denoted as w_0^* in Eq. (10)), the smoothed condition number (SCN, denoted as τ in Eq. (32)), and the practical identifiability score (PIS, denoted as w^* in Eq. (37)), to solve the aforementioned problems. ICIS only uses A and $\mathbf{x}_0$, therefore it does not require numerically solving the ODE before the identifiability analysis. We think ICIS is most suitable for designing simulated ODE systems independent of a specific set of real data. SCN and PIS use data from all time points, which are more suitable for practical identifiability analysis with real

data. Using extensive simulation studies, we showed that SCN and PIS correlated with practical identifiability significantly better than $\kappa(X_1)$ when there was noise in the data.

In addition, we studied the asymptotic properties of practical identifiability for high-dimensional systems with randomly generated A and $\mathbf{x}_0$. We reached the following interesting conclusions: a) almost every system is $(A, \mathbf{x}_0)$-identifiable in the sense that ICIS > 0; and b) when $d \to \infty$, almost all systems are practically unidentifiable in the sense that ICIS $\to 0$. These two seemly contradictory conclusions suggest that the practical identifiability of high-dimensional ODE systems is very different from that of low-dimensional systems, and classical mathematical identifiability analyses are insufficient for analyzing *high-dimensional* real world applications. The focus must be shifted towards practical identifiability analyses characterized by *continuous* scores, especially with the considerations from the stochastic perspective.

Last but not the least, we provide a user-friendly R package `ode.ident`, with full documentation and examples, so practitioners with minimum programming skills can analyze the identifiability of linear ODE systems. This R package is available at https://github.com/qiuxing/ode.ident.

2. $(A, \mathbf{x}_0)$-identifiability

In this section, we focus on the mathematical inverse problem for one fully observed trajectory. Namely, we assume that we have the complete observation of *one* solution curve $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))' \in \mathbb{R}^d)$ governed by Eq. (3) and its derivative on $[0, T]$, with no measurement error.

First, let us define the $(A, \mathbf{x}_0)$-unidentifiable class as follows.

Definition 2.1 (. $(A, \mathbf{x}_0)$-unidentifiable class) For a given system matrix $A \in M_{d\times d}$ and initial condition $\mathbf{x}_0 \in \mathbb{R}^d$, the $(A, \mathbf{x}_0)$-unidentifiable class, denoted by $[A]_{\mathbf{x}_0}$, is a subset of matrices in $M_{d\times d}$ such that

$$B \in [A]_{\mathbf{x}_0} \quad \text{iff} \quad \mathbf{x}(t|A, \mathbf{x}_0) = \mathbf{x}(t|B, \mathbf{x}_0). \tag{5}$$

In other words, two system matrices A, B are in the same unidentifiable class if and only if they produce the same solution trajectory at $\mathbf{x}_0$.

The overarching goal of this section is to understand the structure of $[A]_{\mathbf{x}_0}$, and the conditions under which this class contains only one member, therefore A can be uniquely determined by the trajectory $\mathbf{x}(t|A, \mathbf{x}_0)$. To this end, we need to introduce an important geometric concept called invariant subspace, which is a generalization of eigenvectors, and its connection to the Jordan decomposition of A in Section 2.1.[2]

2.1. Jordan decomposition and invariant subspaces

Definition 2.2 (Invariant subspace)**.** An invariant subspace of a square matrix $A_{d\times d}$ is a linear subspace $L \subseteq R^d$ such that for all $\mathbf{x} \in L$, $A\mathbf{x} \in L$. We say L is a *proper* invariant subspace if $L \neq \mathbb{R}^d$.

By definition, we see that if a vector $\mathbf{x}$ is in a proper invariant subspace L of A, $A\mathbf{x}$ must also stay in L. Using mathematical induction, we see that $A^n\mathbf{x} \in L$ for every positive integer n. With a little more work, it can be proven that $e^{tA}\mathbf{x} \in L$ for $t \in [0, T]$, where e^{tA} is the matrix exponential of tA.

The following proposition states that the intersection and linear span (the combination) of two invariant subspaces are invariant subspaces.

Proposition 2.1. *If L_1 and L_2 are invariant subspaces of A, then*

1. *$L_1 \cap L_2$ is an invariant subspace of A;*
2. *$\text{span}(L_1, L_2)$ is an invariant subspace of A.*

In other words, the collection of invariant subspaces of A forms a lattice.

Based on random matrix theory [29,31,32], we know that almost every (w.r.t. the Lebesgue measure on $M_{d\times d}$) $A \in M_{d\times d}$ has d distinct eigenvalues. This conclusion also holds for probability measures associated with most random matrix ensembles such as Ginibre ensemble, Gaussian orthogonal ensemble, Wishart ensemble, etc. [32].

[2] These concepts and results can be found in many graduate level matrix analysis textbooks, e.g., [30].

Consequently, almost every $A \in M_{d\times d}$ has the following Jordan decomposition

$$A = Q\Lambda Q^{-1}, \qquad \Lambda = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_K \end{pmatrix}, \qquad Q = \begin{pmatrix} [c|c|c|c]Q_1 & Q_2 & \dots & Q_K \end{pmatrix}.$$
$$J_k = \begin{cases} c_k, & k = 1, 2, \dots, K_1, \\ \begin{pmatrix} a_k & -b_k \\ b_k & a_k \end{pmatrix}, & k = K_1+1, K_1+2, \dots, K. \end{cases} \tag{6}$$
$$\dim Q_k = \begin{cases} 1, & k = 1, \dots, K_1, \\ 2, & k = K_1+1, \dots, K. \end{cases}$$

In other words, A can be decomposed into $K = K_1 + K_2$ Jordan blocks, the first K_1 such blocks are 1×1 blocks corresponding with real eigenvalues (those c_k in Eq. (6)); and the rest K_2 blocks are 2×2 blocks corresponding with complex eigenvalues $a_k \pm b_k i$. There is a corresponding column-wise decomposition of matrix Q, such that each Q_k contains: (a) a single column vector of Q which is the eigenvector of c_k, or (b) two column vectors in Q such that $Q_k := (\mathbf{v}_{k1}|\mathbf{v}_{k2})$, which are the "eigenvectors" associated with $a_k \pm b_k i$. To avoid identifiability issues, all column vectors in Q are assumed to have unit Euclidean length.

Note that the word "eigenvector" in case (b) refers to a generalization of true eigenvectors. In fact, those 2×2 Jordan blocks do not have real eigenvectors; instead, each of them is associated with a 2-dimensional invariant subspace of A and $Q_k = (\mathbf{v}_{k1}|\mathbf{v}_{k2})$ is a basis of this 2-dimensional invariant subspace.

We would like to point out that based on simple enumeration of dimensions, we have $d = K_1 + 2K_2$, and

$$Q_k = \begin{cases} Q_k, & k = 1, 2, \dots, K_1, \\ (\mathbf{v}_{k1}|\mathbf{v}_{k2}), \quad \mathbf{v}_{k1} = Q_{2k-K_1-1},\ \mathbf{v}_{k2} = Q_{2k-K_1}, & k = K_1+1, \dots, K. \end{cases} \tag{7}$$

In other words, $\mathbf{v}_{k1}$ and $\mathbf{v}_{k2}$ in Q_k are the $(2k-K_1-1)$th and $(2k-K_1)$th column vectors of Q, respectively. For convenience, we define the following correspondences between i (the original dimension in J) and k (the index of invariant subspaces):

$$i(k) := 2k - K_1 - 1, \qquad k(i) := \begin{cases} i, & i = 1, 2, \dots, K_1, \\ K_1 + \lceil \frac{i-K_1}{2} \rceil, & i = K_1+1, \dots, K. \end{cases} \tag{8}$$

Using the above notation, $\mathbf{v}_{k1} = Q_{i(k)},\ \mathbf{v}_{k2} = Q_{i(k)+1}$.

Theorem 2.1. *Let $L_k := \mathrm{span}(Q_k)$. Each L_k is an invariant subspace of A. Furthermore, if L is an invariant subspace of A, it can always be decomposed as*

$$L = \mathrm{span} \bigcup_{i\in S} L_k, \qquad S \subseteq \{1, 2, \dots, K\}. \tag{9}$$

From now on, we assume that A has d distinct eigenvalues and can be decomposed as $A = Q\Lambda Q^{-1}$ in Eq. (6). The case in which A has repeated eigenvalues will be discussed in Supplementary Text, Section S5. For convenience, we will also denote $L_0 := \{0_d\}$, the trivial proper invariant subspace of A that contains only the origin.

2.2. Initial condition-based identifiability score (ICIS)

One of our main conclusion is that the $(A, \mathbf{x}_0)$-identifiability defined in Definition 1.6 can be determined by the initial condition-based identifiability score (ICIS) defined as follows.

Definition 2.3. Let $\mathbf{x}_0 \in \mathbb{R}^d$ and $\tilde{\mathbf{x}}_0 := Q^{-1}\mathbf{x}_0 \in \mathbb{R}^d$. We define the w_0^* statistic in the following equation as the Initial Condition-based Identifiability Score (ICIS):

$$w_{0,k} := \begin{cases} \tilde{\mathbf{x}}_{0,k} \in \mathbb{R}^1, & k = 1, 2, \dots, K_1, \\ (\tilde{\mathbf{x}}_{0,i(k)}, \tilde{\mathbf{x}}_{0,i(k)+1})' \in \mathbb{R}^2, & k = K_1+1, \dots, K. \end{cases} \tag{10}$$
$$w_0^* := \min_k |w_{0,k}|.$$

Here $|w_{0,k}|$ is the absolute value of $w_{0,k}$ for $k = 1, 2, \dots, K_1$, and the Euclidean norm of $w_{0,k}$ for $k = K_1+1, K_1+2, \dots, K$.

From the geometric perspective, $\mathbf{x}_0$ can be decomposed into a linear combination (oblique projections) of Q_k, and $w_{0,k}$ are the linear coefficients of such a decomposition

$$\mathbf{x}_0 = Q\tilde{\mathbf{x}}_0 = \sum_{k=1}^{K_1} w_{0,k} Q_k. \tag{11}$$

Heuristically speaking, if $w_{0,k} = 0$ (in $\mathbb{R}^1$ or $\mathbb{R}^2$), $\mathbf{x}_0$ does not contain any information from L_k. This is because in this case, $\mathbf{x}_0 \in L_{-k}$, where L_{-k} is the invariant subspace of A that *excludes* L_k, which implies that the *entire trajectory*, $\mathbf{x}(t|A, \mathbf{x}_0)$, is in L_{-k} (see the discussion in the beginning of Section 2.1). These ideas are summarized in Lemma 2.2 below.

