Sloppiness: fundamental study, new formalism and quantification

Prem Jagadeesan[1,3,4], Karthik Raman[2,3,4*], Arun K Tangirala[1,3,4*]

[1]Department of Chemical Engineering, Indian Institute of Technology (IIT) Madras, Chennai – 600 036, India
[2]Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, IIT Madras, Chennai – 600 036, India
[3]Robert Bosch Centre for Data Science and Artificial Intelligence (RBCDSAI), IIT Madras, Chennai – 600 036, India
[4]Centre for Integrative Biology and Systems mEdicine (IBSE), IIT Madras, Chennai – 600 036, India

*kraman@iitm.ac.in, arunkt@iitm.ac.in

Abstract

Precise estimation of parameters in a complex dynamical system is often challenging, even if provided with adequate quality and quantity of data. A major challenge is the possible presence of large regions in the parameter space over which model predictions are nearly identical. This property, known as sloppiness, has been reasonably well-addressed in the past decade, studying its possible impacts and remedies. However, certain critical unanswered questions concerning sloppiness, its quantification and practical implications in identification still prevail. In this work, we systematically examine sloppiness at a fundamental level and formalise a new theoretical definition of sloppiness. Further, we propose a method to quantify sloppiness for non-linear predictors. The proposed method aids in the characterisation of a model structure around a point of interest in the parameter space and detecting local structural unidentifiability. Further, we establish a mathematical relationship between practical identifiability and sloppiness in linear predictors. Finally, we demonstrate the proposed formalism and methods on standard models.

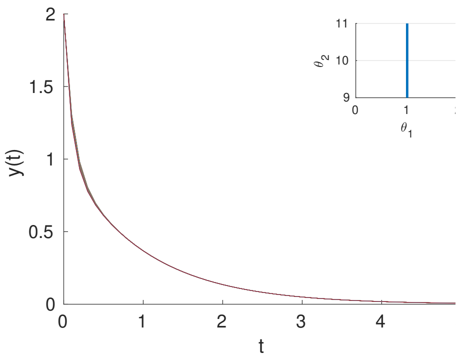
Introduction

Parameter estimation is one of the crucial and challenging steps in the computational modelling of complex dynamical systems. Complex dynamical systems are often modelled as non-linear ordinary differential equations with a large number of states and parameters. Whole-cell modelling is an example of complex dynamical systems [1]. Quantitative systems pharmacology (QSP) is an another example of complex dynamical systems modeling [2, 3]. In a modelling exercise, often, there are many free parameters to be estimated from data. Precise and accurate estimation of parameters depends on the quality and quantity of data, nature of the model structure, and the estimation algorithm [4]. In complex dynamical systems, there are often regions in the parameter space over which the model predictions are identical or nearly identical which result in structural unidentifiabilities [5], and sloppiness [6], respectively. Structural identifiability is a well-established concept in the domain of system identification [4]. The existence of structural unidentifiability implies multiple solutions to the estimation problem. Thus it is imperative to check for structural identifiability prior to parameter estimation. There are several analytical methods available for assessing structural identifiability of a model structure [5, 7, 8, 9, 10]. A differential geometric approach using observability condition is proposed in [11]. However, most of the analytical methods are not scalable to large models. A numerical method is proposed in [12] to assess local structural unidentifiabilities.

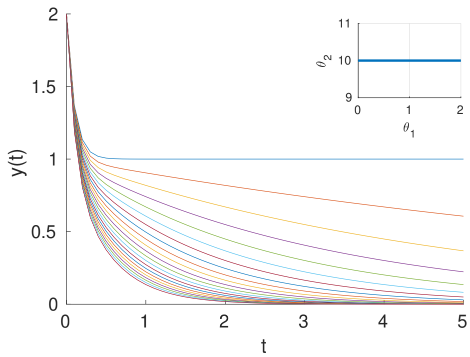
Similarly, the presence of sloppiness often results in huge uncertainties in parameter estimates [6]. Identifiability is a binary situation, whereas sloppiness lies somewhere in between identifiability and loss of identifiability; the closer it is to the loss of identifiability, the more problematic it is. Consider the simple bi-exponential model in (1), with a true parameter vector, $\theta = \begin{bmatrix}1 & 10\end{bmatrix}^T$. The model output $y(t)$ is computed for 5 seconds by fixing one parameter and varying the other one at a time.

