SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2025, Vol. 68, Iss. 2, 122204:1–122204:16
https://doi.org/10.1007/s11432-024-4135-9

Identifiability and sloppiness of structured systems with a matrix fraction description using finite frequency responses

Yunxiang MA[1] & Tong ZHOU[2*]

[1]*Department of Automation, Tsinghua University, Beijing 100084, China*
[2]*Department of Automation and BNRist, Tsinghua University, Beijing 100084, China*

Received 20 March 2024/Revised 19 July 2024/Accepted 16 August 2024/Published online 16 January 2025

Abstract This paper investigates the identifiability of structured systems with a matrix fraction description and attacks practical difficulties in estimating parameters under finite frequency responses. The numerator and denominator polynomial matrices depend affinely on system parameters, which can represent a broad range of systems, especially circuits and mechanical systems. Except for assuming that the denominator matrix is invertible, no further assumptions are adopted. A criterion is provided for validating this assumption utilizing graph theories and matroids. A sufficient and necessary criterion for global identifiability based on the rank of a numerical matrix is given and can be recursively verified with each frequency point, which is computationally attractive for large-scale systems. For an identifiable system, an ellipsoid approximation is obtained for the set of parameters whose corresponding system frequency responses deviate within a specific range from those of the system at a specific parameter value. Using this approximation, explicit expressions for absolute and relative sloppiness have been derived, which can be used to quantify the difficulty of identifying system parameters. Comparisons with the well-known Fisher information matrix have also been performed through a mechanical system, revealing significant differences in quantifying parameter estimation hardness.

Keywords identifiability, sloppiness, structured system, matrix fraction, graph theory, matroid

Citation Ma Y X, Zhou T. Identifiability and sloppiness of structured systems with a matrix fraction description using finite frequency responses. Sci China Inf Sci, 2025, 68(2): 122204, https://doi.org/10.1007/s11432-024-4135-9

1 Introduction

Identifiability is a prerequisite for estimating model parameters from input-output data. In other words, if parameters in the model are unidentifiable, no matter how well-designed the identification experiments are or how advanced the identification algorithms used, it is impossible to estimate them accurately using input-output data. Even with highly effective numerical optimization algorithms, estimated values of the parameters may still be unreliable. Investigating identifiability can provide theoretical guidance for model selection, experimental design, and other aspects. These include enhancing model performance [1], optimizing data measurement conditions [2, 3], and improving controller design [4, 5]. However, with system scales expanding and data volumes increasing, existing methods [6, 7] gradually become computationally infeasible. Developing scalable algorithms remains a challenge. As early as 1987, Ljung [6] highlighted challenges in parameter identifiability verifications. It was not until 2021 that Anstett-Collin et al. [8] reiterated its complexity.

On the other hand, in practice, the data utilized in a parameter estimation process can never be perfect. Factors like modeling errors and measurement noise are inevitable. In some cases, the input signal must also satisfy certain constraints. Thus, it is necessary to consider practical identifiability. According to [9, 10], practical identifiability is generally more challenging and has not yet been well studied.

In addition, for identifiable systems, the difficulty of identification often varies with different systems and even different parameters within the same system. As observed by Transtrum et al. [11] and Waterfall et al. [12], the trajectories of system outputs for all possible parameter values generally exhibit a so-called hyper-ribbon structure, indicating that different parameter variations may have significantly

* Corresponding author (email: tzhou@mail.tsinghua.edu.cn)

different effects on system outputs. There are likely sloppy variations when identifying many parameters simultaneously, which means that even if there are significant changes in system parameter values, they may only cause slight variations in system outputs. This phenomenon is widespread in multivariable and nonlinear systems. In practical identification processes, it is difficult to distinguish whether these variations in system outputs are caused by sloppy parameter variations or by measurement noise. In other words, parameters with sloppy variations are challenging to estimate from actual experimental data with an acceptable accuracy, even if they are theoretically identifiable. Therefore, it is necessary to further study the system's sloppiness, as it can be used to quantify the difficulty of identifying the system's parameters.

Sloppiness is usually considered to be associated with the distribution of eigenvalues of Fisher information matrix (FIM) [9, 11, 13]. Generally, most eigenvalues of FIM are very small, corresponding to sloppy parameter variations that barely affect system behavior. However, the computations involved in this process are often complex and sometimes even unfeasible. Moreover, Chis et al. [14] noted that this measure of sloppiness is not very suitable for system identification and may mislead experimental design. This is because systems with the same degree of sloppiness can exhibit significantly different parameter estimation accuracies. To overcome this difficulty, Zhou et al. [15] proposed a method that uses two metrics, absolute and relative sloppiness, to efficiently and quantitatively assess the sloppiness of a linear fractional transformation (LFT) described descriptor system based on its frequency responses.

The matrix fraction model employed in this paper is widely used in system analysis and synthesis [16], which can describe the constraints among various natural variables. These constraints can be naturally derived from the system's differential equations and have a broad applicability. Moreover, a structured model can be viewed as a collection of many systems, where the non-zero entries typically have clear physical interpretations [6, 17–19], and many practical systems, particularly mechanical and electrical systems, can be modeled in an affine structure, such as compartmental, power, and hydraulic systems [17, 20]. Despite extensive research on analyzing and synthesizing matrix fraction models, such as in describing the stability of multivariable systems [21], there is little work addressing their parameter identifiability and sloppiness. In particular, Arambel and Tadmor [22] analyzed the identifiability of matrix fraction models, providing a lower bound for the persistent excitation of input-output data.

This paper investigates whether the parameters can be uniquely determined from limited system frequency responses, which are closer to practical identification scenarios, and how to calculate the system's sloppiness, which can be used to quantify parameter estimation hardness, in both continuous and discrete time. For the matrix fraction model used in this paper, the numerator and denominator polynomial matrices depend affinely on system parameters, assuming that the denominator polynomial matrix is invertible. This assumption can be verified using the relationship between the generic rank of a matrix and independent matching, when the denominator matrix can be decomposed into a sum of rank-one matrices with respect to θ. Utilizing properties of matrix fraction descriptions, coprime factorizations, and polynomial matrices, a sufficient and necessary condition is derived for the global identifiability of structured systems at a specific parameter value with finite frequency responses. This paper adopts the definition of sloppiness as described in [15]. For an identifiable system, a method for calculating its sloppiness is proposed, utilizing a parametrization of solutions to a linear matrix equation and properties of singular value decomposition (SVD). Specifically, an ellipsoid approximation of a parameter set is explicitly derived, where all parameters within the set result in deviations within a predetermined range between frequency responses of corresponding systems and those at a specific parameter value.

The discussion on sloppiness in this paper is based on the definition of sloppiness provided in [15]. While the matrix fraction model used here can be considered a special case of the LFT described descriptor system, the assumption in this paper is more lenient compared to those in [15]. Specifically, adopting the matrix fraction model instead of the LFT described descriptor system removes an assumption in [15] that a specific transfer function matrix is of full normal row rank. A system described by the matrix fraction model may not meet this specific assumption when recast as an LFT descriptor system. In other words, the matrix fraction model in this paper can characterize systems that do not meet the assumption in [15], extending its applicability beyond the LFT framework. Therefore, the findings of this paper have practical engineering significance and are mathematically non-trivial.

The contributions of this paper are summarized as follows. (1) Aside from the invertibility of the denominator polynomial matrix, which is necessary for a system to function properly, there are no other assumptions. This assumption can be verified for almost all parameters using graph theory and matroids. (2) The condition for global identifiability is both sufficient and necessary, whereas many existing results

are either sufficient or necessary. This condition is based on the rank of a numerical matrix and can be recursively verified with each frequency point, making it computationally attractive for large-scale systems. (3) Explicit expressions for absolute and relative sloppiness have been derived, which are more suitable to quantify the difficulty of identifying system parameters than the widely used FIM. To the best of the author's knowledge, this paper represents the first study on the sloppiness of matrix fraction models.

The rest of this paper is organized as follows. In Section 2, problem descriptions, assumptions and their verification, and some preliminary results are provided. Section 3 further discusses identifiability at finite frequency points. Computations of suggested absolute and relative sloppiness metrics are dealt with in Section 4. Section 5 presents a numerical example. Some concluding remarks are given in Section 6. Finally, Appendixes A–C include proofs of the theorems proposed in this paper.

The following notation and symbols are adopted in this paper. $\operatorname{col}\{X_i|_{i=1}^n\}$ is the vector/matrix stacked by $X_i|_{i=1}^n$ with its ith row block vector/matrix being X_i, and $\operatorname{vec}\{X\}$ represents the vector stacked from left to right by the columns of the matrix. $\star^\dagger$ denotes the Moore-Penrose inverse of a matrix. The superscripts "T" and "H" represent the transpose and conjugate transpose of a matrix/vector, respectively, while subscripts r and j of a complex matrix/vector are real and imaginary parts. $\operatorname{span}\{\star\}$ denotes the linear space spanned by the vector $\star$.

2 Problem formulation and preliminaries

2.1 System description

For a linear time-invariant system with p-dimensional input and q-dimensional output, its matrix fraction model is specifically as follows:

$$D(\lambda,\theta)\,y(\lambda)=N(\lambda,\theta)\,u(\lambda), \tag{1}$$

where λ represents the Laplace operator s for continuous time systems or the Z-transform operator z for discrete time systems. $u(\lambda)$ and $y(\lambda)$ correspond to the input and output data after Laplace or Z transform. $D(\lambda,\theta)$ and $N(\lambda,\theta)$ are polynomial matrices of dimensions $p\times p$ and $p\times q$, respectively, and assumed to be affine with respect to $\theta=(\theta_1,\theta_2,\ldots,\theta_n)^{\mathrm{T}}\in\boldsymbol{\Theta}$. That is,

$$D(\lambda,\theta)=D_0(\lambda)+\sum_{i=1}^{n}\theta_i D_i(\lambda),\ N(\lambda,\theta)=N_0(\lambda)+\sum_{i=1}^{n}\theta_i N_i(\lambda), \tag{2}$$

where n is the number of unknown parameters in the model. For $i=0,1,2,\ldots,n$, $N_i(\lambda)$ and $D_i(\lambda)$ are known polynomial matrices, representing prior information about the system structure. $\theta_i|_{i=1}^n$ are independent unknown parameters in the model, and the set $\boldsymbol{\Theta}\subseteq\mathbb{R}^n$ delineates the permissible range for parameters $\theta_i|_{i=1}^n$.

The transfer function matrix is represented by $G(\lambda,\theta)=D^{-1}(\lambda,\theta)\,N(\lambda,\theta)$, and the frequency response is represented by $G(\mathrm{j}\omega,\theta)$. For continuous time systems, $G(\mathrm{j}\omega,\theta)$ is the value of $G(\lambda,\theta)$ at $\lambda=\mathrm{j}\omega$, where $\omega\geqslant 0$; for discrete time systems, $G(\mathrm{j}\omega,\theta)$ is the value of $G(\lambda,\theta)$ at $\lambda=\mathrm{e}^{\mathrm{j}\omega}$, where $\omega\in[0,\pi)$.