Lemma 2.2. *The following two statements are equivalent*

1. *There exists a proper invariant subspace $L \subsetneq \mathbb{R}^d$ of A, such that $\mathbf{x}_0 \in L$.*
2. *There exists $k \in \{1, 2, \ldots, K\}$, such that $|w_{0,k}| = 0$, or equivalently, $w_0^* = 0$.*

Now we are ready to present the following theorem:

Theorem 2.3 (Computational criterion for. $(A, \mathbf{x}_0)$-*identifiability*) *Assuming that an ODE system A has d distinct eigenvalues. This system is identifiable at $\mathbf{x}_0$ if and only if the ICIS is nonzero.*

Proof. Based on Lemma 3.2 and Theorem (3.4) in [20], we know that system A is identifiable at $\mathbf{x}_0$ if and only if $\mathbf{x}_0$ is not contained in a proper invariant subspace of A, which is equivalent to $w_0^* \neq 0$ based on our Lemma 2.2. □

Theorem 2.3 implies that, when A is not identifiable at $\mathbf{x}_0$, there must be a nonempty subset $S_0 \subset \{1, 2, \ldots, K\}$ such that $|w_{0,k}| = 0$, for $k \in S_0$. WLOG, we assume that $\mathbf{x}_0 \neq 0_d$, so its complement set $S_+ := \{1, 2, \ldots, K\} \setminus S_0$ must be nonempty. By construction, $L_+ := \operatorname{span} \bigcup_{k \in S_+} L_k$ is a proper invariant subspace, and $\mathbf{x}_0 \in L_+$.

The two index sets S_0 and S_+ induce the following diagonal binary matrices

$$I_0 := \operatorname{diag}(a_i), \quad a_i = \begin{cases} 1, & k(i) \in S_0, \\ 0, & k(i) \neq S_0. \end{cases} \qquad I_+ := I_{d\times d} - I_0. \tag{12}$$

They can be used to construct the following decomposition of the eigenvector matrix Q and the Jordan block matrix J

$$\begin{aligned} Q_0 &:= QI_0, \qquad Q_+ := QI_+, \qquad Q = Q_0 + Q_+. \\ J_0 &:= I_0 J I_0 = J I_0, \qquad J_+ := I_+ J I_+ = J I_+, \qquad J = J_0 + J_+. \end{aligned} \tag{13}$$

Intuitively, Q_0 and J_0 replace column vectors in Q and blocks in J into zeros if they belong to L_+. Q_+ and J_+ are defined in exactly the opposite way.

Using this notation, we describe the explicit structure of $(A, \mathbf{x}_0)$-unidentifiable class in the following Theorem.

Theorem 2.4 (Structure of the unidentifiable classes)**.** *$(A, \mathbf{x}_0)$-unidentifiable class has the following explicit structure*

$$[A]_{\mathbf{x}_0} \quad = Q(J_+ + I_0(J_0 + D)I_0)Q^{-1} = A + Q(I_0 D I_0)Q^{-1}, \qquad D \in M_{d\times d}. \tag{14}$$

In other words, two matrices A_1 and A_2 are in the same $(A, \mathbf{x}_0)$-unidentifiable class (written as $A_1 \sim A_2$) iff there exist $D \in M_{d\times d}$, such that $A_1 - A_2 = Q(I_0 D I_0)Q^{-1}$.

Proof. See Section S4, Supplementary Text. □

Remarks 2. The degrees of freedom in $[A]_{\mathbf{x}_0}$ is controlled by $D_0 := I_0 D I_0$, which has d_0^2 degrees of freedom (not d^2). This is because by construction, D_0 is a sparse matrix such that its ijth element satisifies

$$D_{0,ij} = 0, \qquad \text{if } k(i), k(j) \in S_0. \tag{15}$$

2.3. A 3-Dimensional example

Example 1. In this example, the system matrix A and its Jordan canonical form are given as follows:

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 & -1 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \end{pmatrix} = QJQ^{-1}, \qquad J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1/2 & \sqrt{7}/2 \\ 0 & -\sqrt{7}/2 & 1/2 \end{pmatrix}. \\ Q &\approx \begin{pmatrix} 0.408 & 0 & 0.316 \\ -0.816 & 0.418 & 0.158 \\ -0.408 & 0.837 & 0 \end{pmatrix}, \qquad Q^{-1} \approx \begin{pmatrix} 0.612 & -1.225 & 0.612 \\ 0.299 & -0.598 & 1.494 \\ 2.372 & 1.581 & -0.791 \end{pmatrix}. \end{aligned} \tag{16}$$

Based on earlier discussions, A has two proper invariant subspaces. $L_1 := \operatorname{span}(Q_1)$ is a one-dimensional space corresponding with the real eigenvalue $\lambda_1 = -1$, and $L_2 := \operatorname{span}(Q_2, Q_3)$ is a two-dimensional space corresponding with $\lambda_2, \lambda_3 = 1/2 \pm \sqrt{7}i/2$. Here Q_j is the jth column vector of matrix Q.

Let us consider the following two initial conditions:

$$\mathbf{x}_0^{(a)} := Q\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0.816 \\ -2.051 \\ -1.653 \end{pmatrix}, \qquad \mathbf{x}_0^{(b)} := Q\begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix} \approx \begin{pmatrix} 0.949 \\ -0.362 \\ -1.673 \end{pmatrix}.$$

Notice that $\mathbf{x}_0^{(a)}$ contains information from both L_1 and L_2, but $\mathbf{x}_0^{(b)}$ only contains information from L_2. Based on earlier discussions, A is identifiable at $\mathbf{x}_0^{(a)}$ but not $\mathbf{x}_0^{(b)}$. Using Eq. (14), the $(A, \mathbf{x}_0)$-unidentifiable class in the latter case can be represented as follows.

$$\begin{aligned}
&I_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
&I_0 D I_0 = \begin{pmatrix} b & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad b \in \mathbb{R}. \\
&J_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & \sqrt{7}/2 \\ 0 & -\sqrt{7}/2 & 1/2 \end{pmatrix}, \qquad Q(I_0 D I_0)Q^{-1} = bQ_0 Q^{-1}, \\
&[A]_{\mathbf{x}_0^{(b)}} = QJ_+Q^{-1} + bQ_0Q^{-1} \\
&\qquad = \frac{1}{4}\begin{pmatrix} 1 & 2 & -3 \\ 6 & 4 & -2 \\ 11 & 6 & -1 \end{pmatrix} + \frac{b}{4}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{pmatrix}.
\end{aligned} \tag{17}$$

Here $b \in \mathbb{R}$ is an arbitrary parameter, and the $(A, \mathbf{x}_0)$-unidentifiable class is characterized by b times a *full matrix*, therefore different choices of b affects all nine elements in A. Consequently, we cannot determine the value of *any* entry in A without additional information. In fact, we cannot even determine whether a particular A_{ij} is zero or nonzero, which is sometimes referred to as the network topology of A. For example, if we set $b = -1$, we get the original system A specified in Eq. (16) with four zero entries. When we set $b = 3$, we obtain an equivalent matrix with completely different topology and values than A

$$\tilde{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 4 & -2 \\ 2 & 3 & -1 \end{pmatrix}. \tag{18}$$

It is easy to check that:

$$e^{t\tilde{A}}\mathbf{x}_0^{(a)} \neq e^{tA}\mathbf{x}_0^{(a)}, \qquad e^{t\tilde{A}}\mathbf{x}_0^{(b)} = e^{tA}\mathbf{x}_0^{(b)}.$$

Before we move on to the next topic (practical identifiability), we would like to present two auxiliary results that are useful in practice.

1. In Supplementary Text, Section S5, we provide a detailed analysis of identifiability issues induced by repeated eigenvalues in A, and provided the closed-form structure of unidentifiable class for these matrices in Eq. (S.23). Based on these results, we recommend researchers avoid systems that have nearly identical eigenvalues in designing of simulation studies.
2. in Supplementary Text, Section S6, we show that while it is possible to use prior information in the form of structural constraints to resolve the identifiability issue of a problematic system $(A, \mathbf{x}_0)$, such constraints must be compatible with the said system, otherwise: (a) the system may still suffer from the identifiability issue; or (b) under these constraints, no system can generate the observed solution curve.

3. Data-based identifiability scores

We now focus on the following practical problem: to quantify the $(A, \mathbf{x}_0)$-identifiability from imperfect observations in real world applications. To this end, we assume that the observed data is a set of discrete and noisy observations on a time grid $\{t_1, \ldots, t_j\}$:

$$y_{ij} := x_i(t_j) + \epsilon_{ij}, \qquad \epsilon_{ij} \sim F_\epsilon, \qquad i = 1, 2, \ldots, d, \quad j = 1, 2, \ldots, n. \tag{19}$$

In the above equation, F_ϵ, the probability distribution of measurement error, is assumed to be absolutely continuous w.r.t. the Lebesgue measure on $M_{d\times J}$. For convenience, we will also use collective notations $X = \{x_i(t_j)\} \in M_{d\times n}$, $Y = \{y_{ij}\} \in M_{d\times n}$, and $\boldsymbol{\epsilon} = \{\epsilon_{ij}\} \in M_{d\times n}$. With these matrix notations, Eq. (19) can be simplified as $Y = X + \boldsymbol{\epsilon}$.

3.1. Minimal signals for reconstructing A

In this section, we demonstrate that even for a theoretically identifiable system, if the "signal" in a subspace is too small, we are still not able to reconstruct A in practice.