$$y(t) = e^{-\theta_1 t} + e^{-\theta_2 t} \tag{1}$$

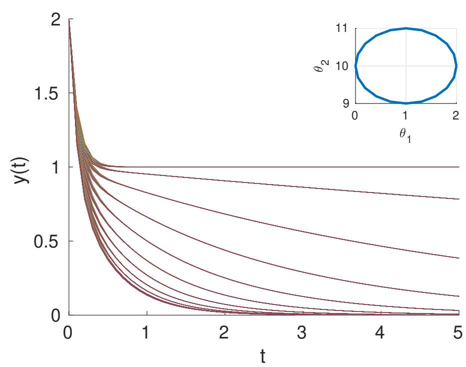
From Fig.1a, it is evident that for a range of θ_2 values in parameter space the model predictions are nearly identical, and hence, θ_2 direction is considered sloppy. However, in Fig. 1b, the change in parameter θ_1 results in clearly distinctive model outputs. Here, θ_1 is stiff direction. From Fig.1c, it is seen that there are certain directions over which the model predictions are nearly identical and significantly vary in other directions. However, in many cases, instead of individual parameters being sloppy and stiff, there will be directions in the parameter space that are sloppy and stiff. The nearly identical model outputs for a significant range of parameter sets might reflect as large standard errors in a subset of parameter estimates. These large standard errors in the parameter estimates are one of the crucial challenges while modelling sloppy models.



(a) $\theta_1 = 1$ and θ_2 varied from 9 to 11



(b) $\theta_2 = 10$ and θ_1 varied from 0 to -2



(c) θ_1 and θ_2 are varied around the circumference of a circle with radius $r = 1$ and $\theta_1 = 1, \theta_2 = 10$ as center

Figure 1: (a) shows the model output while varying θ_2 and fixing θ_1. The output of the model does not significantly vary and is qualitatively indistinguishable. (b) shows the model output while varying θ_1 by fixing θ_2. The output significantly varies while changing θ_1 in the same range. (c) shows the outputs for a range of parameters varied over the circumference of a circle with a fixed radius. Visually evident chunks of outputs that are nearly identical. This implies certain directions in the parameter space over which model outputs are nearly identical.

Study of sloppy models gained interest because of this counter-intuitive nature of uniform predictability even with highly uncertain parameter estimates. The impact of sloppiness on various facets of modelling has been extensively studied in the past decade [6, 12, 13, 14, 15, 16, 17]. Sloppiness in models gives rise to multiple challenges in each stage of a modelling exercise. One of the most common challenges encountered is the loss of practical identifiability [6, 18, 12]. Practical identifiability is defined as the ability to estimate parameters of a structurally identifiable model with acceptable precision given a data set [19, 20]. Practical identifiability is assessed by the width of the confidence interval of the parameter estimate [19]. However, there are other numerical methods proposed to assess practical identifiability [21, 22, 9]. The relationship between structural identifiability and sloppiness is well established in [12]. A sloppy model is always structurally identifiable. Whereas in the case of practical identifiability, though sloppiness is closely related to practical identifiability, the exact relationship is still unclear [12]. This has been one of the sources of ambiguity in using sloppiness analysis as a part of the modelling exercise. The ambiguity is due to measure of sloppiness as it does not guarantee the loss of practical identifiability[19].