It is worth noting that although the numerator and denominator matrices depend on θ in a simple affine structure, the relationship between system outputs and parameters is in general nonlinear and complex due to the inverse of $D(\lambda,\theta)$.

2.2 Assumption and its verification

Throughout this paper, the following assumption is adopted to investigate the identifiability and sloppiness of structured systems with a matrix fraction description.

Assumption 1. The denominator matrix $D(\lambda,\theta)$ is invertible, i.e., the system responds in a deterministic manner to an initial state x_0 and input signal $u(t)|_{t=0}^{\infty}$.

For a matrix fraction model, the invertibility of $D(\lambda,\theta)$ is a prerequisite for its normal operation. It implies that the system's frequency domain model is well-defined, which is necessary to identify system parameters.

Before identification, the specific value of θ is unknown, making it impossible to determine the invertibility of $D(\lambda,\theta)$ through numerical computation. Symbolic computation can be employed for small-scale systems; however, as system size increases, the computational complexity of symbolic computation grows significantly. In fact, by discarding the numerical information of matrices and focusing solely on their structural properties, an effective method for validating Assumption 1 can be derived utilizing graph theory and matroids.

When the rank of $D_i(\lambda)$ $(i=0,1,2,\ldots,n)$ is one, the matrix has a rank-one decomposition, $D_i(\lambda)=\alpha_i(\lambda)\beta_i^{\mathrm{T}}(\lambda)$, where $\alpha_i(\lambda)$ and $\beta_i(\lambda)$ are column vectors. Thus, $D(\lambda,\theta)$ can be expressed as

$$D(\lambda,\theta)=\begin{bmatrix}\alpha_o(\lambda) & \alpha_1(\lambda) & \cdots & \alpha_n(\lambda)\end{bmatrix}\begin{bmatrix}1 & & & & \\ & \theta_1 & & & \\ & & \theta_2 & & \\ & & & \ddots & \\ & & & & \theta_n\end{bmatrix}\begin{bmatrix}\beta_0^{\mathrm{T}}(\lambda)\\ \beta_1^{\mathrm{T}}(\lambda)\\ \vdots \\ \beta_n^{\mathrm{T}}(\lambda)\end{bmatrix}=Q_1(\lambda)T(\theta)Q_2(\lambda), \tag{3}$$

where $Q_1(\lambda)$ and $Q_2(\lambda)$ are known polynomial matrices, and $T(\theta)$ is a diagonal matrix containing unknown variables $\theta_i|_{i=1}^n$. Since $T(\theta)$ is a diagonal matrix with only one non-zero constant, when calculating the rank of $D(\lambda,\theta)$, it can be treated as a general matrix, where the non-zero elements are algebraically independent.

Define a bipartite graph $G=(V^+,V^-;A)$ and two matroids $\boldsymbol{M}^+=(V^+,\mathcal{B}^+)$ and $\boldsymbol{M}^-=(V^-,\mathcal{B}^-)$. Here, vertex sets of the bipartite graph G, defined as V^+ and V^-, respectively, correspond to the row and column sets of the matrix $T(\theta)$, while the arc set is defined based on the zero and non-zero elements of $T(\theta)$: $A=\{(i,j)\mid T_{ij}\neq 0\}$. The matroids $\boldsymbol{M}^+$ and $\boldsymbol{M}^-$ associated with V^+ and V^- are constructed from polynomial matrix $Q_1(\lambda)$ and transpose of $Q_2(\lambda)$, respectively. Specifically, V^+ represents the column set of $Q_1(\lambda)$, while V^- corresponds to the row set of $Q_2(\lambda)$; likewise, $\mathcal{B}^+$ denotes the column basis of $Q_1(\lambda)$, while $\mathcal{B}^-$ forms the row basis of $Q_2(\lambda)$.

The rank of a triple matrix product in (3) is closely related to the independent matching problem, as pointed out by Tomizawa–Iri [23] and Murota [24].

Lemma 1. The generic-rank of $D(\lambda,\theta)$, defined as $\max_{\theta\in\boldsymbol{\Theta}}\mathrm{rank}D(\lambda,\theta)$, matches the size of the largest independent matching in the bipartite graph $G=(V^+,V^-;A)$ and its associated matroids $\boldsymbol{M}^+=(V^+,\mathcal{B}^+)$ and $\boldsymbol{M}^-=(V^-,\mathcal{B}^-)$. In other words, $\mathrm{rank}(D)=\mathrm{rank}(Q_1TQ_2)=\max\{|M|:\text{independent matching}\}$.

The unique structure of $T(\theta)$, whose corresponding bipartite graph is depicted as Figure 1, enables a direct determination of the invertibility of $D(\lambda,\theta)$ without searching for a maximum independent matching. This leads to the following corollary.

Corollary 1. The matrix $D(\lambda,\theta)$ is invertible for almost all $\theta=(\theta_1,\theta_2,\ldots,\theta_n)^{\mathrm{T}}\in\mathbb{R}^n$ when the rank of matrices $D_i(\lambda)$ $(i=0,1,2,\ldots,n)$ is one, if and only if, $Q_1(\lambda)$ is full row rank, $Q_2(\lambda)$ is of full column rank, and the intersection $\mathcal{B}^+\cap\mathcal{B}^-\neq\emptyset$.

The property of independent matching ensures that the corresponding submatrices for $Q_1(\lambda)$ and $Q_2(\lambda)$ are full-rank, while algebraic independence among non-zero elements of $T(\theta)$ guarantees that the determinant of its corresponding submatrix will not cancel out to zero. However, Corollary 1 may not hold true for all parameters $\theta\in\mathbb{R}^n$. Due to properties of graph theory and matroids, there may exist a set of Lebesgue measure zero, within which $D(\lambda,\theta)$ is singular. Moreover, the non-uniqueness of rank-one decomposition does not impede the validity of Corollary 1. This is because different rank-one decompositions only differ by a common polynomial, which does not affect the linear independence between columns of $Q_1(\lambda)$ and rows of $Q_2(\lambda)$.

Remark 1. When rank $D_i(\lambda)>1$, it is still possible to decompose $D(\lambda,\theta)$ into a triple matrix product; however, due to the absence of algebraic independence among non-zero elements in $T(\theta)$, determining whether $D(\lambda,\theta)$ is invertible becomes a combinatorial problem, with computational complexity increasing rapidly as the matrix dimensions and the number of unknown parameters grow. In this case, the determinant of $D(\lambda,\theta)$ remains a polynomial with respect to θ and λ. When it is not identically zero, the values of θ that make the determinant zero form the solution set of a polynomial equation, which has a Lebesgue measure of zero. Consequently, $D(\lambda,\theta)$ is of full rank with probability one or zero with respect to θ. Therefore, in practical applications, if a particular $\theta\in\mathbb{R}^n$ can be found such that $D(\lambda,\theta)$ is

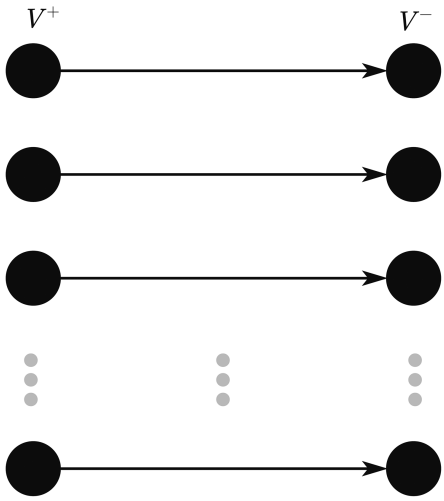

Figure 1 Bipartite graph corresponding to $T(\theta)$.

nonsingular, it can be directly inferred that $D(\lambda,\theta)$ is invertible for almost all $\theta\in\mathbb{R}^n$, without needing to address the aforementioned combinatorial problem. However, the converse may not hold true; if no $\theta\in\mathbb{R}^n$ can be found that makes $D(\lambda,\theta)$ nonsingular, it does not imply that $D(\lambda,\theta)$ is singular for almost all $\theta\in\mathbb{R}^n$.

2.3 Definition of identifiability and sloppiness

Identifiability is an inherent property of models, independent of experimental data and identification algorithms, which must be ensured before any identification experiment. A widely accepted definition of identifiability [6, 7, 25, 26] is as follows.

Definition 1. A system is called locally identifiable at a specific parameter vector value θ, if there exists an ε-neighborhood $\boldsymbol{\Theta}_\varepsilon$ containing θ, which is a subset of $\boldsymbol{\Theta}$, such that for any other parameter vector value $\widetilde{\theta}\in\boldsymbol{\Theta}_\varepsilon$ and any initial state vector $x(0)$, there exists at least one input signal $u(t)|_{t=0}^{\infty}$ such that the output trajectory $y(t,\theta)|_{t=0}^{\infty}$ diverges from the output trajectory $y(t,\tilde{\theta})|_{t=0}^{\infty}$. Otherwise, this system is called locally unidentifiable at θ. If $\boldsymbol{\Theta}_\varepsilon=\boldsymbol{\Theta}$, the system is called globally identifiable or unidentifiable at the parameter vector value θ.

For identifiable systems, the sloppiness can quantitatively measure the difficulty of estimating system parameters from finite frequency responses, as first proposed by Zhou et al. [15].

Definition 2. Supposing $\theta^{[0]}$ is an element of the set $\boldsymbol{\Theta}$ and $w_i|_{i=1}^N$ represents N distinct frequencies, a parameter set $\widehat{\boldsymbol{\Theta}}(\varepsilon,\omega_i|_{i=1}^N,\theta^{[0]})$ is defined as follows:

$$\widehat{\boldsymbol{\Theta}}\left(\varepsilon,\omega_i|_{i=1}^N,\theta^{[0]}\right)=\left\{\theta\,\middle|\,\theta\in\boldsymbol{\Theta},\ \left\|\mathrm{col}\left\{G\left(\mathrm{j}\omega_i,\theta\right)|_{i=1}^N\right\}-\mathrm{col}\left\{G\left(\mathrm{j}\omega_i,\theta^{[0]}\right)\Big|_{i=1}^N\right\}\right\|_F\leqslant\varepsilon\right\},$$

where $\|\cdot\|_F$ stands for the Frobenius norm of a matrix and the dimension of this set is defined as n_θ.

Additionally, a sequence of vectors $\theta_v^{[k]}|_{k=1}^{n_\theta}$ is recursively defined as follows:

$$\theta_v^{[k]}=\arg\max_{\theta\in\widehat{\boldsymbol{\Theta}}(\varepsilon,\omega_i|_{i=1}^N,\theta^{[0]})\setminus\mathrm{span}\{\theta_v^{[l]}|_{l=0}^{k-1}\}}\left\|\theta-\theta^{[0]}\right\|_2,$$

where the initial value $\theta_v^{[0]}=0$ and $\|\cdot\|_2$ stands for the Euclidean norm of a column vector.