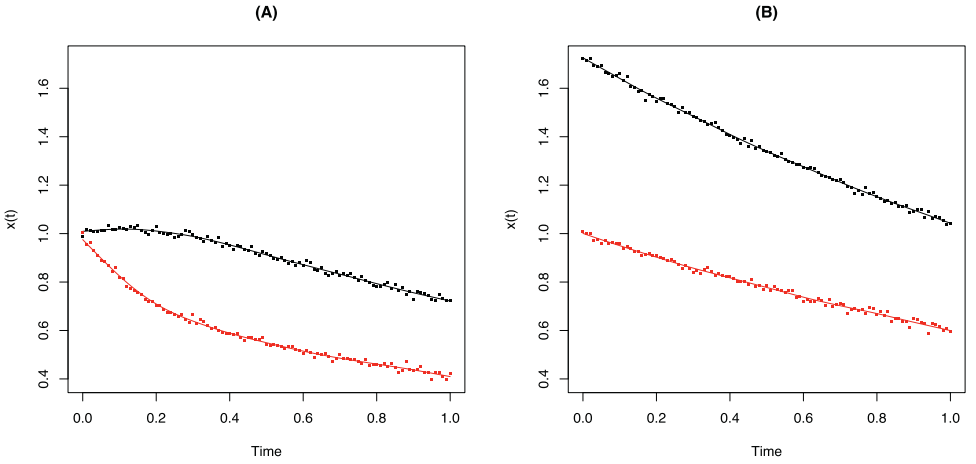

Fig. 1. An illustration of the fitted solution trajectories in Example 2. The black curve is $\hat{x}_1(t)$ and the red curve is $\hat{x}_2(t)$ in both sub-figures. The discrete data (illustrated by dots in this figure) are observed on a total number of $n = 101$ time points evenly assigned on [0,1]. A small normal measurement error, $\epsilon_{ij} \sim N(0, 0.01^2)$, is added to each observation. (A): $\mathbf{x}_0^{(A)} = (1, 1)'$. (B): $\mathbf{x}_0^{(B)} = (1.72, 1)'$. Case (A) is practically identifiable with the functional two-stage method while Case (B) is not. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Example 2. We consider a two-dimensional system

$$\begin{aligned} A &= Q\Lambda Q^{-1} = \begin{pmatrix} -1.875 & 2.382 \\ 2.382 & -4.625 \end{pmatrix}, \\ Q &= \mathrm{Rot}_{\pi/6} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}, \quad Q^{-1} = \mathrm{Rot}_{-\pi/6} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} -1/2 & 0 \\ 0 & -6 \end{pmatrix}. \end{aligned} \tag{20}$$

We generate the solution curve $\mathbf{x}(t)$ from this system and record its values $x_i(t_j)$ at $n = 101$ equally spaced time points on [0,1], $t_1 = 0, t_2 = 0.01, \ldots, t_{101} = 1$. A small normal measurement error, $\epsilon_{ij} \sim N(0, 0.01^2)$, is added to each observation.

It is easy to see that A has two one-dimensional proper invariant subspaces, $L_1 = \mathrm{span}((\frac{\sqrt{3}}{2}, \frac{1}{2})')$, $L_2 = \mathrm{span}((-\frac{1}{2}, \frac{\sqrt{3}}{2})')$. Let us consider two initial conditions

$$\mathbf{x}_0^{(A)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{x}_0^{(B)} = \begin{pmatrix} 1.72 \\ 1 \end{pmatrix}.$$

It can be shown that A is identifiable at both $\mathbf{x}_0^{(A)}$ and $\mathbf{x}_0^{(B)}$. Of note, we would like to mention that this analysis can be done by applying the `ICISAnalysis()` function in our R package.

Using the functional two-stage method (see Section S7.2, Supplementary Text), we are able to estimate A and produce two fitted curves for both cases. The fitted curves, denoted by $(\hat{x}_1(t), \hat{x}_2(t))'$, look reasonable in Fig. 1 for both cases.

However, the two reconstructed system matrices are quite different:

$$\begin{aligned} \hat{A}^{(A)} &= \begin{pmatrix} -1.58 & 1.93 \\ 1.93 & -3.91 \end{pmatrix}, \qquad \|\hat{A}^{(A)} - A\|_F^2 = 1.01, \\ \hat{A}^{(B)} &= \begin{pmatrix} -0.69 & 0.33 \\ -1.81 & 2.62 \end{pmatrix}, \qquad \|\hat{A}^{(B)} - A\|_F^2 = 75.65. \end{aligned}$$

Here $\|\cdot\|_F$ is the Frobenius norm of a matrix. It is clear that A is practically identifiable at $\mathbf{x}_0^{(A)}$ only, not at $\mathbf{x}_0^{(B)}$. From this example, we see that even when the ODE system is mathematical identifiable, its practical identifiability may still be an issue.

In-depth analysis shows that the unidentifiability issue in case (B) is due to the fact that $\mathbf{x}_0^{(B)}$ is "almost" contained in L_1, so that L_2 had only a tiny bit of information. Denote the basis in L_1 and L_2 as $Q_1 = (\frac{\sqrt{3}}{2}, \frac{1}{2})'$ and $Q_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})'$, respectively. We have

$$\begin{aligned} w_{0,1}^{(A)} &= \langle \mathbf{x}_0^{(A)}, Q_1 \rangle = 1.366, \qquad \mathrm{ICIS}^{(A)} = w_{0,2}^{(A)} = \langle \mathbf{x}_0^{(A)}, Q_2 \rangle = 0.366. \\ w_{0,1}^{(B)} &= \langle \mathbf{x}_0^{(B)}, Q_1 \rangle = 1.990, \qquad \mathrm{ICIS}^{(B)} = w_{0,2}^{(B)} = \langle \mathbf{x}_0^{(B)}, Q_2 \rangle = \mathbf{0.006}. \end{aligned}$$

Based on the above analysis, it is easy to see that the small value of $\mathrm{ICIS}^{(B)} = 0.006$ causes the numerical problem in estimating A in case (B).

3.2. Identifiability for high-dimensional systems

Based on Theorem 2.3, A is identifiable at $\mathbf{x}_0$ if and only if $\mathbf{x}_0$ is not located in a proper invariant subspace of A. Because there are only finitely many ($2^K - 1$ of them, to be more precise) proper invariant subspaces of A, and each of them has dimension strictly less than d (the "proper" part of the definition), the union of all proper invariant subspaces is only a zero-measure set of $\mathbb{R}^d$. In this regard, as long as A does not have repeated eigenvalues (which is true for almost every $A \in M_{d\times d}$), A is mathematically identifiable at almost every $\mathbf{x}_0 \in \mathbb{R}^d$. This fact is probably the main reason why not many mathematicians have paid much attention to the identifiability problem of linear ODE systems.

However, as is shown in Example 2, to have a reliable estimate of A requires more than just a *qualitative* statement that $\mathbf{x}_0$ does not lie in any proper invariant subspace of A. We need to ensure that when we decompose $\mathbf{x}_0$ into a linear combination of components from L_k, each one of them has *enough* information, so that we can reconstruct the corresponding sub-system on L_k with noisy observations. This is the main motivation for us to propose ICIS, a *quantitative* measure of identifiability.

Knowing that the collection of all proper invariant subspaces have measure zero in $\mathbb{R}^d$, the readers may think that while practical identifiability issues do exist, they must be rare in practice. Unfortunately, these issues are not that unusual when d is large, in which case those practically identifiable systems are the exceptions instead. In Supplementary Text, Section S2, we proved that a large class of symmetric random ODE systems are practically unidentifiable when $d \to \infty$, as stated in the following theorem

Theorem 3.1. *Let us assume that:*

(a) *The system matrix $A \in M_{d\times d}$ is sampled from a symmetric, real-valued random matrix ensemble with probability measure $p(A)$ that is statistically invariant to orthogonal transformations, namely,*

$$p(A) = p(TAT'), \qquad \forall T \in O(d). \tag{21}$$

(b) *The initial condition $\mathbf{x}_0 \in \mathbb{R}^d$ is sampled from a random distribution that is independent of A and satisfies*

$$\lim_{d\to\infty} \frac{E|\mathbf{x}_0|^2}{d^3} = 0. \tag{22}$$

Based on the above two assumptions, the ICIS converges to zero in L^2, namely,

$$E(w_0^*(A, \mathbf{x}_0))^2 \longrightarrow 0, \quad \text{when} d \to \infty. \tag{23}$$

Proof. The proof is provided in Section S2, Supplementary Text. □

Remarks 3. Perhaps the most well known random matrix ensemble that satisfies Assumption (a) is the Gaussian Orthogonal Ensemble (GOE, [32]). Many other ensembles also satisfy this condition, such as the Wishart ensemble, Jacobi orthogonal ensemble, etc. In fact, according to Weyl's lemma [33,34], a random matrix ensemble is orthogonally invariant as long as its distribution function has the following trace representation

$$p(H) = \phi\left(\mathrm{tr}H, \mathrm{tr}H^2, \ldots, \mathrm{tr}H^n\right). \tag{24}$$

Assumption (b) is a very weak condition that should be satisfied in almost all practical applications. If $\mathbf{x}_{0,i}$ has finite second order moments, and

$$\sup_i Ex_{0,i}^2 = \mu_2 < \infty,$$

we have

$$E|\mathbf{x}_0|^2 := \sum_{i=1}^{d} Ex_{0,i}^2 \leqslant d \cdot \mu_2 = O(d^1), \qquad \frac{E|\mathbf{x}_0|^2}{d^3} = O(d^{-2}) \to 0.$$

In this case, $x_{0,i}$ do not have to be independent nor identically distributed.

We use the following simple and concrete example to illustrate the issue of practical identifiability described by Theorem 3.1.

Example 3. Let $d = 100$ and assume that A is an arbitrary diagonal matrix in $M_{d\times d}$. By construction, all eigenvalues are real and $Q = I_d$, therefore $w_{0,i} = x_{0,i}$, $i = 1, 2, \ldots, d$. Let $\mathbf{x}_0 \sim N(0_d, I_d)$, in other words, $x_{0,i}$ are generated from *i.i.d.* $N(0, 1)$. For simplicity, we write $r_i = |w_{0,i}|$. Apparently, r_i are standard half normals with relatively large expectations $Er_i = \sqrt{\frac{2}{\pi}} \approx 0.8$. This fact seems to suggest that, as long as the measurement error is small ($\sigma \ll 0.8$), we would have enough information to reconstruct A.

In reality, the *smallest* member of r_i (denoted by $r_{(1)}$), has a distribution that is statistically much smaller than a standard half normal. Using numerical integration, we found that $Er_{(1)} \approx 0.012$, which is 66 times smaller than Er_i. Based on the lessons we learned from Example 2, we anticipate that it is almost impossible to estimate A accurately in this case.

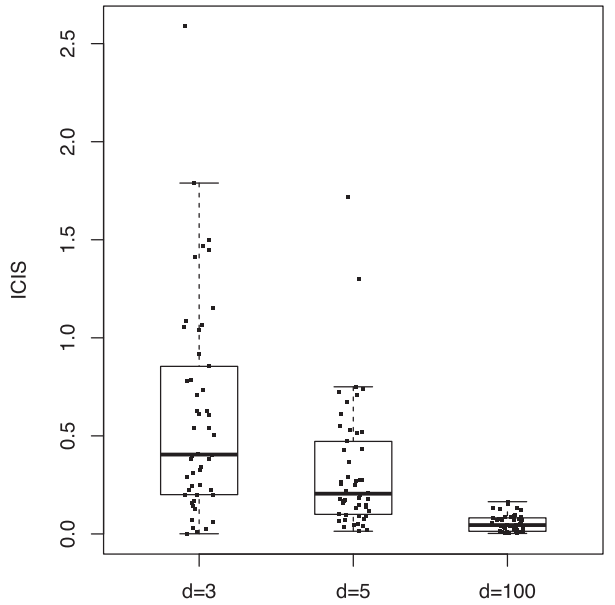

Fig. 2. Larger dimension is associated with smaller ICIS (w_0^*), which implies that high-dimensional ODE systems are more difficult to be reconstructed numerically.