The impact of sloppiness in a modelling exercise has been addressed widely in the following articles. It has been shown that the sloppy models have nearly flat cost surfaces in the vicinity of the optimal parameter set [13]. In such cases, it is observed that non-linear optimization algorithms may get stuck in the sloppy region. A differential geometric approach has been employed to improve the convergence of the Levenberg-Marquardt algorithm. In the case of validating models, it is observed that the uncertainty estimates are unreliable in the case of sloppy models [14]. They have suggested modified Markov Chain Monte Carlo (MCMC) simulations to circumvent this issue. An attempt to design optimal experiments for a precise estimate of parameters in sloppy models may result in compromising prediction accuracy [18]. The consequences of sloppiness on system identification can be summarised in the two following points (i) obtaining precise parameter estimates is challenging because the model's behaviour is highly insensitive to many parameter combinations/directions in the parameter space (ii) even though many parameters cannot be estimated with good precision, the prediction uncertainty in sloppy models is considerably low. The predictions depend only on a few stiff directions in the parameter space. Despite the significant progress, there are some critical answered questions that call for attention.

Even though the qualitative definition of sloppiness is unambiguous, when a model is found to be sloppy, the current sloppiness analysis does not answer the key questions such as (i) what is the source of sloppiness? (ii) what is the relationship of sloppiness with parameter uncertainty / practical identifiability? (iii) does the measure of sloppiness indicate the goodness of the estimated parameters? Once we obtain answers to these questions, the usefulness of sloppiness analysis in modelling and subsequent applications is significantly enriched. In this work, we attempt to answer these questions and bring some more clarity to the phenomenon of sloppiness and its role in a modelling exercise.

Numerous attempts have been made to resolve the issues arising from sloppy models, but very few works have attempted to find the root cause of sloppiness [23, 19, 15]. Though the source of sloppiness is attributed to both model and data [6], in most cases, the source of sloppiness is attributed to the model structure [17, 6, 12, 18, 24, 25, 14]. A computational study is carried out to study the relationship of sloppiness measure with structural identifiability, practical identifiability, and experimental design [19]. The result suggests that the relationship of sloppiness measure with practical identifiability is inconclusive. Tönsing, C. et.al. have shown that the root cause of sloppiness is both model and experimental condition [23]. They suggest that by changing the experimental condition, it is possible to cure sloppiness. Apgar et al. have shown it is possible to estimate parameters precisely on sloppy models by careful design of experiments [24]. Sloppiness have also been reported to have connections with biological phenomenon such as robustness and evolvability [9].

The utility of sloppiness analysis in modelling has been questioned in [19, 26]. They argue through computational study that the presence of sloppiness does not guarantee either structural or practical unidentifiability. Based on the result, they conclude that using sloppiness analysis to assess the identifiability of parameters can be misleading and suggest identifiability analysis as a better tool. Though the study is convincing, the effect of sloppiness in errors of the parameter estimate is inevitable. The relationship of sloppiness with practical identifiability is widely accepted but not formally established [19, 12]. This motivates us to revisit the definition and measure of sloppiness to find a remedy.

In this work, we revisit the definition of sloppiness and show that with the current measure of sloppiness, in the case of non-linear predictors, it is not possible to attribute sloppiness to model structure alone decisively. In order to circumvent the ambiguity, first, we formulate two new theoretical definitions of sloppiness (i) Sloppiness (ii) Conditional sloppiness. The proposed definitions of sloppiness for autonomous differential equation systems is defined in an augmented space of parameter and initial conditions. This implies that both model and experimental conditions together responsible for sloppiness. Secondly, we establish a mathematical relationship between practical identifiability and sloppiness in the case of linear predictors. Further, we propose a new measure for sloppiness and a visual tool that can detect sloppiness and insensitive parameters in addition to structural unidentifiability. Given a set of experimental conditions, the insensitive parameters are more likely to become practically unidentifiable. The proposed definitions of sloppiness and the visual tool can help identify whether the true model is in a sloppy region and corresponding insensitive parameters. The benefit of analysing the true model is two-fold; once we know that the true model is in a sloppy region, the probability of the estimated model landing in the sloppy region is high. Then appropriate experiments can be designed to circumvent the effects. On the other hand, if the true model is not in the sloppy region and if the estimated model is sloppy, we can fine-tune the estimation algorithm and other controllable experimental conditions. The proposed tool can also detect multi-scale sloppiness [12].