For $k=1,2,\ldots,n_\theta-1$, the absolute and relative sloppiness of matrix fraction models at the parameter vector $\theta^{[0]}$ with respect to the frequencies $\omega_i|_{i=1}^N$ are respectively defined as follows:

$$\mathbf{Sm}^{[\mathrm{a}]}\left(\omega_i|_{i=1}^N,\theta^{[0]}\right)=\lim_{\varepsilon\to0}\frac{\|\theta_v^{[1]}-\theta^{[0]}\|_2}{\varepsilon},\ \mathbf{Sm}^{[\mathrm{r},k]}\left(\omega_i|_{i=1}^N,\theta^{[0]}\right)=\lim_{\varepsilon\to0}\frac{\|\theta_v^{[k]}-\theta^{[0]}\|_2}{\|\theta_v^{[k+1]}-\theta^{[0]}\|_2},\tag{4}$$

where the superscript [a] represents absolute sloppiness, $[\mathrm{r},k]$ represents the k-th relative sloppiness.

In fact, the parameter set $\widehat{\mathbf{\Theta}}(\varepsilon, \omega_i|_{i=1}^N, \theta^{[0]})$ is generally non-convex and exhibits a complex structure, corresponding to the hyper-ribbon structure described by Transtrum et al. [11] and Waterfall et al. [12]. The vectors $\theta_v^{[k]} - \theta^{[0]}|_{k=1}^{n_\theta}$ represent the extreme directions of this set, making the physical meaning of the defined sloppiness quite clear. Absolute sloppiness represents the maximum ratio of the magnitude of system parameter variations to the amplitude of system frequency response deviations. In contrast, relative sloppiness represents the ratio of the magnitude of parameter variations in two succeeding extreme cases that result in identical magnitude of frequency response deviations.

The significance of sloppiness is obvious in practical identification processes. For instance, when the maximum amplitude of estimation errors for system frequency responses is known, the absolute sloppiness provides an approximation of the highest estimation accuracy theoretically achievable by an identification algorithm. The relative sloppiness offers an approximate ratio of estimation precisions between two succeeding extreme estimation errors. Alternatively, Ref. [27] implied that the absolute sloppiness must be less than a certain threshold to estimate parameters with specific confidence intervals and identification precision. Moreover, distinct weights can be assigned to various parameters, reflecting different identification precisions for each. Similarly, frequency responses can be assigned varying weights to reflect different measurement accuracies at each frequency point.

2.4 Preliminary results

For the matrix fraction model adopted in this paper, a similar conclusion about the identifiability under finite frequency responses in Zhou et al. [15] can be obtained.

Lemma 2. Under Assumption 1, parameters of the structured system described by (1) and (2) are globally identifiable, if and only if there exists a finite positive integer N and a set of distinct frequency points $\omega_i|_{i=1}^N$, such that for any two different parameter vector values θ and $\widetilde{\theta}$ in the set $\mathbf{\Theta}$, for $i = 1, 2, \ldots, N$, $G(\mathrm{j}\omega_i, \theta) = G(\mathrm{j}\omega_i, \widetilde{\theta})$ simultaneously. Then it implies $\theta = \widetilde{\theta}$.

The following conclusion is helpful to calculate sloppiness, which is stated in Zhou et al. [28], giving a complete parametrization of all solutions to a linear matrix equation.

Lemma 3. A, B, and C are complex matrices with compatible dimensions. There exists a matrix X such that $AXB = C$, if and only if

$$\left(I - AA^\dagger\right) C = 0,\ C\left(I - B^\dagger B\right) = 0. \tag{5}$$

Let Z be an arbitrary matrix of appropriate dimensions. All the matrices satisfying $AXB = C$ can be expressed as follows:

$$X = A^\dagger C B^\dagger + Z - A^\dagger A Z B B^\dagger. \tag{6}$$

The SVD of matrices A and B is respectively defined as follows:

$$A = \begin{bmatrix} U_{A,1} & U_{A,2} \end{bmatrix} \begin{bmatrix} \Sigma_A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{A,1} & V_{A,2} \end{bmatrix}^{\mathrm{H}},\ B = \begin{bmatrix} U_{B,1} & U_{B,2} \end{bmatrix} \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{B,1} & V_{B,2} \end{bmatrix}^{\mathrm{H}},$$

where 0 represents zero matrix with compatible dimensions.

Based on Moore-Penrose inverse, Eqs. (5) and (6) can be further expressed respectively as $U_{A,2}^{\mathrm{H}} C = 0,\ C V_{B,2} = 0$, and $X = V_{A,1}\Sigma_A^{-1} U_{A,1}^{\mathrm{H}} C V_{B,1} \Sigma_B^{-1} U_{B,1}^{\mathrm{H}} + Z - V_{A,1} V_{A,1}^{\mathrm{H}} Z U_{B,1} U_{B,1}^{\mathrm{H}}$.

3 Parameter identifiability

In practical identification processes, acquired input-output data are typically not infinite. This section investigates whether the parameters of structured systems with a matrix fraction description can be uniquely determined by their frequency responses at a finite number of frequencies.

For the transfer function matrix $G\left(\lambda, \theta^{[0]}\right) \in \mathbb{R}^{p\times q}[\lambda]$, which is a rational fraction matrix, there must be a set of right coprime polynomial matrices $\hat{N}(\lambda, \theta^{[0]}) \in \mathbb{R}^{p\times q}[\lambda]$ and $\hat{D}(\lambda, \theta^{[0]}) \in \mathbb{R}^{q\times q}[\lambda]$, such that

$$G(\lambda, \theta^{[0]}) = \hat{N}(\lambda, \theta^{[0]}) \hat{D}^{-1}(\lambda, \theta^{[0]}). \tag{7}$$

Define a sequence of polynomial matrices $X_i|_{i=1}^{n}$ as follows:

$$X_i\left(\lambda,\theta^{[0]}\right)=N_i\left(\lambda\right)\hat{D}\left(\lambda,\theta^{[0]}\right)-D_i\left(\lambda\right)\hat{N}\left(\lambda,\theta^{[0]}\right), \tag{8}$$

where n is the number of independent parameters θ_i in the system.

As Eqs. (1) and (2) are determined by differential equations corresponding to the system, degrees of polynomials in matrices $D_i(\lambda)$ and $N_i(\lambda)$ are finite. Consequently, $X_i\left(\lambda,\theta^{[0]}\right)$ can be further expressed as $X_i\left(\lambda,\theta^{[0]}\right)=\sum_{j=1}^{v_{\max}}X_{ij}\left(\theta^{[0]}\right)\lambda^j$, where $X_{ij}\left(\theta^{[0]}\right)$, a real matrix, is the coefficient matrix of $X_i(\lambda,\theta^{[0]})$ for λ^j, and $v_{\max}$ denotes the maximum degree among all elements in $X_i\left(\lambda,\theta^{[0]}\right)$.

In accordance with the frequency response $G(\mathrm{j}\omega)$ of the system, for continuous-time systems, let $\lambda=\mathrm{j}\omega$, where $\omega\geqslant 0$; for discrete-time systems, let $\lambda=\mathrm{e}^{\mathrm{j}\omega}$, where $\omega\in[0,\pi)$. Substituting it into $X_i(\lambda,\theta^{[0]})$ yields a complex matrix, defined as $X_i(\mathrm{j}\omega,\theta^{[0]})$.

By employing vectorization, a complex polynomial matrix can be defined as $\hat{\Pi}\left(\mathrm{j}\omega,\theta^{[0]}\right)=(\mathrm{vec}\left(X_1\right),\mathrm{vec}\left(X_2\right),\ldots,\mathrm{vec}\left(X_n\right))$, where X_i represents the matrix $X_i\left(\mathrm{j}\omega,\theta^{[0]}\right)$ in an abbreviated form. Separating its real and imaginary parts yields: $\hat{\Pi}\left(\mathrm{j}\omega,\theta^{[0]}\right)=\hat{\Pi}_{\mathrm{r}}\left(\omega,\theta^{[0]}\right)+\mathrm{j}\hat{\Pi}_{\mathrm{j}}\left(\omega,\theta^{[0]}\right)$. Consequently, a real matrix $\Pi(\omega,\theta^{[0]})$ can be defined as $\Pi\left(\omega,\theta^{[0]}\right)=\left[\begin{smallmatrix}\hat{\Pi}_{\mathrm{r}}(\omega,\theta^{[0]})\\ \hat{\Pi}_{\mathrm{j}}(\omega,\theta^{[0]})\end{smallmatrix}\right]$. For a given set of N distinct frequency points $\omega_i|_{i=1}^{N}$, a sequence of real matrices $\Pi\left(\omega_i,\theta^{[0]}\right)|_{i=1}^{N}$ can be obtained, thereby defining a numerical matrix $\bar{\Pi}(\omega_i|_{i=1}^{N},\theta^{[0]})$ as $\bar{\Pi}(\omega_i|_{i=1}^{N},\theta^{[0]})=\mathrm{col}\{\Pi\left(\omega_i,\theta^{[0]}\right)|_{i=1}^{n}\}$.

Given the symbols defined above, the following sufficient and necessary conditions based on matrix rank for frequency-domain identifiability of matrix fraction models can be derived. Its proof is given in Appendix A.

Theorem 1. Under Assumption 1, the structured system with a matrix fraction description described by (1) and (2) is globally identifiable at a specific parameter vector value $\theta^{[0]}$, if and only if there exists a finite positive integer N and a set of distinct frequency points $\omega_i|_{i=1}^{N}$ such that the numerical matrix $\bar{\Pi}(\omega_i|_{i=1}^{N},\theta^{[0]})$ is of full column rank.

The right coprime decomposition of $G(\lambda,\theta^{[0]})$ given by (7) is not unique. However, this non-uniqueness does not affect the validity of Theorem 1. Specifically, consider any unimodular polynomial matrix $W(\lambda)\in\mathbb{R}^{q\times q}[\lambda]$, and define two polynomial matrices $\tilde{N}(\lambda,\theta^{[0]})=\hat{N}(\lambda,\theta^{[0]})W(\lambda)$ and $\tilde{D}(\lambda,\theta^{[0]})=\hat{D}(\lambda,\theta^{[0]})W(\lambda)$. These can represent all possible right coprime decompositions of $G(\lambda,\theta^{[0]})$ [18, 29]. The non-uniqueness of the right coprime decomposition impacts (A3), as detailed below:

$$\sum_{i=1}^{n}\delta_{\theta_i}\left[N_i\left(\lambda\right)\hat{D}\left(\lambda,\theta^{[0]}\right)W\left(\lambda\right)-D_i\left(\lambda\right)\hat{N}\left(\lambda,\theta^{[0]}\right)W\left(\lambda\right)\right]=0\Leftrightarrow\sum_{i=1}^{n}\delta_{\theta_i}\bar{X}_i\left(\lambda,\theta^{[0]}\right)=0, \tag{9}$$

where $\bar{X}_i\left(\lambda,\theta^{[0]}\right)=X_i\left(\lambda,\theta^{[0]}\right)W\left(\lambda\right)$. Given that $W(\lambda)$ is an unimodal matrix and the right side of (9) equals zero, the value of $W(\lambda)$ does not influence the condition where Eq. (9) has only a zero solution.