Finally, we conduct a mini-simulation to illustrate that ODEs with random *asymmetric* system matrices also suffer from the identifiability issues stated in Theorem 3.1. Specifically, we randomly generate 50 A from the standard GinOE with $d = 3, 5, 100$ dimensions, and pair them with 50 $\mathbf{x}_0$ sampled from $N(0_d, I_d)$. We compute the ICIS for those $(A, \mathbf{x}_0)$ and plot them in Fig. 2. We see that larger dimension is associated with smaller ICIS, which implies that these ODE systems are more difficult to be numerically identified.

In light of the above discussions, to design a well-behaved, identifiable high-dimensional ODE system in simulations, one needs to ensure that:

1. A has nice mathematical properties, such as distinct eigenvalues; stability; no high-frequency components, etc.
2. $\mathbf{x}_0$ should not be "randomly" generated; instead, it should be generated in a way such that ICIS$(A, \mathbf{x}_0)$ is not too small.

3.3. Practical identifiability score (PIS)

Recall that ICIS does not depend on the full trajectory $\mathbf{x}(t)$, therefore it is most useful in designing simulation studies. In real world applications, it is preferable to define identifiability scores that use the entirety of Y (the discrete data measured at all time points) to quantify the practical identifiability of the system. Let $\hat{A}(Y)$ be an estimator of A given the discrete observation. One way to quantify the practical identifiability is to use the numerical sensitivity of $\hat{A}$, which can be defined as the mean squared error, MSE $:= E\|\hat{A} - A\|_F^2$. Unfortunately, there are many different ways to estimate A, thus it is impossible to develop a *universal* quantity that works for all estimators. In this section, we propose two scores based on a class of the two-stage methods, under the assumption that σ^2 is small enough so that $\hat{A} - A$ is small. These proposed scores are compared to $\kappa(Y_1)$, the practical identifiability measure proposed by Stanhope and colleagues in [20] in our simulation studies.

3.3.1. Stanhope's condition number

Stanhope and colleagues proposed to use $\kappa(Y_1)$ as a practical identifiability measure in [20]. Here $Y_1 := [Y_{\cdot 1}|Y_{\cdot 2}|\dots|Y_{\cdot d}] \in M_{d\times d}$ is the matrix of discrete data evaluated at the first d time points; $\kappa(Y_1) := \|Y_1\|_F\|Y_1^{-1}\|_F$ is the *condition number* of Y_1, which is a quantitative measure of numerical stability of Y_1^{-1}. It was motivated by the fact that if Y_1 is invertible and there is no noise in the discrete observations,

$$e^{\Delta t A} = Y_2 Y_1^{-1}, \qquad Y_2 := [Y_{\cdot 2}|Y_{\cdot 3}|\dots|Y_{\cdot d+1}] \in M_{d\times d}. \tag{25}$$

3.3.2. Functional two-stage methods

Definition 3.1 (pairwise. L^2-inner product matrix) Let $\mathbf{x}(t) \in \mathbb{R}^d$ and $\mathbf{y}(t) \in \mathbb{R}^{d'}$ be two multidimensional functions defined on $[0, T]$. We use notation $\langle \mathbf{x}(t), \mathbf{y}(t)\rangle$ to refer to the pairwise L^2-inner product matrix between them, which is a $d \times d'$

matrix in which

$$\langle \mathbf{x}(t),\ \mathbf{y}(t)\rangle_{i,j} := \langle x_i(t),\ y_j(t)\rangle_{L^2} = \int_0^T x_i(t)y_j(t)\mathrm{d}t.$$

For convenience, we denote $\langle \mathbf{x}(t),\ \mathbf{x}(t)\rangle$ by $\Sigma_{\mathbf{xx}}$ when there is no confusion. Obviously, $\Sigma_{\mathbf{xx}}$ is a symmetric positive semi-definite matrix. It is singular if and only if zero is one of its eigenvalues, in which case we know

$$\mathbf{v}_0'\Sigma_{\mathbf{xx}}\mathbf{v}_0 = 0, \qquad \mathbf{v}_0 \in \mathbb{R}^d, \quad \mathbf{v}_0 \neq 0_d. \tag{26}$$

where $\mathbf{v}_0$ is an eigenvector associated with the zero eigenvalue.

Theorem 3.2. *Let $\mathbf{x}(t)$ be an observed solution trajectory governed by ODE system A and initiated at $\mathbf{x}_0 := \mathbf{x}(0)$. This ODE system is identifiable at $\mathbf{x}_0$ if and only if the pairwise inner product matrix $\Sigma_{\mathbf{xx}}$ is nonsingular (invertible).*

Proof. See Section S4, Supplementary Text. □

Notice that the ODE $D\mathbf{x}(t) = A\mathbf{x}(t)$ implies

$$\langle D\mathbf{x}(t),\ \mathbf{x}(t)\rangle = A\langle \mathbf{x}(t),\ \mathbf{x}(t)\rangle \Longrightarrow A = \Sigma_{D\mathbf{x},\mathbf{x}} \cdot \Sigma_{\mathbf{xx}}^{-1}. \tag{27}$$

Here $\Sigma_{D\mathbf{x},\mathbf{x}} := \langle D\mathbf{x}(t),\ \mathbf{x}(t)\rangle$ is the pairwise inner product matrix between $D\mathbf{x}(t)$ and $\mathbf{x}(t)$. This fact motivated the following two-stage methods, which is a class of estimators of A:

$$\hat{A} := \hat{\Sigma}_{D\mathbf{x},\mathbf{x}} \cdot \hat{\Sigma}_{\mathbf{xx}}^{-1}. \tag{28}$$

Here $\hat{\Sigma}_{D\mathbf{x},\mathbf{x}}$ and $\hat{\Sigma}_{\mathbf{xx}}$ are estimates of $\Sigma_{D\mathbf{x},\mathbf{x}}$ and $\Sigma_{\mathbf{xx}}$, respectively. There are many choices of these estimators, some of which include tuning parameter(s). For example, one can use discretizing techniques and finite differences to estimate them. We call this approach the *simple two-stage method*. Alternatively, we could use roughness penalized basis splines to estimate $\hat{\mathbf{x}}(t)$, then apply the differential operator and integral operator to estimate those terms. We call the latter approach the *functional two-stage method*. These two methods are described in detail in Section S7, Supplementary Text. In either case, $\hat{\Sigma}_{D\mathbf{x},\mathbf{x}}$ and $\hat{\Sigma}_{\mathbf{xx}}$ could be represented by the following matrix operations

$$\hat{\Sigma}_{\mathbf{xx}} = YSY', \quad \hat{\Sigma}_{D\mathbf{x},\mathbf{x}} = YLY', \quad \hat{A} = YLY'(YSY')^{-1}. \tag{29}$$

Here S and L were two $(n \times n)$-dimensional matrices obtained from the particular estimating procedure, such as the smoothing step in the functional two-stage methods.

The following two theorems state that: a) when there is no error in estimating $\mathbf{x}(t)$, $\hat{A}$ defined in Eq. (28) is exact, and b) small errors in estimating $\mathbf{x}(t)$ and $D\mathbf{x}(t)$ only induce a small error in $\hat{A}$ for an identifiable $(A, \mathbf{x}_0)$.

Theorem 3.3. *Assume that a linear ODE system with constant coefficient is identifiable at $\mathbf{x}_0$. A matrix $A \in M_{d\times d}$ is its system matrix if and only if it satisfies*

$$A = \Sigma_{D\mathbf{x},\mathbf{x}} \cdot \Sigma_{\mathbf{xx}}^{-1}. \tag{30}$$

Theorem 3.4. *Let $\hat{\mathbf{x}}(t)$ and $D\hat{\mathbf{x}}(t)$ be the estimates of the solution trajectory and its derivative used in Eq. (28). Let $\delta_1 := \|\hat{\mathbf{x}}(t) - \mathbf{x}(t)\|_{L^2}$ and $\delta_2 := \|D\hat{\mathbf{x}}(t) - D\mathbf{x}(t)\|_{L^2}$ be the estimation errors measured in L^2 norm of $\hat{\mathbf{x}}(t)$ and $D\hat{\mathbf{x}}(t)$, respectively; and define $\delta := \max(\delta_1, \delta_2)$. We have*

$$||\hat{A} - A||_F \leqslant C\delta. \tag{31}$$

Here C is a multiplicative constant that depends on $\|\hat{\mathbf{x}}(t)\|$, $\|D\hat{\mathbf{x}}(t)\|$, and the condition number of $\Sigma_{\mathbf{xx}}$.

Proof. The proofs of the above two theorems are provided in Section S4, Supplementary Text. □

It is well known that, with reasonable knot placement and design points (t_js), $\hat{x}(t)$ and $D\hat{x}(t)$ obtained by roughness penalized smoothing splines converge to $\mathbf{x}(t)$ and $D\mathbf{x}(t)$ in L^2 norm. Given Theorem 3.3, it is reasonable to assume that $\hat{A} - A$ is small in our subsequent analyses.

3.3.3. Smoothed condition number

We first propose a straightforward generalization of Stanhope's condition number, called the smoothed condition number (SCN), to measure practical identifiability:

$$\tau(Y, S) := \kappa(\hat{\Sigma}_{\mathbf{xx}}) = \kappa(YSY'). \tag{32}$$

Apparently, Eq. (28) is the main motivation of this generalization. Compared with Stanhope's κ statistic, SCN incorporates the information contained in the smoothing operator S, therefore captures more information of the parameter estimation procedure.