The rest of the paper is organized as follows: Section 2 provides perspectives on sloppiness by revisiting the concept of sloppiness and rightly positioning it relative to identifiability. Section 3 presents key results including two motivating examples to highlight the challenges in using the current measure of sloppiness. Further, it provides two new definitions of sloppiness and their relationship with practical identifiability in case of linear predictors. Section 4 illustrates the proposed method to assess the sloppiness of non-linear predictors. The paper ends with concluding remarks in Section 5.

Perspectives

Sloppiness and identifiability

The end goal of a modelling exercise is a useful model. The quality and usefulness of an estimated model depend on a few crucial properties. Identifiability is one such property that guarantees the existence of a unique model. The fundamental requirement is that the model structure being unique with respect to the parameters (structural property). This is known as structural identifiability (SI). Given a structurally identifiable model structure, the ability to recover unique model from data is associated with practical identifiability (PI). While there are several methods to assess structural identifiability, practical identifiability is usually assessed from the precision of the parameter estimates.

Within the class of structurally identifiable models, sloppiness is another important model property. Sloppiness at its core quantifies the sensitivity of output with respect to change in parameters or directions in the parameter space. When the gradient of the predictor qualitatively does not change as the parameter is varied significantly, then the system is sloppy. Sloppiness can be observed as pockets of regions in the parameter space, in such cases, the gradient of the predictor vary significantly in the parameter directions.

In the existing literature, a method of quantifying this insensitivity of predictions to changes in parameters is captured by (2).

$$S = \frac{\lambda_{min}(H)}{\lambda_{max}(H)} \tag{2}$$

Where H is the Hessian of the cost function, in general, for a non-linear parameter space, it is difficult to identify all such pockets of sloppy regions. Hence, sloppiness is determined locally around a parameter of interest by computing the sensitivity of the output to the change in the parameter. The sloppiness is characterized by equally spaced eigenvalues of the Hessian of the cost function in the log scale and is quantified by 2.

Even though there is no strict cut-off, a model is considered to be sloppy if $S \leq 10^{-6}$ [25]. The eigenvectors corresponding to maximum and minimum eigenvalues indicate the stiff and sloppy direction, respectively. It can be observed that the Hessian of the cost function is representative of the derivative of the output with respect to the parameters. The current measure of sloppiness is valid only if the estimation algorithm is least-squares; for a different estimation algorithm, the Hessian of the cost function may not be representative of the sensitivity of the output with respect to the parameters.

Table 1 shows the list of model properties, their definitions, and the corresponding method of assessment. Even though qualitative definition for practical identifiability and sloppiness are available, a formal mathematical framework for defining sloppiness is not available. In this work, we formalize sloppiness mathematically.

Table 1: List of model properties and their definitions

#	Model Property	Definition	Assessment
1	SI	$\hat{y}(t,\theta_1) = \hat{y}(t,\theta_2) \iff \theta_1 = \theta_2$	Direct application
2	PI	-	trace$(\Sigma_{\hat{\theta}})$
3	Sloppiness	-	$H_{i,j} = \frac{\partial^2 C(\theta)}{\partial log\theta_j \partial log\theta_i}$

As a first step in answering the questions raised above, the notion of sloppiness needs to be rightly positioned relative to well-established concepts such as structural identifiability and practical identifiability. Additionally, when a model is sloppy, it is important to attribute the source of sloppiness to the appropriate factors. All the three properties of interest that are under discussion arise due to either one or more of data, model, and estimation algorithm. The data itself is characterized by (i) signal to noise ratio (SNR) (ii) sample size, and (iii) the input. While input and SNR determine the quality of data, the quantity of data is determined by sample size. The model's contribution towards these properties is characterized by predictor gradient, and finally, the estimation method is generally characterized by the cost function.