Let $r(\bar{\Pi})$ denote the dimension of the right null space basis of $\bar{\Pi}(\omega_i|_{i=1}^{N},\theta^{[0]})$. Moreover, a set $\mathscr{U}\left(\theta^{[0]}\right)$ is defined as follows: $\mathscr{U}(\theta^{[0]})=\{\theta|\ =\theta^{[0]}+\bar{\Pi}^{\perp}\gamma\}$, where $\bar{\Pi}^{\perp}$ is a basis for the right null space of the matrix $\bar{\Pi}(\omega_i|_{i=1}^{N},\theta^{[0]})$ and γ is an arbitrary vector in $\mathbb{R}^{r(\bar{\Pi})}$. Evidently, this set includes $\theta^{[0]}$ as an element. According to Theorem 1, the following corollary can be obtained.

Corollary 2. Under Assumption 1, when the structured system with a matrix fraction description defined by (1) and (2) is unidentifiable at a specific parameter vector value $\theta^{[0]}$, the set $\mathscr{U}\left(\theta^{[0]}\right)$ represents the unidentifiable region of the system at the specific parameter $\theta^{[0]}$.

Corollary 2 reveals that, at a specific parameter vector value $\theta^{[0]}$, the global identifiability of structured systems with a matrix fraction description is equivalent to their local identifiability.

Remark 2. The proof of Theorem 1 reveals that the solution space corresponding to (A6) is essentially the intersection of solution spaces for (A5) at each frequency point ω_i. By gradually increasing the number of frequency points, the solution space related to (A6) progressively reduces until it diminishes to null space, thereby verifying the identifiability of the system. In other words, when frequencies $\omega_i|_{i=1}^{N}$ are given, the condition of Theorem 1, that is, whether $\bar{\Pi}(\omega_i|_{i=1}^{N},\theta^{[0]})$ is of full column rank, can be verified recursively. It is pretty computationally appealing, especially when the number of frequency points N is large and/or the dimension of the system is high.

However, determining the number and value of frequency points is still a significant problem. Ordinarily, the number of frequency points determined according to the order of systems is relatively large [6], especially when the order of the transfer function matrix is relatively high, which is a common situation in large-scale dynamic network systems. Zhou et al. [15] proposed an iterative algorithm to identify as few frequency points as possible, whose frequency responses ensure the system's identifiability. Utilizing Theorem 1 and the Smith form of polynomial matrices, this algorithm can be efficiently performed under the matrix fraction model adopted in this paper, and the desired frequency points can be obtained within n iterations at most. It typically results in significantly fewer frequencies than those obtained directly from system orders.

4 Computation of the sloppiness

Identifiability only provides a binary answer: yes or no. For identifiable models, sloppiness can further quantify the difficulty of parameter identification. In this section, under the premise of identifiability, based on the definition of sloppiness in Zhou et al. [15], explicit formulas for calculating absolute and relative sloppiness of matrix fraction models are given, which quantitatively measure the difficulty of identifying system parameters from finite frequency responses.

To address this problem, let $\theta^{[0]}$ be a specific element in the set $\boldsymbol{\Theta}$, and $\omega_i|_{i=1}^N$ represent a set of distinct frequencies. In this section, the following two assumptions are adopted.

Assumption 2. The parameter set $\boldsymbol{\Theta}$ is open, convex, and includes 0 as an element.

Assumption 3. The system, at a specific parameter vector value $\theta^{[0]}$, is globally identifiable with respect to the given N distinct frequency points $\omega_i|_{i=1}^N$ and corresponding system frequency responses. Both $\theta^{[0]}$ and $\omega_i|_{i=1}^N$ are considered to be known.

Similarly to $X_i(\mathrm{j}\omega, \theta^{[0]})$, complex polynomial matrices $\hat{N}\left(\mathrm{j}\omega, \theta^{[0]}\right)$, $\hat{D}\left(\mathrm{j}\omega, \theta^{[0]}\right)$, $N\left(\mathrm{j}\omega, \theta^{[0]}\right)$, and $D(\mathrm{j}\omega, \theta^{[0]})$ can be obtained. Subscripts r and j represent the real and imaginary parts of the corresponding complex matrix, respectively. Based on the symbols above, two polynomial matrices are defined as follows:

$$\mathcal{P}_{\mathrm{r}}\left(\omega, \theta^{[0]}\right) = \left(\hat{D}_{\mathrm{r}}^{\mathrm{T}} \otimes [D_{\mathrm{r}}, -D_{\mathrm{j}}] \ -\hat{D}_{\mathrm{j}}^{\mathrm{T}} \otimes [D_{\mathrm{j}}, D_{\mathrm{r}}]\right), \ \mathcal{P}_{\mathrm{j}}\left(\omega, \theta^{[0]}\right) = \left(\hat{D}_{\mathrm{r}}^{\mathrm{T}} \otimes [D_{\mathrm{j}}, D_{\mathrm{r}}] \ \hat{D}_{\mathrm{j}}^{\mathrm{T}} \otimes [D_{\mathrm{r}}, -D_{\mathrm{j}}]\right),$$

where the variable $(\omega, \theta^{[0]})$ is omitted for brevity in the formula. In addition, a polynomial matrix is defined: $\mathcal{P}\left(\omega, \theta^{[0]}\right) = \mathrm{diag}\left\{\mathcal{P}_{\mathrm{r}}, \mathcal{P}_{\mathrm{j}}\right\}$. For a set of distinct frequency points $\omega_i|_{i=1}^N$, a real matrix can be obtained: $\bar{\mathcal{P}}(\omega_i|_{i=1}^N, \theta^{[0]}) = \mathrm{diag}\{\mathcal{P}\left(\omega_i, \theta^{[0]}\right)\big|_{i=1}^N\}$.

The SVD of $\bar{\Pi}(\omega_i|_{i=1}^N, \theta^{[0]})$ is presented as follows:

$$\bar{\Pi}\left(\omega_i|_{i=1}^N, \theta^{[0]}\right) = U_{\bar{\Pi}} \begin{bmatrix} \Sigma_{\bar{\Pi}} & 0 \\ 0 & 0 \end{bmatrix} V_{\bar{\Pi}}^{\mathrm{T}} = \begin{bmatrix} U_{\bar{\Pi},1} & U_{\bar{\Pi},2} \end{bmatrix} \begin{bmatrix} \Sigma_{\bar{\Pi}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{\bar{\Pi},1} & V_{\bar{\Pi},2} \end{bmatrix}^{\mathrm{T}} = U_{\bar{\Pi},1} \Sigma_{\bar{\Pi}} V_{\bar{\Pi},1}^{\mathrm{T}}.$$

Let $U_{\bar{\Pi},2}^{\mathrm{T}} \bar{\mathcal{P}}(\omega_i|_{i=1}^N, \theta^{[0]})$ as $\Psi(\omega_i|_{i=1}^N, \theta^{[0]})$, which has n columns. Applying SVD to it gives

$$\begin{bmatrix} U_{\Psi,1} & U_{\Psi,2} \end{bmatrix} \begin{bmatrix} \Sigma_{\Psi} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{\Psi,1} & V_{\Psi,2} \end{bmatrix}^{\mathrm{T}},$$

where $V_{\Psi,2}^{\mathrm{T}} V_{\Psi,2} = I$. Then, a matrix is defined as follows:

$$S_\theta\left(\omega_i|_{i=1}^N, \theta^{[0]}\right) = V_{\bar{\Pi}} \Sigma_{\bar{\Pi}}^{-1} U_{\bar{\Pi},1}^{\mathrm{T}} \bar{\mathcal{P}}\left(\omega_i|_{i=1}^N, \theta^{[0]}\right) V_{\Psi,2}. \tag{10}$$

Using the symbols defined above, the following parameter set is obtained, which includes all parameters whose corresponding system frequency responses deviate from those of $\theta^{[0]}$ within a specific range.

Theorem 2. ε is a sufficiently small positive real number, and a parameter set is defined as follows:

$$\overline{\boldsymbol{\Theta}}\left(\varepsilon, \omega_i|_{i=1}^N, \theta^{[0]}\right) = \left\{\theta \left| \sum_{i=1}^N \left\| G\left(\mathrm{j}\omega_i, \theta\right) - G\left(\mathrm{j}\omega_i, \theta^{[0]}\right) \right\|_F^2 \leqslant \varepsilon^2 \right.\right\}. \tag{11}$$

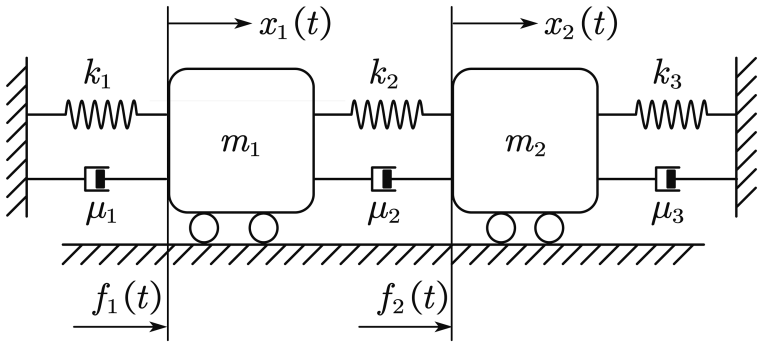

Figure 2 Spring-damping system with two carts.

Under Assumptions 1–3, this set can be equivalently expressed as

$$\overline{\boldsymbol{\Theta}}\left(\varepsilon, \omega_i|_{i=1}^N, \theta^{[0]}\right) = \left\{\theta \mid \theta = \theta^{[0]} + S_\theta\left(\omega_i|_{i=1}^N, \theta^{[0]}\right)\xi + O\left(\|\xi\|_2^2\right), \xi^{\mathrm{T}}\xi \leqslant \varepsilon^2, \xi \in \mathbb{R}^n\right\}. \tag{12}$$

The proof of this theorem is given in Appendix B.

It is worth noting that the non-uniqueness of the right coprime decomposition of $G\left(\mathrm{j}\omega_i, \theta^{[0]}\right)$ does not affect the conclusion of Theorem 2. The proof is fairly straightforward and is omitted here for brevity.

Define a set as $\tilde{\boldsymbol{\Theta}}(\varepsilon, \omega_i|_{i=1}^N, \theta^{[0]}) = \left\{\xi \mid \xi \in \mathbb{R}^n, \xi^{\mathrm{T}} V_{\Psi,2}^{\mathrm{T}} V_{\Psi,2}\xi = \xi^{\mathrm{T}}\xi \leqslant \varepsilon^2\right\}$. For any positive ε, this set is always an ellipsoid, which is a highly attractive property. This feature enables the derivation of explicit computation for absolute and relative sloppiness under relatively weak conditions via the set $\tilde{\boldsymbol{\Theta}}(\varepsilon, \omega_i|_{i=1}^N, \theta^{[0]})$. More specifically, the following theorem is derived.