3.3.4. Practical identifiability score

Based on Eq. (29), a small perturbation of data could induce a small $||\hat{A}-A||_F^2$. To conduct a formal sensitivity analysis, we need to make the following additional assumptions:

1. Measurement errors are uncorrelated: $\text{cor}(\epsilon_{ij},\epsilon_{i'j'})=0$ if $i\neq i'$ or $j\neq j'$.
2. These errors are relatively small, namely, $\text{var}(\epsilon_{ij})\leqslant\sigma^2\ll 1$ for all i,j.
3. $\|\hat{A}(X)-A\|_F^2\ll\|\hat{A}(Y)-A\|_F^2$, namely, the *numerical error* in $\hat{A}$ due to the use of discrete data is much smaller than the *variance* of $\hat{A}$ caused by measurement error. This assumption can also be expressed as $A\approx XLX'(XSX')^{-1}$, which is a reasonable assumption for cases in which n is large based on Theorem 3.3.

Denote $\boldsymbol{\epsilon}_S=\boldsymbol{\epsilon}SX'+XS\boldsymbol{\epsilon}'$, $\boldsymbol{\epsilon}_L=\boldsymbol{\epsilon}LX'+XL\boldsymbol{\epsilon}'$, and $N=\left(XSX'\right)^{-1}$. Based on the above assumptions, we have

$$\begin{aligned}\hat{A}(X+\boldsymbol{\epsilon})-A&=(X+\boldsymbol{\epsilon})L(X+\boldsymbol{\epsilon})'\left((X+\boldsymbol{\epsilon})S(X+\boldsymbol{\epsilon})'\right)^{-1}-A\\&\approx(XLX'+\boldsymbol{\epsilon}LX'+XL\boldsymbol{\epsilon}')\left(XSX'+\boldsymbol{\epsilon}SX'+XS\boldsymbol{\epsilon}'\right)^{-1}-A\\&\approx(XLX'+\boldsymbol{\epsilon}_L)(N-N\boldsymbol{\epsilon}_SN)-A\\&\approx-XLX'N(\boldsymbol{\epsilon}_S)N+\boldsymbol{\epsilon}_LN\\&\approx(\boldsymbol{\epsilon}_L-A\boldsymbol{\epsilon}_S)N.\end{aligned}\tag{33}$$

Using Proposition S3.1, we have

$$\begin{aligned}E(\boldsymbol{\epsilon}_L'\boldsymbol{\epsilon}_L)&=E((\boldsymbol{\epsilon}L'X'+XL'\boldsymbol{\epsilon}')(\boldsymbol{\epsilon}LX'+XL\boldsymbol{\epsilon}'))\\&=E\left(\boldsymbol{\epsilon}L'X'\boldsymbol{\epsilon}LX'+\boldsymbol{\epsilon}L'X'XL\boldsymbol{\epsilon}'+XL'\boldsymbol{\epsilon}'\boldsymbol{\epsilon}LX'+XL'\boldsymbol{\epsilon}'XL\boldsymbol{\epsilon}'\right)\\&=\sigma^2\left(XL^2X'+\text{tr}(L'X'XL)\cdot I_d+d\cdot XL'LX'+X(L')^2X'\right).\\E\left(\boldsymbol{\epsilon}_S'A'\boldsymbol{\epsilon}_L\right)&=E((\boldsymbol{\epsilon}S'X'+XS'\boldsymbol{\epsilon}')A'(\boldsymbol{\epsilon}LX'+XL\boldsymbol{\epsilon}'))\\&=E\left(\boldsymbol{\epsilon}S'X'A'\boldsymbol{\epsilon}LX'+XS'\boldsymbol{\epsilon}'A'\boldsymbol{\epsilon}LX'+\boldsymbol{\epsilon}S'X'A'XL\boldsymbol{\epsilon}'+XS'\boldsymbol{\epsilon}'A'XL\boldsymbol{\epsilon}'\right)\\&\leqslant\sigma^2\left(AXSLX'+\text{tr}(A)\cdot XS'LX'+\text{tr}(S'X'A'XL)\cdot I_d+XS'L'X'A\right).\\E\left(\boldsymbol{\epsilon}_S'A'A\boldsymbol{\epsilon}_S\right)&=E((\boldsymbol{\epsilon}S'X'+XS'\boldsymbol{\epsilon}')A'A(\boldsymbol{\epsilon}SX'+XS\boldsymbol{\epsilon}'))\\&=E\left(\boldsymbol{\epsilon}S'X'A'A\boldsymbol{\epsilon}SX'+XS'\boldsymbol{\epsilon}'A'A\boldsymbol{\epsilon}SX'\right.\\&\quad\left.+\boldsymbol{\epsilon}S'X'A'AXS\boldsymbol{\epsilon}'+XS'\boldsymbol{\epsilon}'A'AXS\boldsymbol{\epsilon}'\right)\\&\leqslant\sigma^2\left(A'AXS^2X'+\text{tr}(A'A)\cdot XS'SX'\right.\\&\quad\left.\text{tr}\left(S'X'A'AXS\right)\cdot I_d+X(S')^2X'A'A\right).\end{aligned}\tag{34}$$

Therefore

$$\begin{aligned}||\hat{A}(X+\boldsymbol{\epsilon})-A||_F^2&\approx\text{tr}\left((\boldsymbol{\epsilon}_L-A\boldsymbol{\epsilon}_S)N^2\left(\boldsymbol{\epsilon}_L'-\boldsymbol{\epsilon}_S'A'\right)\right)\\&=\text{tr}\left(\left(\boldsymbol{\epsilon}_L'-\boldsymbol{\epsilon}_S'A'\right)(\boldsymbol{\epsilon}_L-A\boldsymbol{\epsilon}_S)N^2\right)\\&=\text{tr}\left(\left(\boldsymbol{\epsilon}_L'\boldsymbol{\epsilon}_L-\boldsymbol{\epsilon}_S'A'\boldsymbol{\epsilon}_L-\boldsymbol{\epsilon}_L'A\boldsymbol{\epsilon}_S+\boldsymbol{\epsilon}_S'A'A\boldsymbol{\epsilon}_S\right)N^2\right)\\E\|\hat{A}(X+\boldsymbol{\epsilon})-A\|_F^2&\approx\text{tr}\left(\left(E\left(\boldsymbol{\epsilon}_L'\boldsymbol{\epsilon}_L\right)-2E\left(\boldsymbol{\epsilon}_S'A'\boldsymbol{\epsilon}_L\right)+E\left(\boldsymbol{\epsilon}_S'A'A\boldsymbol{\epsilon}_S\right)\right)N^2\right)\\&\leqslant\sigma^2W(X).\end{aligned}\tag{35}$$

Here

$$\begin{aligned}W(X|A,S,L)&:=\text{tr}\left((XSX')^{-2}\left(XL^2X'+d\cdot XL'LX'+X(L')^2X'\right.\right.\\&\quad-2AXSLX'-2\text{tr}(A)\cdot XS'LX'-2XS'L'X'A\\&\quad+A'AXS^2X'+\text{tr}(A'A)\cdot XS'SX'+X(S')^2X'A'A)\\&\quad\left.\left.+\text{tr}(L'X'XL-2S'X'A'XL+S'X'A'AXS)\cdot I_d\right)\right.\\&=\text{vec}((XS'X')^{-2})'\text{vec}\left(XL^2X'+d\cdot XL'LX'+X(L')^2X'\right.\\&\quad-2AXSLX'-2\text{tr}(A)\cdot XS'LX'-2XS'L'X'A\\&\quad\left.+A'AXS^2X'+\text{tr}(A'A)\cdot XS'SX'+X(S')^2X'A'A\right)\\&\quad+\text{tr}(L'X'XL-2S'X'A'XL+S'X'A'AXS)\text{tr}(N^2).\end{aligned}\tag{36}$$

By construction, $W(X)$ is a scalar that quantifies the MSE of $\hat{A}$ as a function of σ^2, the maximum variance of ϵ_{ij} for all $i=1,2,\ldots,d$ and $j=1,2,\ldots,n$. Smaller values of $W(X)$ imply better practical identifiability in reconstructing A. Motivated by this fact, we define the practical identifiability score (PIS) as the sample version of $W(X)$. Specifically, PIS (denoted as w^* in Eq. (37)) is computed by replacing X and A in Eq. (36) with Y and $\hat{A}:=YLY'(YSY')^{-1}$, respectively:

$$w^*(Y|S,L):=W(Y|\hat{A},S,L).\tag{37}$$

Compared with Stanhope's κ and SCN, PIS depends not only on the observed data (Y), but also the S and L matrix of the particular two-stage method used in reconstructing $\hat{A}$, therefore it is a more accurate indicator of practical identifiability. A simulation study was designed to demonstrate this point in Section 4.2.

4. Simulation studies

4.1. ICIS is inversely correlated with the relative estimation error (REE)

We design SIM1 to demonstrate that the ICIS is inversely correlated with estimation error, as predicted in Section 3.2. In this simulation, the system matrix and its two invariant subspaces are:

$$A = \begin{pmatrix} -0.1 & 3 & 0 \\ -3 & -0.1 & 0 \\ 0 & 0 & -0.5 \end{pmatrix}, \qquad L_1 = \text{span}(\mathbf{e}_1, \mathbf{e}_2), \quad L_2 = \text{span}(\mathbf{e}_3).$$

Here $\mathbf{e}_k$ is the kth column vector of I_3, a.k.a. the kth natural basis vector of $\mathbb{R}^3$. We generate $\mathbf{x}_0$ from $N(0_3, I_3)$ first, then standardize it to have unit length to reduce the variation in ICIS due to different $|\mathbf{x}_0|$. This is equivalent to sampling $\mathbf{x}_0$ from a uniform distribution on S^2.

Once $\mathbf{x}_0$ is generated, we compute $X_{\cdot j} := e^{t_j A}\mathbf{x}_0$ at a time grid with range $[0, 6]$ and step $\Delta t = 0.1$, i.e., $t_j = 0, 0.1, \ldots, 6$ for $j = 1, \ldots, 61$. We also add a small noise $\epsilon_{ij} \sim N(0, \sigma^2)$, $\sigma = 0.05$ to each observation to create noisy data $Y = X + \boldsymbol{\epsilon}$. A functional two-stage method based on cubic splines with roughness penalty $\lambda = 0.001$ is used to estimate $\hat{A}$ from both noisy (Y) and noise-free (X) data.

The accuracy of estimation is measured by relative estimation error (REE), defined as follows

$$\text{REE}\big(\hat{A}, A\big) := \frac{\|\hat{A} - A\|_F}{\|A\|_F}. \tag{38}$$

We repeat SIM1 for 100 times, each with randomly generated $\mathbf{x}_0$ and measurement error. We find that ICIS (w_0^*) is strongly negatively correlated with REE. The Spearman correlation between these two quantities is $\rho_1 = -0.803$ for the noisy data and $\rho_2 = -0.843$ for the noise-free data. This inverse correlation is visualized in Fig. 3. Other than a few outliers, $1/w_0^*$ has an almost perfect linear relationship with REE in the noise-free case (the second row of Fig. 3). The correlation between $1/w_0^*$ and REE is weaker but still quite apparent for the noisy data (the first row of Fig. 3).