Table 2: Factors influencing model properties

#	Model Property	SNR	Sample size	Input	$\nabla(y(\theta))$	$C(\theta)$
1	SI	-	-	-	✓	-
2	PI	✓	✓	✓	✓	✓
3	Sloppiness	-	-	✓	✓	-

Table 2 lists the factors that affect structural identifiability, practical identifiability and sloppiness. Structural identifiability analysis is searching for two or more parameter sets that result in identical predictions in the absence of data. Hence, only the gradient of the predictions with respect to parameters influences structural identifiability. At the same time, all three factors influence practical identifiability: the data, model, and estimation algorithm. A structurally identifiable model may end up practically unidentifiable due to noise, sample size, insufficient excitation, or a combination of these factors. Sloppiness of the true model is influenced by both input and gradient of the predictor with respect to parameters. In the case of autonomous differential equations, the system's initial conditions can be considered as impulse inputs. While the sloppiness of the estimated model can be attributed using the matrix in Fig.2.

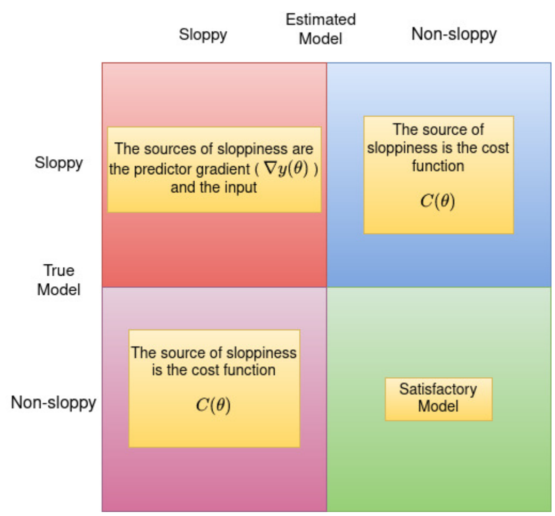


Figure 2: Matrix for finding the sources of sloppiness to an estimated model. The rows of the matrix are sloppy and non-sloppy categories of the true model, while the columns are sloppy and non-sloppy categories of the estimated model.

When the estimated model is sloppy, and the true model is also sloppy, the possible reasons are predictor gradient, input, or both. When the true model is non-sloppy and the estimated model is sloppy, then the only possible reason is the cost function. The third case is when the true model is sloppy and the estimated model is non-sloppy, again cost-function is the possible reason. Finally, when both the models are non-sloppy then we obtain a satisfactory model. The role of the cost function in sloppiness is indirect. When the true model is not sloppy, the cost function may push the estimated model into a sloppy region while the cost function itself does not affect the gradient of the predictor.

The relationship of sloppiness with practical identifiability is still an open problem [12]. From Table 2 it is clear that both the input and the gradient of the output with respect to parameters affect sloppiness and practical identifiability. This is the reason why most parameters in a sloppy model become practically unidentifiable when the gradient is large.

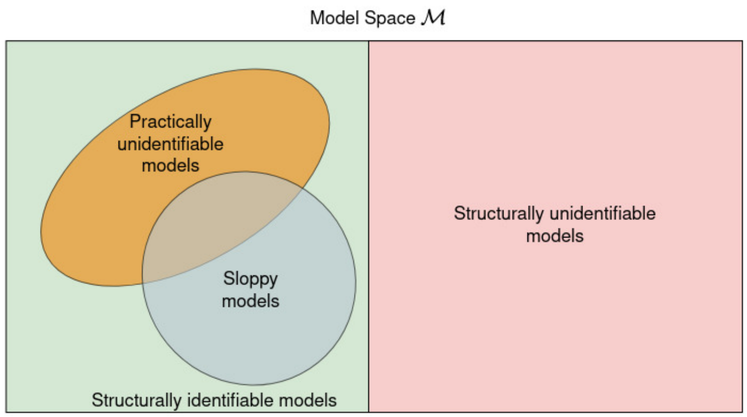


Figure 3: The relationship between structural identifiability, practical identifiability, and sloppiness

The relationship between sloppiness structural and practical identifiability is depicted in the Fig.3. Sloppy models intersect with practically unidentifiable models. However, how much sloppiness affects practical identifiability is the question that needs to be answered. In the following section, using two motivating examples, we show that the current measure of sloppiness does not answer the above question, and we also argue that the lack of formal mathematical definition is another reason.