Theorem 3. Under Assumptions 1–3, let the k-th largest singular value of $S_\theta(\omega_i|_{i=1}^N, \theta^{[0]})$ be denoted as $\mu^{[k]}$, $k = 1, 2, \ldots, n$. The explicit expressions for absolute and relative sloppiness at the specific parameter $\theta^{[0]}$ and frequency points $\omega_i|_{i=1}^N$ are as follows:

$$\mathbf{Sm}^{[\mathrm{a}]}\left(\omega_i|_{i=1}^N, \theta^{[0]}\right) = \mu^{[1]}, \ \mathbf{Sm}^{[\mathrm{r},k]}\left(\omega_i|_{i=1}^N, \theta^{[0]}\right) = \frac{\mu^{[k]}}{\mu^{[k+1]}}.$$

The proof of this theorem is provided in Appendix C.

Remark 3. According to practical identification needs, the vector and matrix norms can adopt varying definitions, such as requiring the estimation error of system frequency responses to be bounded at each frequency in its maximum singular value, which is widely employed in robust control and estimation theory [28, 29]. Although all norms for finite-dimensional vectors/matrices are equivalent [28, 30, 31], different norms typically result in different values of absolute/relative sloppiness. Moreover, during computations, the parameter set corresponding to $\tilde{\boldsymbol{\Theta}}(\varepsilon, \omega_i|_{i=1}^N, \theta^{[0]})$ may no longer constitute an ellipsoid, making explicit expressions for sloppiness unobtainable and resulting in an optimization problem. For further details on sloppiness computation with other vector/matrix norms, please refer to [15].

5 Numerical simulations

To validate the calculation method of absolute and relative sloppiness presented in Section 4, and to highlight the differences between the proposed sloppiness metrics and FIM, a spring-damping system is used as an example for comparison, which has wide applications in engineering, physics, and other fields [32].

To simplify computation without loss of generality, the situation with only two carts is considered, as depicted in Figure 2. The corresponding calculations can be naturally extended to systems with N carts. The system dynamics is described by the following differential equations:

$$\begin{aligned} m_1\frac{\mathrm{d}^2x_1(t)}{\mathrm{d}t^2} &= f_1(t) - \mu_1\frac{\mathrm{d}x_1(t)}{\mathrm{d}t} - k_1x_1(t) - \mu_2\frac{\mathrm{d}\left(x_1(t) - x_2(t)\right)}{\mathrm{d}t} - k_2\left(x_1(t) - x_2(t)\right), \\ m_2\frac{\mathrm{d}^2x_2(t)}{\mathrm{d}t^2} &= f_2(t) - \mu_2\frac{\mathrm{d}\left(x_2(t) - x_1(t)\right)}{\mathrm{d}t} - k_2\left(x_2(t) - x_1(t)\right) - \mu_3\frac{\mathrm{d}x_2(t)}{\mathrm{d}t} - k_3x_2(t), \end{aligned} \tag{13}$$

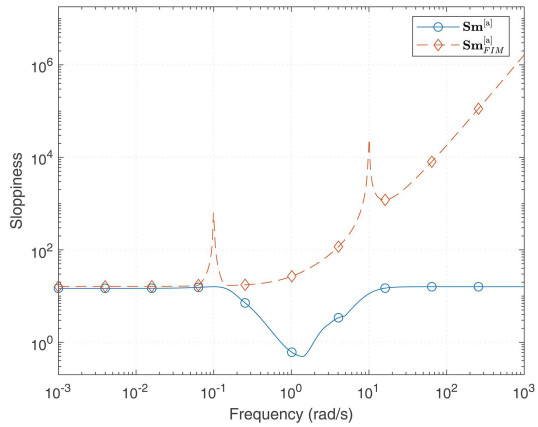

Figure 3 (Color online) Absolute sloppiness and FIM-based absolute sloppiness.

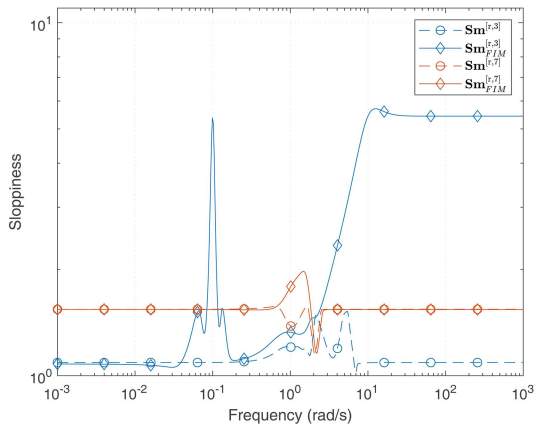

Figure 4 (Color online) Relative sloppiness and FIM-based relative sloppiness.

where the system inputs are external forces $f_1(t)$ and $f_2(t)$ applied to two carts, respectively, and the outputs are the carts' displacements, $x_1(t)$ and $x_2(t)$. The unknown parameters of the structured system in Figure 2 are defined as $\theta = (m_1, m_2, \mu_1, \mu_2, \mu_3, k_1, k_2, k_3)^{\mathrm{T}}$, with $n = 8$. According to (13), the rank of matrices $D_i(\lambda)\,(i = 0, 1, 2, \ldots, n)$ is one. Thus, Corollary 1 ascertains that $D(\lambda, \theta)$ is invertible for almost all $\theta = (\theta_1, \theta_2, \ldots, \theta_n)^{\mathrm{T}} \in \mathbb{R}^n$. Let $\theta^{[0]} = (0.3, 0.4, 0.2, 0.2, 0.4, 0.5, 0.2, 0.4)^{\mathrm{T}}$, the system outputs do not diverge, indicating that $\theta^{[0]}$ is reasonable.

To highlight the difference between FIM and the proposed sloppiness matrices, let $\lambda_{\mathrm{FIM}}^{[k]}$, where $1 \leqslant k \leqslant n$, denote the k-th largest eigenvalue of FIM. According to [11, 12, 14], FIM-based absolute and relative sloppiness metrics are detailed as follows:

$$\mathbf{Sm}_{\mathrm{FIM}}^{[\mathrm{a}]}\left(\omega_i|_{i=1}^{N}, \theta^{[0]}\right) = 1/\sqrt{\lambda_{\mathrm{FIM}}^{[n]}},\ \mathbf{Sm}_{\mathrm{FIM}}^{[\mathrm{r},k]}\left(\omega_i|_{i=1}^{N}, \theta^{[0]}\right) = \sqrt{\lambda_{\mathrm{FIM}}^{[k]}/\lambda_{\mathrm{FIM}}^{[k+1]}},$$

where $k = 1, 2, \ldots, n-1$.

According to Remark 2, the identifiability can be ensured with two or more distinct non-zero frequency points and their corresponding frequency responses, using the frequency selection approach proposed by Zhou et al. [15]. However, FIM requires at least three distinct frequency points to be well defined. Therefore, three frequency points have been selected for comparison with the FIM. To facilitate plotting, two frequency points are set in the low and high-frequency ranges: $\omega_1 = 0.1$ rad/s and $\omega_2 = 10$ rad/s. The third frequency point ω_3 is logarithmically sampled at 500 points within the frequency band 1×10^{-3}–1×10^{3} rad/s. The calculation results are presented in Figures 3–5. According to Definition 2, both absolute and relative sloppiness are dimensionless, as are FIM-based absolute and relative sloppiness.

Figure 3 shows that absolute sloppiness and FIM-based absolute sloppiness provide similar information in the low-frequency range. However, significant differences are observed in mid and high-frequency ranges. The two peaks of FIM-based absolute sloppiness correspond to the fixed frequency points $\omega_1 = 0.1$ rad/s and $\omega_2 = 10$ rad/s, with their values reaching infinity. As discussed earlier, FIM requires three distinct frequency points to be well-defined, and as ω_3 approaches the values of ω_1 and ω_2, the value of FIM-based absolute sloppiness increases. The trend of absolute sloppiness in Figure 3 indicates that frequency points ω_1 and ω_2 prevent absolute sloppiness from increasing in the low and high-frequency ranges.

For graphical clarity, the 3rd and 7th relative sloppiness and FIM-based relative sloppiness were chosen for comparison. Figure 4 shows significant differences between relative sloppiness and FIM-based relative sloppiness, indicating they provide different information about extreme directions of the system's parameter variations that lead to identical deviations in the magnitude of the system's frequency responses. Specifically, the 3rd exhibits significant discrepancies in mid and high frequencies, especially in high frequencies, consistent with Figure 3, while the 7th shows considerable differences at 1 rad/s.

Additionally, sloppiness is usually associated with the condition number of FIM. Hence, there is a comparison between the product of relative sloppiness and the condition number of FIM. In Figure 5, the trends of curves in low and high-frequency ranges align with Figure 3; however, in the mid-frequency

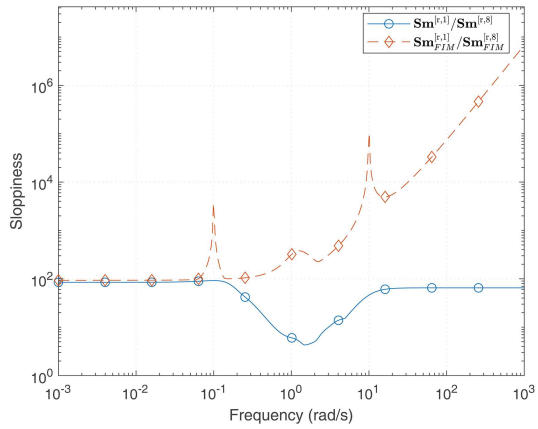

Figure 5 (Color online) Product of relative sloppiness and condition number of FIM.

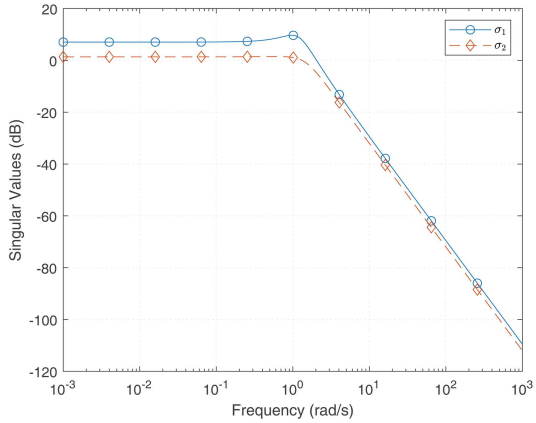

Figure 6 (Color online) Singular values of frequency responses.

range, especially at 1 rad/s, there are significant numerical differences and opposing trends, with one exhibiting a peak and the other a trough.