4.2. Using SCN and PIS to classify identifiable and un-identifiable systems

We design SIM2 to demonstrate that, when data collected at all time points are available, SCN and PIS have better performance in classifying identifiable and un-identifiable Systems than ICIS and Stanhope's κ. SIM2 contains one identifiable case and two unidentifiable cases, which are described as follows.

1. In each one of 200 repetitions, we generate two 4×4-dimensional system matrices A and B, and two initial conditions $\mathbf{x}_0^{(a)}$ and $\mathbf{x}_0^{(b)}$ on S^3.
2. Both A and B have one pair of complex eigenvalues and two real eigenvalues. The eigenvalues of A, $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, are generated in this way

$$\begin{aligned} &\lambda_1, \lambda_2 = -0.1 \pm bi, \quad b \sim \text{Unif}([2, 4]). \\ &\lambda_3 \sim \text{Unif}([-0.8, -0.4]), \quad \lambda_4 \sim \text{Unif}([-2, -1.2]). \end{aligned} \tag{39}$$

   The eigenvalues of B are set to be $(\lambda_1, \lambda_2, \lambda_3, \lambda_3)$, namely, B has a pair of *repeated eigenvalues* by construction. Therefore it is not identifiable at any initial condition.
3. After generating the eigenvalues, we sample an orthogonal matrix Q from the standardized Haar measure (the uniform distribution) on the orthogonal group $O(4)$, and create two system matrices A and B as follows:

$$A = Q\begin{pmatrix} -0.1 & b & 0 & 0 \\ -b & -0.1 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}Q', \quad B = Q\begin{pmatrix} -0.1 & b & 0 & 0 \\ -b & -0.1 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}Q'. \tag{40}$$

4. Like SIM1, we sample $\mathbf{x}_0^{(a)}$ from the uniform distribution on S^{d-1} ($d = 4$ in this case). Furthermore, we only keep those $\mathbf{x}_0^{(a)}$ with relatively large ICIS, namely $w_0^*(A, \mathbf{x}_0^{(a)}) > 0.2$. This ensures the practical identifiability of A at $\mathbf{x}_0^{(a)}$.
5. Once $\mathbf{x}_0^{(a)}$ is sampled, we define

$$\mathbf{x}_0^{(b)} = \frac{(I_4 - Q_4 Q_4')\mathbf{x}_0^{(a)}}{\big|(I_4 - Q_4 Q_4')\mathbf{x}_0^{(a)}\big|}. \tag{41}$$

   By construction, $\mathbf{x}_0^{(b)}$ is a unit vector such that $\mathbf{x}_0^{(b)} \perp Q_4$, so that ICIS equals zero in this case. According to our theoretical derivations, A is not identifiable at $\mathbf{x}_0^{(b)}$ due to ill-positioned initial conditions.
6. We compute three sets of solution trajectories: case (A) corresponds with $(A, \mathbf{x}_0^{(a)})$, case (B) with $(A, \mathbf{x}_0^{(b)})$, and case (C) with $(B, \mathbf{x}_0^{(a)})$.

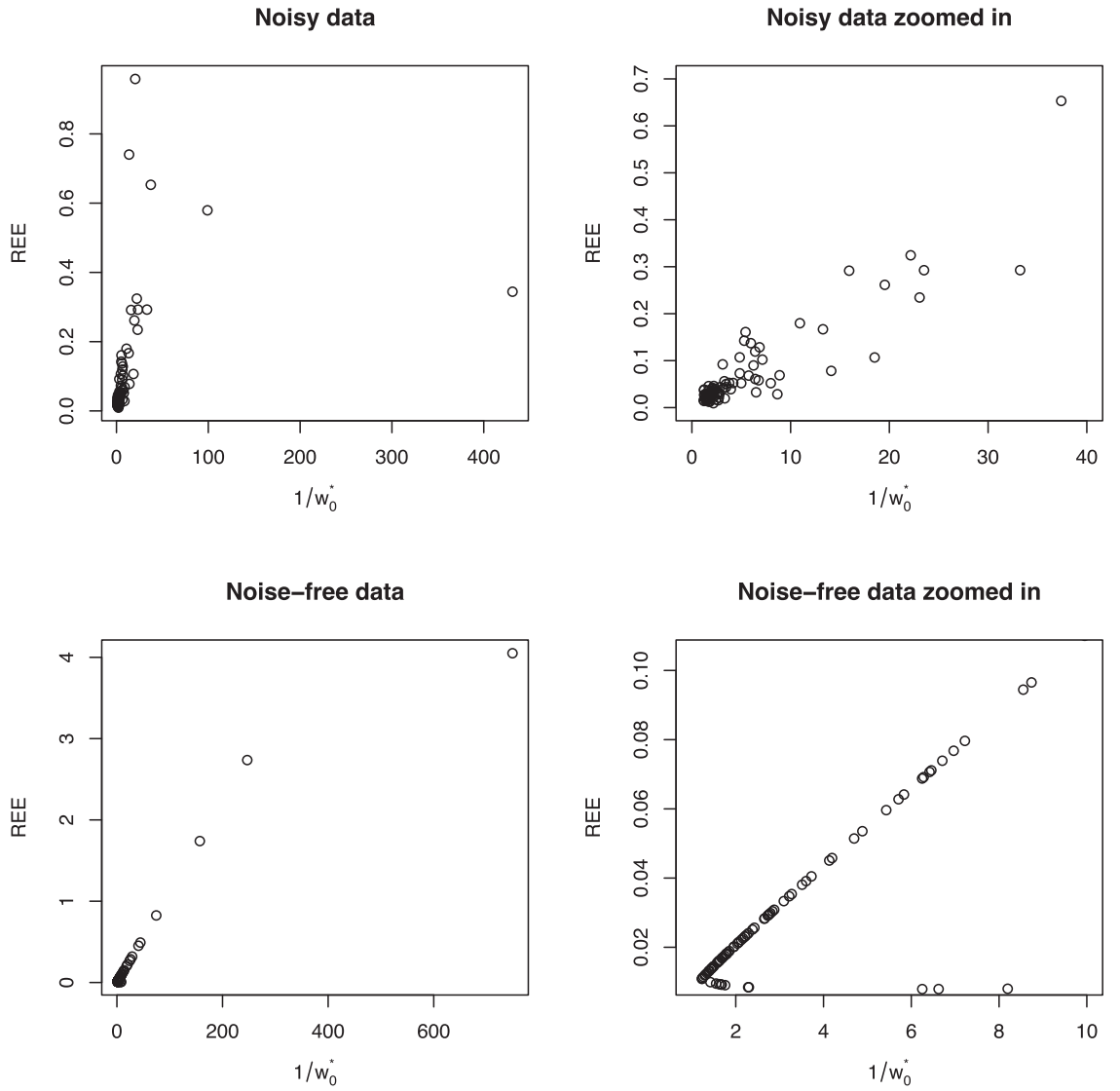

Fig. 3. ICIS (w_0^*) is inversely correlated with the relative estimation error (REE). In all four subfigures, the x-axis is $1/w_0^*$, y-axis is REE. Each dot represents one of the 100 repetitions of SIM1. Top two subfigures are generated from noisy data (Y); bottom two subfigures uses noise-free data (X). The right subfigures are the zoomed-in version of the left subfigures.

For each case, we compute ICIS (w_0^*), SCN (τ), PIS (w^*), and Stanhope's κ, based on both noisy and noise-free data. The results are illustrated in Figs. 4 and 5. We find that for noise-free data, SCN, PIS, and κ perform very well, with almost perfect area under the curve (AUC) in receiver operating characteristic (ROC) analyses. However, ICIS only has a relatively small AUC=0.773. This is not a surprise at all because ICIS is designed to detect un-identifiability issues associated with ill-positioned *initial conditions* (case B), not un-identifiable systems that have *repeated eigenvalues* (case C). This fact is also revealed in the corresponding boxplot in Fig. 4 (second column).

For data with noise, κ is almost uninformative (AUC=0.503), but SCN, which is a smoothed extension of κ, works very well (AUC=0.946). It suggests that taking the smoothing effect into the consideration in SCN improves its utility as a classifier of identifiable systems.

While SCN has significantly better performance than κ and ICIS, it is still an *ad hoc* metric of practical identifiability that does not account for the uncertainty in $\hat{A}$ due to measurement error. In contrast, PIS is designed based on rigorous asymptotic analysis on the variance of $\hat{A}$, therefore PIS has the best performance (AUC=0.962). That being said, we need to point out that from the computational perspective, SCN is more efficient and numerically robust, because SCN does not contain $(YSY')^{-1}$ and $(YSY')^{-2}$ terms used in PIS, which could have numerical issues if the dimension of the ODE system is large. In summary, SCN could be considered as a simplified version of PIS that is less vulnerable to computational issues.