Motivating examples

This section provides two motivating examples that illustrate the challenges in the current measure of sloppiness in answering the questions of interest in this work. We consider a non-sloppy model and show that it can be turned sloppy by changing experimental conditions, and a model that is regarded as sloppy can be made non-sloppy again by changing the experimental condition.

Example 1

Consider the linear predictor given in (3). Let $\mathbf{z}$ be a vector of m observations. The parameters of the model are estimated using the method of least squares. The Hessian of the cost function is given in (5).

$$y = a_1x_1 + a_2x_2 + ... + a_nx_n \tag{3}$$

Now, let us define

$$C(a) = \sum_{i=1}^{m}(z_i - y)^2 = \sum_{i=1}^{m}(z_i - (a_1 x_{1_i} + a_2 x_{2_i} + ... + a_n x_{n_i}))^2 \tag{4}$$

$$\nabla^2(C(a)) = \begin{bmatrix} 2\sum_{i=1}^{m} x_{1_i}^2 & 0 & \dots & 0 \\ 0 & 2\sum_{i=1}^{m} x_{2_i}^2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 2\sum_{i=1}^{m} x_{n_i}^2 \end{bmatrix} \tag{5}$$

For the purpose of discussion, we use an existing measure of sloppiness given in (6), the ratio of smallest to largest eigenvalues of the Hessian of the cost function.

$$S = \frac{\|x_{min}\|_2^2}{\|x_{max}\|_2^2} \tag{6}$$

Linear predictors are not generally observed to be sloppy [18]. However in the above example, the ratio of eigenvalues in only a function of data and not the model parameters and hence we can see that it is possible for the linear predictor to show sloppiness for some experimental condition. Moreover, sloppiness in linear predictors given in the above example is purely an artefact of data/input, and not due to the nature of the model structure.

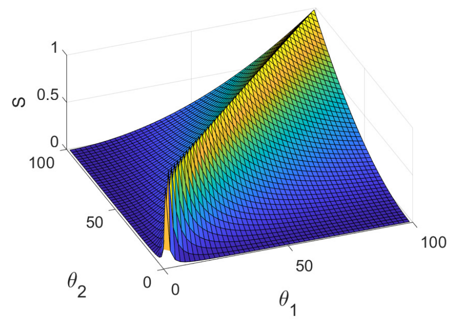
Example 2

Consider the state space model in (7). The Hessian of the least-squares cost function is given in (8).

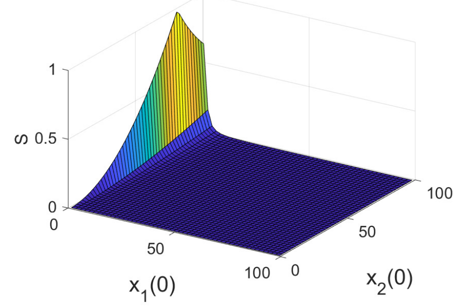
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{7}$$

$$y(t) = x_1(0)e^{-\theta_1 t} + x_2(0)e^{-\theta_2 t}$$

$$\nabla^2(C(a)) = \begin{bmatrix} \int_0^t \theta_1^2 x_1(0)^2 e^{-2\theta_1 t} dt & \int_0^t \theta_2\theta_2 x_1(0)x_2(0)e^{-(\theta_2+\theta_2)t} dt \\ \int_0^t \theta_2\theta_2 x_1(0)x_2(0)e^{-(\theta_2+\theta_2)t} dt & \int_0^t \theta_2^2 x_2(0)^2 e^{-2\theta_2 t} dt \end{bmatrix} \tag{8}$$



(a) Sloppiness in the parameter space



(b) Sloppiness in initial condition space

Figure 4: (a) Shows sloppiness in the parameter space. There are regions in the parameter space where system is non-sloppy. (b) Shows sloppiness for various initial conditions. For a subset of initial conditions the model becomes non-sloppy

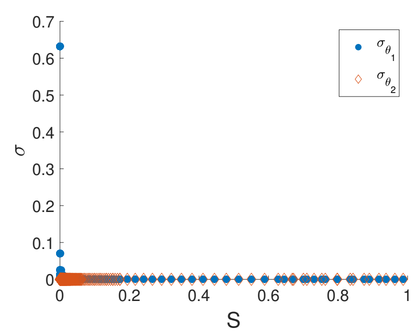


Figure 5: Standard errors of the parameter estimates obtained from various initial conditions (Fig. 4b) plotted against the corresponding sloppiness value. It seen that there is no specific relationship between sloppiness and standard errors.