Figure 6 illustrates the variation of singular values of system frequency responses with frequency. It exhibits significant fluctuations and relatively large amplitudes within the 0.1–10 rad/s frequency range. The input-output data in this frequency range are generally believed to reveal the most about the system's internal structure. Moreover, practical engineering experience suggests that, for physical systems, mid-frequency data usually carries more information, while low and high-frequency data is less. The trends of sloppiness curves in Figures 3 and 5 align more closely with the observations above. Hence, there is a reason to believe that the proposed sloppiness matrices provide more precise information than FIM, especially within the 0.1–10 rad/s range. This is similar to the situation mentioned in [14], where the condition number of FIM yielded opposite results, particularly at 1 rad/s.

Remark 4. From a definition perspective, the suggested matrices in Definition 2 are more suitable for analyzing system sloppiness than the widely adopted FIM. In Definition 2, the magnitudes of frequency response deviations are confined within a slight range, indicating that the definitions of sloppiness apply when system frequency responses deviate only slightly from the original after parameter changes. However, slight frequency response deviations do not necessarily imply minor corresponding parameter variations. In other words, the definitions of sloppiness can help detect whether there are substantial parameter variations that lead to minor system frequency response deviations, as well as whether there are significant magnitude differences in parameter variations that result in the same magnitude of system frequency response deviations. Compared to Definition 2, FIM restricts the magnitudes of parameter variations within a tiny range. It can help detect whether there are minor system parameter variations that cause significant changes in system outputs, which is more suitable for measuring the sensitivity of system outputs to parameter variations.

Furthermore, it is observed that the number and values of frequency points can influence the magnitude of sloppiness. Figure 7 illustrates the impact of the number of frequency points on absolute sloppiness by adding fixed frequency points $\omega_4 = 0.5$ rad/s and $\omega_5 = 5$ rad/s. As the number of frequency points increases, the magnitude of absolute sloppiness decreases, indicating that more frequency points generally provide more information about system parameters. In other words, increasing frequency points helps reduce sloppiness and improve estimation precision. However, selecting the appropriate additional frequency points is essential to maximize this improvement. Figure 8 depicts how different combinations of frequency point values affect absolute sloppiness by changing the fixed frequency points ω_1 and ω_2 within the low, mid, and high-frequency ranges. With the same number of frequency points, absolute sloppiness is lowest for the mid-frequency and highest for the high-frequency. It indicates that for the system in Figure 2, frequency responses in the mid-frequency range provide the most information, consistent with Figures 3 and 5.

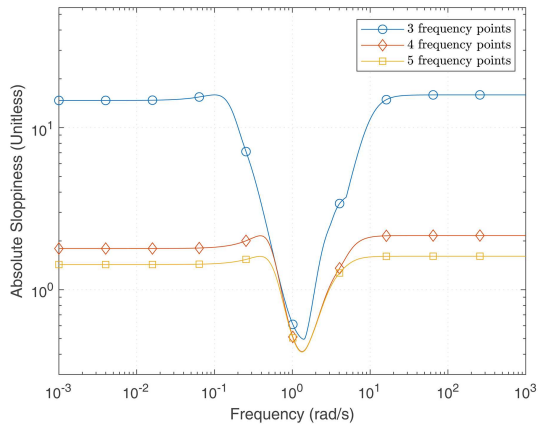

Figure 7 (Color online) Absolute sloppiness with various numbers of frequency points. ○: ω_1–ω_3; ◇: ω_1–ω_4; □: ω_1–ω_5.

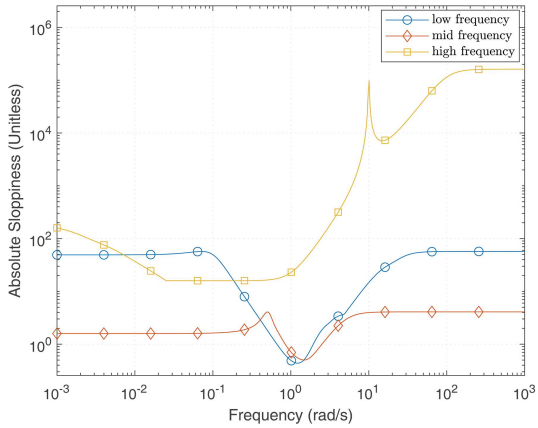

Figure 8 (Color online) Absolute sloppiness with various combinations of frequency points. ○: $\omega_1 = 0.01$ rad/s, $\omega_2 = 0.1$ rad/s; ◇: $\omega_1 = 0.5$ rad/s, $\omega_2 = 5$ rad/s; □: $\omega_1 = 10$ rad/s, $\omega_2 = 100$ rad/s.

6 Concluding remarks

This paper investigates the identifiability of structured systems using finite frequency responses with a matrix fraction description and proposes a method for quantitatively calculating system sloppiness. Building on the definition of sloppiness introduced in [15], this paper extends the analysis of system sloppiness to the matrix fraction model. Specifically, by considering the matrix fraction model instead of the LFT-described descriptor system, this paper not only allows for more lenient assumptions on the system but also simplifies the computation of sloppiness. First, an assumption for the problem is provided, which is quite lenient. By employing graph theory and matroids, a condition that the assumption is valid for almost all parameters is derived when the system-related matrices are rank-one. Second, under the assumption that the denominator polynomial matrix is invertible, a global identifiability criterion based on the rank of a numerical matrix is given, so that the parameters can be uniquely determined from finite system frequency responses. The criterion can be recursively verified with each frequency point, which is computationally attractive in the analysis and synthesis of large-scale systems. Moreover, employing this criterion can further determine the number and values of ideal frequency points, whose frequency responses ensure the system's identifiability. Last, utilizing properties of matrix fraction models and linear system theory, an ellipsoidal approximation of a parameter set is provided, including all parameters whose corresponding system frequency responses deviate within a prescribed range from those at a specific parameter value. As a result, analytical expressions of absolute and relative sloppiness are derived. It is worth mentioning that all the conclusions in this paper can be easily extended to the analysis and synthesis of time-delay systems without any difficulty in formula derivation. However, further efforts are still required to determine the ideal number and values of frequency points based on sloppiness requirements to achieve the desired estimation accuracy.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 62127809, 62373212, 61733008, 52061635102).

References

1 Zhang K, Zhu H, Guo S. Dependency analysis and improved parameter estimation for dynamic composite load modeling. IEEE Trans Power Syst, 2016, 32: 3287–3297
2 Nabavi S, Chakrabortty A. A graph-theoretic condition for global identifiability of weighted consensus networks. IEEE Trans Automat Contr, 2016, 61: 497–502
3 Xia X, Moog C H. Identifiability of nonlinear systems with application to HIV/AIDS models. IEEE Trans Automat Contr, 2003, 48: 330–336
4 Sun D, Hwang I, Corless M. Controlled mode distinguishability for cybersecurity. IEEE Control Syst Lett, 2021, 6: 998–1003
5 Battistelli G, Tesi P. Switching control for parameter identifiability. IEEE Trans Automat Contr, 2017, 63: 849–856
6 Ljung L. System Identification: Theory for the User. 2nd ed. Upper Saddle River: Prentice-Hall,1999
7 Pintelon R, Schoukens J. System Identification: A Frequency Domain Approach. Hoboken: John Wiley & Sons, 2012
8 Anstett-Collin F, Denis-Vidal L, Millérioux G. A priori identifiability: an overview on definitions and approaches. Annu Rev Control, 2020, 50: 139–149
9 Wieland F G, Hauber A L, Rosenblatt M, et al. On structural and practical identifiability. Curr Opin Syst Biol, 2021, 25: 60–69

10 Guillaume J H A, Jakeman J D, Marsili-Libelli S, et al. Introductory overview of identifiability analysis: a guide to evaluating whether you have the right type of data for your modeling purpose. Environ Model Software, 2019, 119: 418–432
11 Transtrum M K, Machta B B, Brown K S, et al. Perspective: sloppiness and emergent theories in physics, biology, and beyond. J Chem Phys, 2015, 143: 010901
12 Waterfall J J, Casey F P, Gutenkunst R N, et al. Sloppy-model universality class and the Vandermonde matrix. Phys Rev Lett, 2006, 97: 150601
13 Raman D V, Anderson J, Papachristodoulou A. Delineating parameter unidentifiabilities in complex models. Phys Rev E, 2017, 95: 032314
14 Chis O T, Villaverde A F, Banga J R, et al. On the relationship between sloppiness and identifiability. Math Biosci, 2016, 282: 147–161
15 Zhou T, Li Y, Yin K. Frequency domain identifiability and sloppiness of descriptor systems with an LFT structure. Automatica, 2024, 159: 111362
16 Rugh W J. Linear System Theory. Upper Saddle River: Prentice-Hall, Inc., 1996
17 Bellman R, Åström K J. On structural identifiability. Math Biosci, 1970, 7: 329–339
18 Kailath T. Linear Systems. Englewood Cliffs: Prentice-Hall, 1980
19 Verhaegen M, Verdult V. Filtering and System Identification: A Least Squares Approach. Cambridge: Cambridge University Press, 2007
20 Dankers A G. System identification in dynamic networks. Dissertation for Ph.D. Degree. Delft: Delft University of Technology, 2014
21 Shieh L S, Tsay Y T, Barnett S. A review of some matrix continued-fraction descriptions and their applications to the stability of multivariable control systems. IMA J Math Control Info, 1985, 2: 1–23
22 Arambel P O, Tadmor G. Identifiability and persistent excitation in full matrix fraction parameter estimation. Automatica, 1997, 33: 689–692
23 Tomizawa N, Iri M. An algorithm for determining the rank of a triple matrix product AXB with application to the problem of discerning the existence of the unique solution in a network (in Japanese). Trans Inst Electr Comm Engin Japan, 1974, 57: 834–841; English translation: Electronics and Communications in Japan, 1974, 57: 50–57
24 Murota K. Matrices and Matroids for Systems Analysis. Berlin: Springer, 2010
25 Zhou T, Yin K. Global structure identifiability and reconstructibility of an NDS with descriptor subsystems. Automatica, 2022, 142: 110356
26 Zhou T. Structure identifiability of an NDS with LFT parametrized subsystems. IEEE Trans Automat Contr, 2022, 67: 6682–6697
27 Jedra Y, Proutiere A. Finite-time identification of linear systems: fundamental limits and optimal algorithms. IEEE Trans Automat Contr, 2022, 68: 2805–2820
28 Zhou T, You K, Tao L. Estimation and Control of Large-scale Networked Systems. Oxford: Butterworth-Heinemann, 2018
29 Khalil I S, Doyle J C, Glover K. Robust and Optimal Control. Upper Saddle River: Prentice Hall, 1996
30 Golub G H, van Loan C F. Matrix Computations. Baltimore: Johns Hopkins University Press, 2013
31 Hiriart-Urruty J B, Lemaréchal C. Fundamentals of Convex Analysis. Berlin: Springer Science & Business Media, 2004
32 Kibble T, Berkshire F H. Classical Mechanics. Singapore: World Scientific Publishing Company, 2004