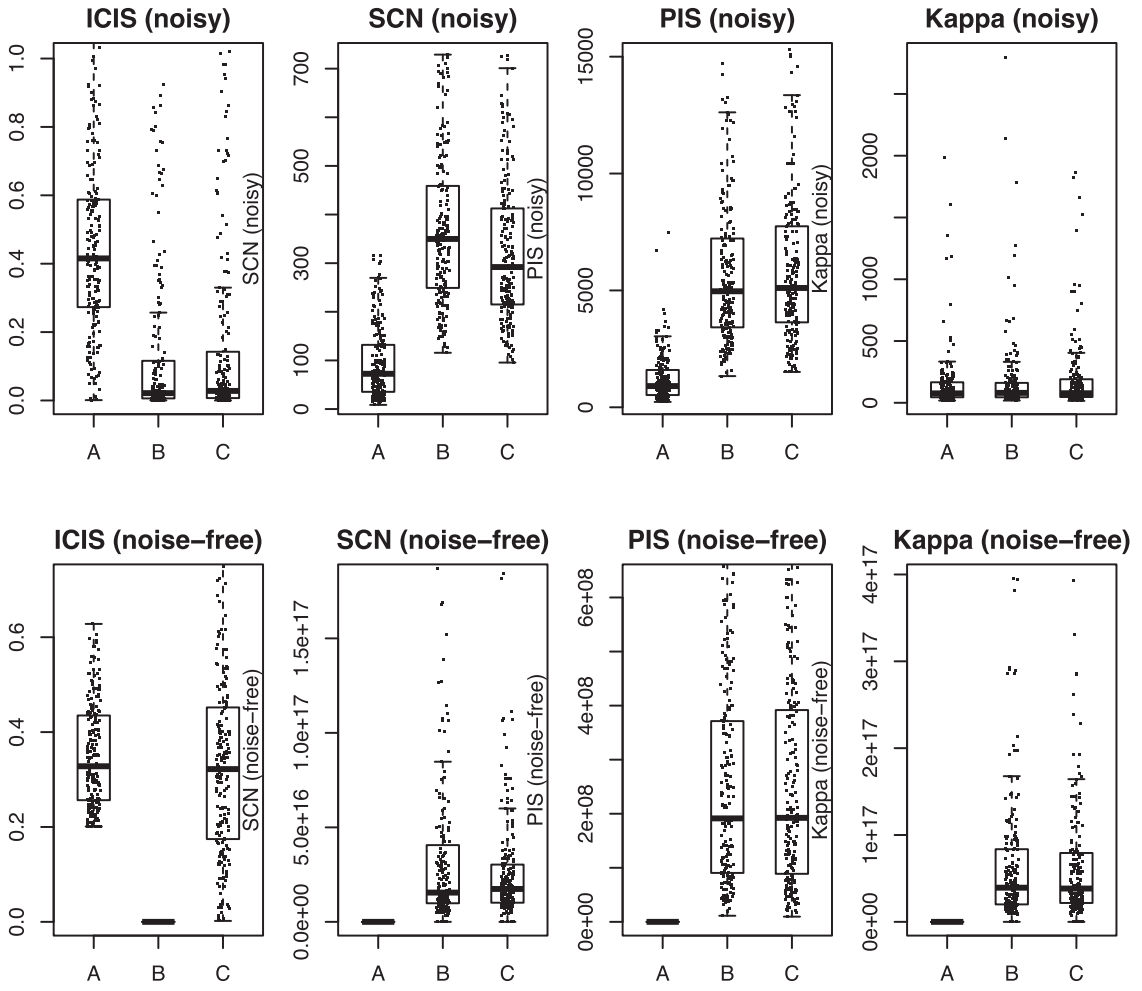

Fig. 4. Values of four identifiability measures (ICIS, SCN, PIS, and Stanhope's κ) computed from 200 repetitions of SIM2. For all four scores, smaller values imply better identifiability. The first row of subfigures are computed from noisy data (Y), the second row of subfigures are computed from noise-free data (X). Each subfigure has three cases: (A) is the identifiable case generated by $(A, \mathbf{x}_0^{(a)})$; (B) is the unidentifiable case generated by $(A, \mathbf{x}_0^{(b)})$; (C) is the unidentifiable case generated by $(B, \mathbf{x}_0^{(b)})$.

Both noisy and noise-free data for all three cases were illustrated in Fig. 6. Notably, visual examinations did not reveal apparent differences between the three cases, suggesting that the identifiability of the ODE system does not depend on obvious features in the solution trajectories.

5. Conclusions

Classical identifiability analyses for ODE systems typically depend on the availability of solution trajectories from arbitrarily many initial points. However, in many real world problems, the system matrix must be estimated from just one observed trajectory. In this case, identifiability depends not only on the properties of A, but also the initial condition $\mathbf{x}_0$. In this case, the $(A, \mathbf{x}_0)$-identifiability used in our study is more appropriate than classical identifiability measures.

We develop an explicit formula of all matrices that are unidentifiable with A at a given $\mathbf{x}_0$ in this study. It enables researchers to gain better insight into identifiability analysis and help them design more practical simulation studies.

Another notable finding of our study is that when A is coupled (not diagonal), an identifiability issue in just a one-dimensional invariant subspace could cause issues in many other elements of A (e.g., Example 1). Consequently, identifiability analyses that only depend on the topology of the network are insufficient in practice.

For high-dimensional cases, even if $\mathbf{x}_0$ is generated in a "completely random" fashion (e.g., $\mathbf{x}_0 \sim N(0_d, I_d)$), by chance, one invariant subspace of A may have very little information, which in turn leads to practical identifiability issues. In fact, we are able to prove that when $d \to \infty$, $\mathrm{ICIS}(A, \mathbf{x}_0) \to 0$ for a large class of random ODE systems, which suggests that the practical identifiability properties of low-dimensional and high-dimensional systems are fundamentally different. We believe it will be rewarding to derive more accurate convergence rates for ICIS as a function of d in a future study. It will require

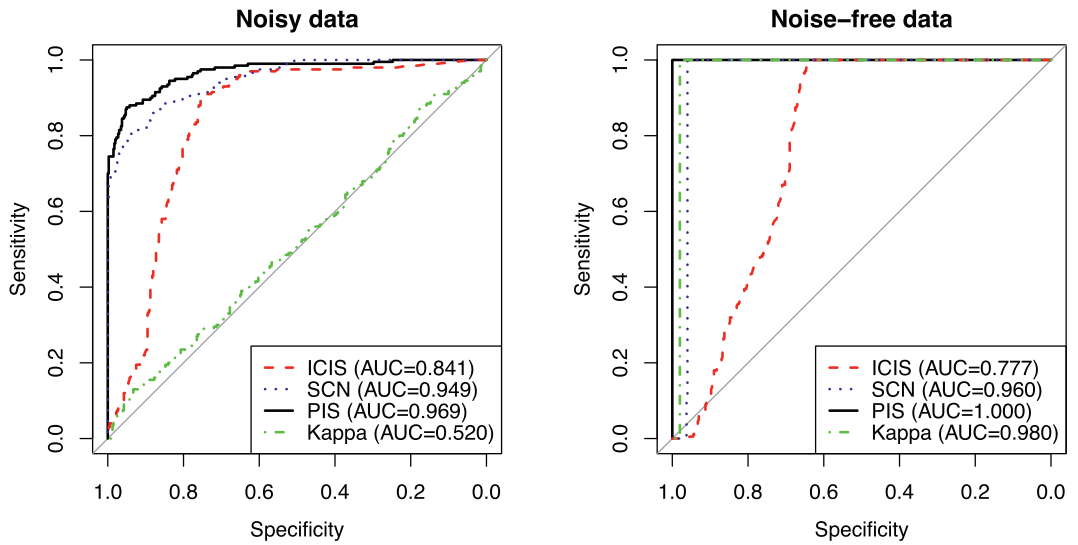

Fig. 5. ROC curves of four identifiability measures (ICIS, SCN, PIS, and Stanhope's κ) in classifying case A (identifiable systems) from cases B and C (both are unidentifiable systems) in SIM2, with 200 repetitions. The left panel use data with noise; the right panel use noise-free data.

combining advanced techniques in random matrix theory, especially for ensembles of asymmetric matrices (e.g., Conjecture S2.1 in Supplementary Text) in which the Q matrix are no longer orthogonal, with the identifiability analysis of ODE systems.

In this study, we also developed two scores, SCN and PIS, that use the entire dataset obtained at all time points, to quantify the practical identifiability for real world applications. Both SCN and PIS are more accurate than Stanhope's κ when noise is present in the data, as shown by extensive simulation studies.

While our methods are developed for homogeneous systems, it should be relatively easy to generalize them for the following inhomogeneous linear ODE system

$$\begin{cases} D\mathbf{x}(t) = A\mathbf{x}(t) + b, \quad b \in \mathbb{R}^d, \quad t \in (0, T], \\ \mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^d. \end{cases} \tag{42}$$

This is because Eq. (42) can be transformed into an *equivalent homogeneous system* with a simple mathematical technique. Let $\mathbf{z}(t) = (\mathbf{x}(t), 1)' = (x_1(t), \ldots, x_d(t), 1)' \in \mathbb{R}^{d+1}$. It satisfies the following ODE

$$\begin{cases} D\mathbf{z}(t) = \check{A}\mathbf{z}(t), \qquad t \in (0, T], \\ \mathbf{z}(0) = (\mathbf{x}_0, 1)' \in \mathbb{R}^{d+1}. \end{cases} \qquad \check{A} := \begin{pmatrix} A & b \\ 0_d & 0 \end{pmatrix}. \tag{43}$$

Therefore, the identifiability of Eq. (42) is the same as the identifiability of Eq. (43), which is a homogeneous equation with the constraint that the last row of $\check{A}$ must be zeros. Let $M_0 = \left\{B \in M_{(d+1)\times(d+1)} : M_{(d+1)\cdot} = 0_{d+1}\right\}$ be the set of $(d+1)\times(d+1)$-dimensional matrices such that their last rows equal 0_{d+1}. The unidentifiability class associated with system $(A, b, \mathbf{x}_0)$, denoted by $[A, b]_{\mathbf{x}_0}$, is the following subset of $M_{d\times d}$:

$$[A, b]_{\mathbf{x}_0} = [\check{A}]_{(\mathbf{x}_0, 1)'} \cap M_0. \tag{44}$$

More future work is required to extend ICIS, SCN, and PIS for constrained systems, so that they can be used as practical guidance for applications with *a priori* information.

In the near future, we plan to extend our work to the following family of nonlinear ODE system:

$$D\mathbf{x}(t) = Af(\mathbf{x}(t)), \qquad \mathbf{x}(0) = \mathbf{x}_0. \tag{45}$$

Here $f(\cdot) : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ is a known locally Lipschitz function of $\mathbf{x}(t)$, $A \in M_{d_1\times d_2}$ is the system matrix that needs to be estimated. This system has been studied by Stanhope and colleagues, and their main conclusion (Theorem (5.3) in [20]) is very similar to that for the linear ODE systems: A is identifiable at $\mathbf{x}_0$ if and only if the solution curve is not confine in a proper linear subspace of $\mathbb{R}^{d_2}$. To extend the SCN and PIS we developed in this study to Eq. (45), we will need to study the sensitivity of an extended two-stage method that works for Eq. (45).

Using linearization techniques, we believe SCN and PIS can be further extended to other types of nonlinear systems. To this end, we need: a) to approximate a nonlinear ODE system by a linear ODE at $\mathbf{x}(t)$; b) to propose a local version of the $(A, \mathbf{x}_0)$-identifiability that works in a neighborhood of A at $\mathbf{x}(t)$; c) to study the sensitivity of a reasonable parameter estimator for such system, and propose an identifiability score based on the useful information aggregated from all time points.

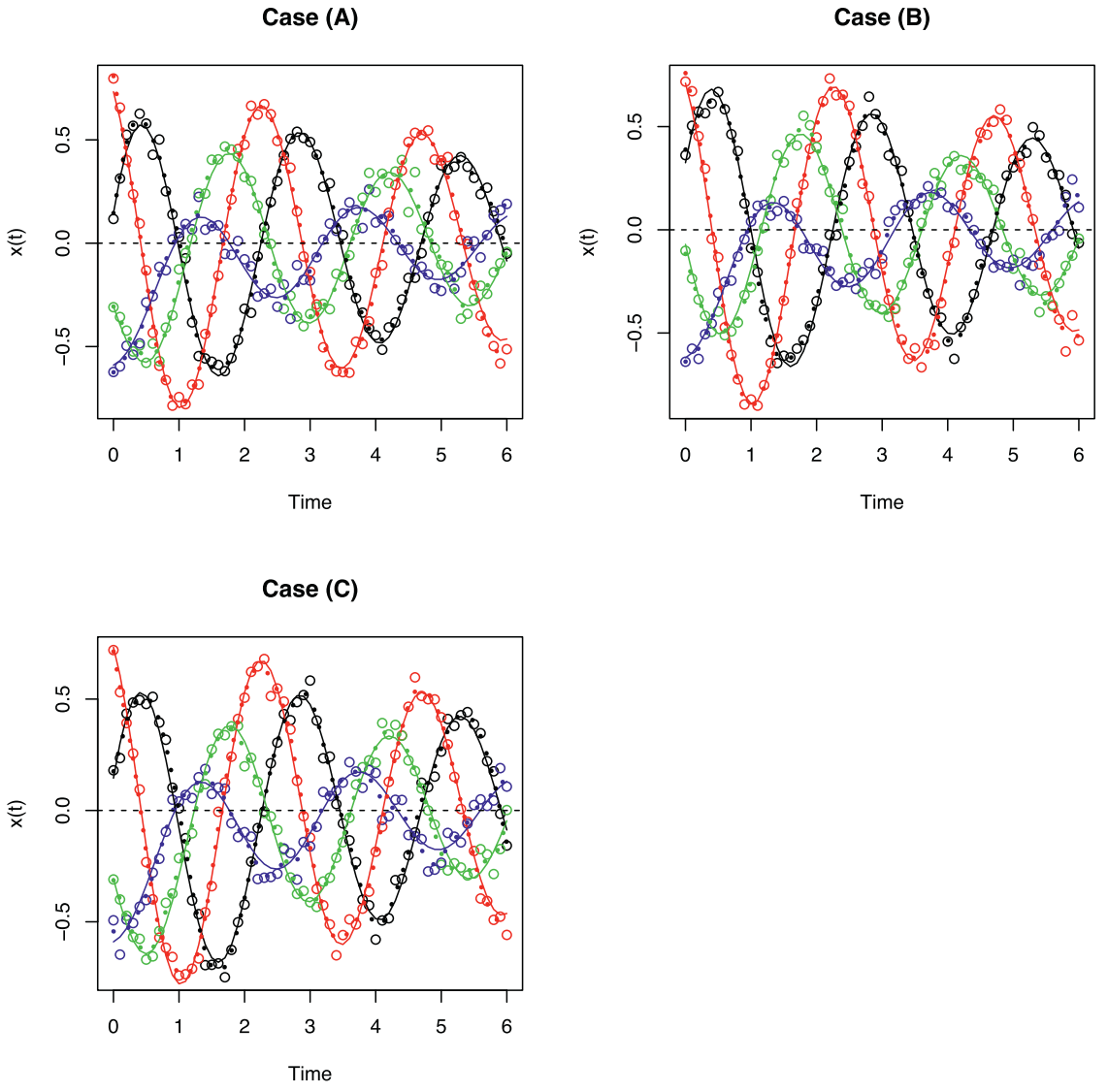

Fig. 6. An illustration of the fitted solution curves in SIM2. Dots represent noisy data y_{ij}, four colors represent four dimensions. Solid curves are the noise-free solutions ($x(t)$) and dotted curves are the smoothed curves ($\hat{x}(t)$) fitted by roughness penalized splines. The same roughness penalty ($\lambda = 0.001$) was used in all three cases. Overall, the fitted curves agree with the noise-free data well.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.amc.2022.127260.

References

[1] J. Butcher, Ordinary differential equations, in: Walter Gautschi, vol. 3, Springer, 2014, pp. 7–8.
[2] D. Commenges, D. Jolly, J. Drylewicz, H. Putter, R. Thiébaut, Inference in HIV dynamics models via hierarchical likelihood, Comput. Stat. Data Anal. 55 (1) (2011) 446–456.
[3] H. De Jong, Modeling and simulation of genetic regulatory systems: a literature review, J. Comput. Biol. 9 (1) (2002) 67–103.
[4] P.W. Hemker, Numerical methods for differential equations in system simulation and in parameter estimation, Anal. Simul. Biochem. Syst. 28 (1972) 59–80.
[5] N.S. Holter, A. Maritan, M. Cieplak, N.V. Fedoroff, J.R. Banavar, Dynamic modeling of gene expression data, Proc. Natl. Acad. Sci. 98 (4) (2001) 1693–1698.
[6] Y. Huang, D. Liu, H. Wu, Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system, Biometrics 62 (2) (2006) 413–423.
[7] M. Lavielle, A. Samson, A. Karina Fermin, F. Mentré, Maximum likelihood estimation of long-term HIV dynamic models and antiviral response, Biometrics 67 (1) (2011) 250–259.
[8] Z. Li, P. Li, A. Krishnan, J. Liu, Large-scale dynamic gene regulatory network inference combining differential equation models with local dynamic Bayesian network analysis, Bioinformatics 27 (19) (2011) 2686–2691.

[9] T. Lu, H. Liang, H. Li, H. Wu, High-dimensional ODEs coupled with mixed-effects modeling techniques for dynamic gene regulatory network identification, J. Am. Stat. Assoc. 106 (496) (2011) 1242–1258.
[10] J.O. Ramsay, G. Hooker, D. Campbell, J. Cao, Parameter estimation for differential equations: a generalized smoothing approach (with discussion), J. R. Stat. Soc. 69 (5) (2007) 741–796.
[11] C. Moler, C. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, SIAM Rev. 45 (1) (2003) 3–49.
[12] Y. Huang, H. Wu, A Bayesian approach for estimating antiviral efficacy in HIV dynamic models, J. Appl. Stat. 33 (2) (2006) 155–174.
[13] Y. Huang, H. Wu, E.P. Acosta, Hierarchical Bayesian inference for HIV dynamic differential equation models incorporating multiple treatment factors, Biom. J. 52 (4) (2010) 470–486.
[14] Z. Li, M.R. Osborne, T. Prvan, Parameter estimation of ordinary differential equations, IMA J. Numer. Anal. 25 (2) (2005) 264–285.
[15] H. Putter, S. Heisterkamp, J. Lange, F. De Wolf, A Bayesian approach to parameter estimation in HIV dynamical models, Stat. Med. 21 (15) (2002) 2199–2214.
[16] L. Wu, X. Qiu, Y.-x. Yuan, H. Wu, Parameter estimation and variable selection for big systems of linear ordinary differential equations: a matrix-based approach, J. Am. Stat. Assoc. 114 (526) (2019) 657–667.
[17] H. Xue, H. Miao, H. Wu, Sieve estimation of constant and time-varying coefficients in nonlinear ordinary differential equation models by considering both numerical error and measurement error, Ann. Stat. 38 (4) (2010) 2351–2387.
[18] H. Miao, X. Xia, A.S. Perelson, H. Wu, On identifiability of nonlinear ode models and applications in viral dynamics, SIAM Rev. 53 (1) (2011) 3–39.
[19] A. THOWSEN, Identifiability of dynamic systems, Int. J. Syst. Sci. 9 (7) (1978) 813–825.
[20] S. Stanhope, J. Rubin, D. Swigon, Identifiability of linear and linear-in-parameters dynamical systems from a single trajectory, SIAM J. Appl. Dyn. Syst. 13 (4) (2014) 1792–1815.
[21] X. Qiu, S. Wu, S.P. Hilchey, J. Thakar, Z.-P. Liu, S.L. Welle, A.D. Henn, H. Wu, M.S. Zand, Diversity in compartmental dynamics of gene regulatory networks: the immune response in primary influenza a infection in mice, PLoS ONE 10 (9) (2015).
[22] X. Sun, F. Hu, S. Wu, X. Qiu, P. Linel, H. Wu, Controllability and stability analysis of large transcriptomic dynamic systems for host response to influenza infection in human, Infect. Dis. Model. 1 (1) (2016) 52–70.
[23] S. Wu, Z.-P. Liu, X. Qiu, H. Wu, High-dimensional ordinary differential equation models for reconstructing genome-wide dynamic regulatory networks, in: Topics in Applied Statistics, Springer, New York, NY, 2013, pp. 173–190.
[24] S. Wu, Z.-P. Liu, X. Qiu, H. Wu, Modeling genome-wide dynamic regulatory network in mouse lungs with influenza infection using high-dimensional ordinary differential equations, PLoS ONE 9 (5) (2014).
[25] J.A. McCullers, J.L. McAuley, S. Browall, A.R. Iverson, K.L. Boyd, B. Henriques Normark, Influenza enhances susceptibility to natural acquisition of and disease due to streptococcus pneumoniae in ferrets, J. Infect. Dis. 202 (8) (2010) 1287–1295.
[26] E. Tunali, T.-J. Tarn, New results for identifiability of nonlinear systems, IEEE Trans. Automat. Control 32 (2) (1987) 146–154.
[27] A.M. Jeffrey, X. Xia, I. Craig, Identifiability of HIV/AIDS models, in: Deterministic and Stochastic models of AIDS epidemics and HIV infections with intervention, 2005, pp. 255–286.
[28] X. Xia, C.H. Moog, Identifiability of nonlinear systems with application to HIV/AIDS models, IEEE Trans. Automat. Control 48 (2) (2003) 330–336.
[29] J. Ginibre, Statistical ensembles of complex, quaternion, and real matrices, J. Math. Phys. 6 (1965) 440.
[30] I. Gohberg, P. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, SIAM, 2006.
[31] N. Lehmann, H.-J. Sommers, Eigenvalue statistics of random real matrices, Phys. Rev. Lett. 67 (8) (1991) 941–944.
[32] T. Tao, Topics in Random Matrix Theory, vol. 132, American Mathematical Society Providence, RI, 2012.
[33] G. Livan, M. Novaes, P. Vivo, Introduction to Random Matrices: Theory and Practice, vol. 26, Springer, 2018.
[34] H. Weyl, The Classical Groups: Their Invariants and Representations, vol. 45, Princeton University Press, 1946.