Appendix A Proof of Theorem 1

A structured system is identifiable at a specific parameter vector value $\theta^{[0]}$, if and only if, for any $\delta_\theta = \left(\delta_{\theta_1}, \delta_{\theta_2}, \ldots, \delta_{\theta_n}\right)^{\mathrm{T}} \in \mathbb{R}^n$ where $\delta_\theta \neq 0$, the following equation has no solution:

$$G\left(\lambda, \theta^{[0]} + \delta_\theta\right) = G\left(\lambda, \theta^{[0]}\right). \tag{A1}$$

According to the matrix fraction model defined by (1) and (2), the transfer function can be written as $G\left(\lambda, \theta^{[0]} + \delta_\theta\right) = D^{-1}\left(\lambda, \theta^{[0]} + \delta_\theta\right) N\left(\lambda, \theta^{[0]} + \delta_\theta\right)$, where the numerator and denominator matrices can be expressed as follows:

$$\begin{aligned} D\left(\lambda, \theta^{[0]} + \delta_\theta\right) &= D_0(\lambda) + \sum_{i=1}^{n}\left(\theta_i^{[0]} + \delta_{\theta_i}\right) D_i(\lambda) = D\left(\lambda, \theta^{[0]}\right) + \sum_{i=1}^{n} \delta_{\theta_i} D_i(\lambda), \\ N\left(\lambda, \theta^{[0]} + \delta_\theta\right) &= N_0(\lambda) + \sum_{i=1}^{n}\left(\theta_i^{[0]} + \delta_{\theta_i}\right) N_i(\lambda) = N\left(\lambda, \theta^{[0]}\right) + \sum_{i=1}^{n} \delta_{\theta_i} N_i(\lambda). \end{aligned}$$

Hence, Eq. (A1) can be further refined as follows:

$$G(\lambda, \theta^{[0]} + \delta_\theta) = G(\lambda, \theta^{[0]}) \Leftrightarrow \sum_{i=1}^{n} \delta_{\theta_i}\left[N_i(\lambda) - D_i(\lambda) G(\lambda, \theta^{[0]})\right] = 0. \tag{A2}$$

Utilizing the right coprime decomposition of $G(\lambda, \theta^{[0]})$ in (7), Eq. (A2) can be derived as follows:

$$\begin{aligned} &\sum_{i=1}^{n} \delta_{\theta_i}\left[N_i(\lambda) - D_i(\lambda) G(\lambda, \theta^{[0]})\right] = 0 \\ &\quad \Leftrightarrow \sum_{i=1}^{n} \delta_{\theta_i}\left[N_i(\lambda) \hat{D}\left(\lambda, \theta^{[0]}\right) - D_i(\lambda) \hat{N}\left(\lambda, \theta^{[0]}\right)\right] = 0 \\ &\quad \Leftrightarrow \sum_{i=1}^{n} \delta_{\theta_i} X_i\left(\lambda, \theta^{[0]}\right) = 0. \end{aligned} \tag{A3}$$

By vectorization, Eq. (A3) can be expressed as matrix multiplication:

$$\sum_{i=1}^{n} \delta_{\theta_i} X_i\left(\lambda, \theta^{[0]}\right) = 0 \Leftrightarrow \left[\mathrm{vec}\left(X_1\left(\lambda, \theta^{[0]}\right)\right), \mathrm{vec}\left(X_2\left(\lambda, \theta^{[0]}\right)\right), \ldots, \mathrm{vec}\left(X_n\left(\lambda, \theta^{[0]}\right)\right)\right] \delta_\theta = 0.$$

From Lemma 2, system parameters are identifiable if and only if, for a given set of distinct frequency points $\omega_i|_{i=1}^N$, there exists solely a zero vector δ_θ which simultaneously satisfies the following equation for $i=1,2,\ldots,N$:

$$\hat{\Pi}\left(\mathrm{j}\omega_i,\theta^{[0]}\right)\delta_\theta=0. \tag{A4}$$

For complex numbers, the real and imaginary parts are relatively independent. Two complex numbers are equal if and only if their real and imaginary parts are respectively equivalent. This also applies to complex matrices. Thus, by separating real and imaginary parts of the complex matrix in (A4), the following equation holds concurrently for $\omega_i|_{i=1}^N$:

$$\left[\begin{array}{c}\hat{\Pi}_{\mathrm{r}}\left(\mathrm{j}\omega_i,\theta^{[0]}\right)\\ \hat{\Pi}_{\mathrm{j}}\left(\mathrm{j}\omega_i,\theta^{[0]}\right)\end{array}\right]\delta_\theta=0\Leftrightarrow \mathit{\Pi}\left(\omega_i,\theta^{[0]}\right)\delta_\theta=0. \tag{A5}$$

Thus, ultimately obtain

$$\mathrm{col}\left\{\mathit{\Pi}\left(\omega_i,\theta^{[0]}\right)\Big|_{i=1}^n\right\}\delta_\theta=0\Leftrightarrow\bar{\Pi}\left(\omega_i|_{i=1}^N,\theta^{[0]}\right)\delta_\theta=0. \tag{A6}$$

Hence, system parameters are identifiable if and only if $\bar{\Pi}(\omega_i|_{i=1}^N,\theta^{[0]})$ is of full column rank.

Appendix B Proof of Theorem 2

For a given $\theta^{[0]}\in\boldsymbol{\Theta}$ and any $\theta=\mathrm{col}\left\{\theta_k|_{k=1}^n\right\}\in\boldsymbol{\Theta}$, a real vector is defined as $\delta_\theta=\theta-\theta^{[0]}$. For a given set of N distinct frequency points $\omega_i|_{i=1}^N$, a sequence of complex matrices $\Delta G\left(\mathrm{j}\omega_i,\theta\right)|_{i=1}^N$ is defined as $\Delta G\left(\mathrm{j}\omega_i,\theta\right)=G\left(\mathrm{j}\omega_i,\theta\right)-G\left(\mathrm{j}\omega_i,\theta^{[0]}\right)$. Therefore, it can be shown that for $i=1,2,\ldots,N$, the following holds:

$$D^{-1}\left(\mathrm{j}\omega_i,\theta^{[0]}+\delta_\theta\right)N\left(\mathrm{j}\omega_i,\theta^{[0]}+\delta_\theta\right)-D^{-1}\left(\mathrm{j}w_i,\theta^{[0]}\right)N\left(\mathrm{j}w_i,\theta^{[0]}\right)=\Delta G\left(\mathrm{j}\omega_i,\theta\right).$$

According to (2) and the existence of $\Delta G\left(\mathrm{j}\omega_i,\theta\right)$, it is derived that

$$\sum_{k=1}^n\left[N_k\left(\mathrm{j}\omega_i\right)-D_k\left(\mathrm{j}\omega_i\right)G\left(\mathrm{j}\omega_i,\theta^{[0]}\right)-D_k\left(\mathrm{j}\omega_i\right)\Delta G\left(\mathrm{j}\omega_i,\theta\right)\right]\delta_{\theta_k}=D\left(\mathrm{j}w_i,\theta^{[0]}\right)\Delta G\left(\mathrm{j}\omega_i,\theta\right). \tag{B1}$$

Since the maximum singular value and the Frobenius norm of a matrix are two equivalent norms [28,30], for a sufficiently small positive number ε, $\bar{\sigma}\left(\Delta G\left(\mathrm{j}\omega_i,\theta\right)\right)$ is also sufficiently small. For given ω_i and $\theta^{[0]}$, the maximum singular values of matrices $D_k\left(\mathrm{j}\omega_i\right)$ and $N_k\left(\mathrm{j}\omega_i\right)-D_k\left(\mathrm{j}\omega_i\right)G\left(\mathrm{j}\omega_i,\theta^{[0]}\right)$ are bounded. Thus, for a sufficiently small ε, it follows that

$$\bar{\sigma}\left(\Delta G\left(\mathrm{j}\omega_i,\theta\right)\right)\ll\bar{\sigma}\left(N_k\left(\mathrm{j}\omega_i\right)-D_k\left(\mathrm{j}\omega_i\right)G\left(\mathrm{j}\omega_i,\theta^{[0]}\right)\right).$$

Therefore, Eq. (B1) can be equivalently rewritten as

$$\begin{aligned}&\sum_{k=1}^n\left[N_k\left(\mathrm{j}\omega_i\right)\hat{D}\left(\mathrm{j}\omega_i,\theta^{[0]}\right)-D_k\left(\mathrm{j}\omega_i\right)\hat{N}\left(\mathrm{j}\omega_i,\theta^{[0]}\right)\right]\delta_{\theta_k}\\&\quad=D\left(\mathrm{j}w_i,\theta^{[0]}\right)\Delta G\left(\mathrm{j}\omega_i,\theta\right)\hat{D}\left(\mathrm{j}\omega_i,\theta^{[0]}\right)+O\left(\Delta G^2\left(\mathrm{j}\omega_i,\theta\right)\right).\end{aligned} \tag{B2}$$

To simplify the derivation, variables $\left(\mathrm{j}w_i,\theta^{[0]}\right)$ and $\left(\mathrm{j}w_i,\theta\right)$ are omitted, and subscripts r and j are added to represent real and imaginary parts of the corresponding complex matrix, respectively. It leads to

$$\begin{aligned}\left(D\Delta G\hat{D}\right)_{\mathrm{r}}&=\left[D_{\mathrm{r}},-D_{\mathrm{j}}\right]\left[\begin{array}{c}\Delta G_{\mathrm{r}}\\ \Delta G_{\mathrm{j}}\end{array}\right]\hat{D}_{\mathrm{r}}-\left[D_{\mathrm{j}},D_{\mathrm{r}}\right]\left[\begin{array}{c}\Delta G_{\mathrm{r}}\\ \Delta G_{\mathrm{j}}\end{array}\right]\hat{D}_{\mathrm{j}},\\\left(D\Delta G\hat{D}\right)_{\mathrm{j}}&=\left[D_{\mathrm{j}},D_{\mathrm{r}}\right]\left[\begin{array}{c}\Delta G_{\mathrm{r}}\\ \Delta G_{\mathrm{j}}\end{array}\right]\hat{D}_{\mathrm{r}}+\left[D_{\mathrm{r}},-D_{\mathrm{j}}\right]\left[\begin{array}{c}\Delta G_{\mathrm{r}}\\ \Delta G_{\mathrm{j}}\end{array}\right]\hat{D}_{\mathrm{j}}.\end{aligned}$$

Let $\zeta_{G,i}\left(\mathrm{j}w_i,\theta\right)=\mathrm{vec}([\begin{smallmatrix}\Delta G_{\mathrm{r}}\left(\mathrm{j}w_i,\,\theta\right)\\ \Delta G_{\mathrm{j}}\left(\mathrm{j}w_i,\,\theta\right)\end{smallmatrix}])$ and $O(\left\|\zeta_{G,i}\right\|_2^2)=O(\left\|\zeta_{G,i}\left(\mathrm{j}w_i,\theta\right)\right\|_2^2)$. Utilizing properties of the Kronecker product and the symbols defined above, let the real and imaginary parts of (B2) equal respectively, and combining these equations, the following can be obtained:

$$\mathit{\Pi}\left(\omega_i,\theta^{[0]}\right)\delta_\theta=\mathcal{P}\left(\omega_i,\theta^{[0]}\right)\zeta_{G,i}\left(\mathrm{j}w_i,\theta\right)+O\left(\left\|\zeta_{G,i}\right\|_2^2\right). \tag{B3}$$

According to that the sloppiness is defined on a set of N distinct frequency points $\omega_i|_{i=1}^N$, for a given $\theta^{[0]}\in\boldsymbol{\Theta}$ and any $\theta=\mathrm{col}\left\{\theta_k|_{k=1}^n\right\}\in\boldsymbol{\Theta}$, Eq. (B3) holds simultaneously for $i=1,2,\ldots,N$. Let $\zeta_G(\omega_i|_{i=1}^N,\theta)=\mathrm{col}\{\zeta_{G,i}\left(\mathrm{j}w_i,\theta\right)|_{i=1}^N\}$, $O(\|\zeta_G\|_2^2)=\mathrm{col}\{O(\left\|\zeta_{G,i}\right\|_2^2)|_{i=1}^N\}$. Based on the definition of $\zeta_{G,i}\left(\mathrm{j}w_i,\theta\right)$ and (11), it is derived that $O(\|\zeta_G\|_2^2)=O(\varepsilon^2)$. By combining these equations, the following can be obtained:

$$\bar{\Pi}\left(\omega_i|_{i=1}^N,\theta^{[0]}\right)\delta_\theta=\bar{\mathcal{P}}\left(\omega_i|_{i=1}^N,\theta^{[0]}\right)\zeta_G\left(\omega_i|_{i=1}^N,\theta\right)+O\left(\varepsilon^2\right). \tag{B4}$$

According to Lemma 3, there exists a parameter vector $\theta \in \boldsymbol{\Theta}$ such that for $i = 1, 2, \ldots, N$, Eq. (B3) is simultaneously satisfied, if and only if the following equation holds:

$$U_{\bar{\Pi},2}^{\mathrm{T}} \bar{\mathcal{P}} \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \zeta_G \left(\omega_i|_{i=1}^{N}, \theta \right) = O\left(\varepsilon^2\right). \tag{B5}$$

It is worth noting that Eq. (B5) actually provides the range that $\zeta_G(\omega_i|_{i=1}^{N}, \theta)$ can take in the parameter space $\mathcal{C}^{2pqn}$ for any parameter vector $\theta \in \boldsymbol{\Theta}$. Hence, the following can be derived:

$$\zeta_G \left(\omega_i|_{i=1}^{N}, \theta \right) = \left(U_{\bar{\Pi},2}^{\mathrm{T}} \bar{\mathcal{P}} \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \right)_{\mathrm{r}}^{\perp} \xi + O\left(\varepsilon^2\right) = \Psi \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \xi + O\left(\varepsilon^2\right), \tag{B6}$$

where ξ is an arbitrary vector in $\mathbb{R}^n$.

Under Assumption 3, the system is globally identifiable at the specific parameter $\theta^{[0]}$ with a set of N distinct frequency points $\omega_i|_{i=1}^{N}$ and their corresponding frequency responses. According to Theorem 1, the real matrix $\bar{\Pi}(\omega_i|_{i=1}^{N}, \theta^{[0]})$ is of full column rank, which means that $V_{\bar{\Pi},2} = 0$ and $V_{\bar{\Pi},1} V_{\bar{\Pi},1}^{\mathrm{T}} = I$. Consequently, all parameters satisfying (B4) can be expressed as

$$\theta = \theta^{[0]} + V_{\bar{\Pi}} \Sigma_{\bar{\Pi}}^{-1} U_{\bar{\Pi},1}^{\mathrm{T}} \bar{\mathcal{P}} \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \zeta_G \left(\omega_i|_{i=1}^{N}, \theta \right). \tag{B7}$$

It is worth noting that since the matrix $\bar{\Pi}(\omega_i|_{i=1}^{N}, \theta^{[0]})$ is of full column rank, and the right-hand side of $\delta_\theta = \theta - \theta^{[0]}$ is a unit matrix, according to Lemma 3, there is no arbitrary matrix Z appearing in (B7). Substituting (B6) into (B7), the following can be derived:

$$\theta = \theta^{[0]} + S_\theta \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \xi + O\left(\varepsilon^2\right).$$

Therefore, based on (11) and the definition of $\zeta_G(\mathrm{j}w_i, \theta)$, it can be derived that

$$\begin{aligned}
&\sum_{i=1}^{N} \left\| G\left(\mathrm{j}\omega_i, \theta\right) - G\left(\mathrm{j}\omega_i, \theta^{[0]}\right) \right\|_F^2 \\
&\quad = \zeta_G^{\mathrm{T}} \left(\omega_i|_{i=1}^{N}, \theta \right) \zeta_G \left(\omega_i|_{i=1}^{N}, \theta \right) \\
&\quad = \xi^{\mathrm{T}} \Psi \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \Psi \left(\omega_i|_{i=1}^{N}, \theta^{[0]} \right) \xi + O\left(\varepsilon^2\right) \\
&\quad = \xi^{\mathrm{T}} \xi + O\left(\varepsilon^2\right).
\end{aligned}$$

Putting (11), (B6), and (B7) together, the proof is completed.

Appendix C Proof of Theorem 3

For a concise presentation, the dependence of all matrices and sets on the frequencies $\omega_i|_{i=1}^{N}$ and the parameter vector $\theta^{[0]}$ is omitted.

According to Assumption 2, that is, the set $\boldsymbol{\Theta}$ is open, it can be inferred that there exists a $\varepsilon^{[0]} > 0$ such that for all positive ε satisfying $\varepsilon \leqslant \varepsilon^{[0]}$ and any $\xi \in \tilde{\boldsymbol{\Theta}}_F(\varepsilon)$, $\theta^{[0]} + S_\theta \xi$ belongs to $\boldsymbol{\Theta}$. On the other hand, for any $\xi \in \tilde{\boldsymbol{\Theta}}_F(\varepsilon)$, $\theta^{[0]} + S_\theta \xi$ belongs to the set $\boldsymbol{\Theta}$, under the premise that ε is sufficiently small, $\bar{\boldsymbol{\Theta}}_F(\varepsilon)$ can be approximated as a subset of the set $\boldsymbol{\Theta}$. Therefore, given that $\xi \in \tilde{\boldsymbol{\Theta}}_F(\varepsilon)$, the problem of maximizing $\|S_\theta \xi\|_2^2$ can be equivalently written as

$$\begin{aligned}
&\underset{\xi}{\text{maximum}} \ \ \xi^{\mathrm{T}} S_\theta^{\mathrm{T}} S_\theta \xi, \\
&\text{subject to} \ \ \xi^{\mathrm{T}} \xi = \varepsilon^2.
\end{aligned} \tag{C1}$$

Let μ be the Lagrange multiplier and construct the Lagrange function $J(\mu, \xi)$ as follows:

$$J(\mu, \xi) = \frac{1}{2} \xi^{\mathrm{T}} S_\theta^{\mathrm{T}} S_\theta \xi - \frac{\mu}{2} \left(\xi^{\mathrm{T}} \xi - \varepsilon^2 \right).$$

Then, by differentiating the Lagrange function $J(\mu, \xi)$, the first-order optimization condition with respect to ξ can be obtained:

$$\frac{\partial J(\mu, \xi)}{\partial \xi} = \left(S_\theta^{\mathrm{T}} S_\theta - \mu \right) \xi = 0. \tag{C2}$$

It implies that the optimal ξ, defined as $\xi_{\mathrm{opt}}^{[\max]}$, is related to singular values of S_θ. Specifically, let $\xi^{[i]}$ be the singular vector corresponding to the ith singular value $\mu^{[i]}$, and let $\sigma^{[i]}$ be ${\xi^{[i]}}^{\mathrm{T}} \xi^{[i]}$, where $i = 1, 2, \ldots, n$. Then there exists a positive scalar $\alpha^{[i]}$ such that

$$\xi_{\mathrm{opt}}^{[\max]} = \alpha^{[i]} \xi^{[i]}. \tag{C3}$$

Substituting (C3) into the constraint of the optimization problem in (C1), it is clear that the constraint is satisfied if and only if ${\alpha^{[i]}}^2 \sigma^{[i]} = \varepsilon^2$. Therefore, we can obtain

$$\alpha^{[i]} = \frac{\varepsilon}{\sqrt{\sigma^{[i]}}}, \ \xi_{\mathrm{opt}}^{[\max]} = \frac{\varepsilon \xi^{[i]}}{\sqrt{\sigma^{[i]}}}. \tag{C4}$$

It implies that the Euclidean norm of the optimal vector $\xi_{\text{opt}}^{[\max]}$ is in the same order of magnitude as ε. Combined with (C2), it can be obtained that

$$\begin{aligned}&\xi_{\text{opt}}^{[\max]\mathrm{T}} S_\theta^{\mathrm{T}} S_\theta \xi_{\text{opt}}^{[\max]\mathrm{T}}\\&= \alpha^{[i]2} \xi^{[i]\mathrm{T}} S_\theta^{\mathrm{T}} S_\theta \xi^{[i]\mathrm{T}}\\&= \mu^{[i]} \alpha^{[i]2} \xi^{[i]\mathrm{T}} \xi^{[i]}\\&= \mu^{[i]} \varepsilon^2.\end{aligned}$$

It implies that the optimal solution to the optimization problem corresponding to (C1) is $\mu^{[1]}\varepsilon^2$. Moreover, the parameter vector that maximizes $\left\|\theta-\theta^{[0]}\right\|_2$ is defined as $\xi_{\text{opt}}^{[1]}$, which is given by (C4), where $\xi^{[i]}=\xi^{[1]}$.

Similarly, for $k=2,3,\ldots,n$, the following optimization problem is constructed recursively:

$$\begin{aligned}\max_{\xi}\ & \xi^{\mathrm{T}} S_\theta^{\mathrm{T}} S_\theta \xi,\\ \text{s.t.}\ & \xi^{\mathrm{T}}\xi=\varepsilon^2,\\ & \xi\in\mathbb{R}^n\backslash\operatorname{span}\left\{\xi_{\text{opt}}^{[1]},\xi_{\text{opt}}^{[2]},\ldots,\xi_{\text{opt}}^{[k-1]}\right\}.\end{aligned}$$

Based on properties of eigenvalues and eigenvectors, along with the aforementioned proof process, the optimal value of this optimization problem equals $\mu^{[k]}\varepsilon^2$. The parameter vector that maximizes $\left\|\theta-\theta^{[0]}\right\|_2$ is defined as $\xi_{\text{opt}}^{[k]}$, which is given by (C4), where $\xi^{[i]}=\xi^{[k]}$.

With the definition of set $\bar{\boldsymbol{\Theta}}_F(\varepsilon,\omega_i|_{i=1}^N,\theta)$, as well as the definitions of absolute and relative sloppiness, the proof is completed.