Sloppiness is computed in the grid $1 \leq \theta_1 \leq 100$ and $1 \leq \theta_1 \leq 100$. From Fig. 4a, it is evident that sloppiness is a function of parameter space. In Fig. 4b, sloppiness is computed over a grid of initial conditions $1 \leq x_1(0) \leq 100$ and $1 \leq x_2(0) \leq 100$ around the point θ_1 =1, $\theta_2 = 100$. The model is simulated for $t = 0$ to $t = 20$ seconds with a sampling interval of 0.5 seconds.

From Fig. 4b, we can observe that the change in experimental condition can cure sloppiness, which indicates that the sloppiness is a function of information contained in a data set. In our previous work [27], we have demonstrated using simulation that parameters contributing to sloppy directions have very low information gain in the Bayesian framework. For initial condition > 50 of x_2, the model becomes non-sloppy. This has been extensively studied in [23, 24, 9]. Tonsing *et al. (2014)*

showed that the structure of the sensitivity matrix is the reason for sloppiness in ordinary differential equations and also conclude that both experiment and model structure are the root causes of sloppiness. Fig. 5 shows that the relationship between sloppiness and standard errors of the parameters does not follow any definitive trend. Though it has been argued that sloppiness and practical identifiability are two distinct concepts and that they are incorrectly conflated [18], the effect of sloppiness on practical identifiability cannot be ignored [6].

Following are two important observations emanate from our study: (i) for a non-linear least-squares estimation problem, it is impossible to attribute sloppiness to the model structure alone. Sloppiness is a function of both parameter space and initial conditions and hence, labelling a model to be sloppy with the current analysis method of sloppiness can be misleading [6, 18] (ii) sloppiness often results in loss of practical identifiability. However, the exact relationship is not revealed by the current measure of sloppiness

Problem statement

The demonstrated challenges of the current measure of sloppiness reaffirms the need to answer the following questions (i) what is the source of sloppiness ? (ii) what is the relationship of sloppiness with parameter uncertainty / practical identifiability? (iii) does the measure of sloppiness indicate the goodness of the estimated parameters? In the following section we propose a new definition of sloppiness to circumvent these challenges.

Results

A new mathematical definition of sloppiness

The above-perceived challenges of the present sloppiness analysis motivated us to formulate a new mathematical definition of sloppiness. The new definition of sloppiness is based on the following remarks.

Remark 1. *Sloppiness is assessed across an augmented space (parameters, initial conditions, and inputs) rather than parameter space alone. In the case of autonomous ODE models, the augmented space is*

$$\phi = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_n & \vdots & x_1(0) & x_2(0) & \dots & x_m(0) \end{bmatrix}$$

Remark 2. *A significant change in the subset of augmented space (parameter space) results in small change in prediction space.*

$\mathcal{D}_{\mathcal{M}}$ - All possible parameter values the model $\mathcal{M}$ can take. $\mathcal{I}$ - Identifiable region, $\mathcal{Z}_{\mathcal{M}}$ - Set of all experiments for which the model structure $\mathcal{M}$ is identifiable.

Definition 1. *A model $\mathcal{M}$ is (ϵ, δ) sloppy with respect to experiment space $\mathcal{Z}_{\mathcal{M}}$ at $\theta^* \in \mathcal{I} \subset \mathcal{D}_{\mathcal{M}}$, if*

$$||\theta^* - \theta_1||_2 > \delta \ \forall \theta \in \mathcal{S} \subset \mathcal{I} \tag{9}$$

$$||y(\theta^*, t) - y(\theta_1, t)||_2^2 < \epsilon \ \forall \mathcal{Z} \in \mathcal{Z}_{\mathcal{M}} \tag{10}$$

for every (θ_1, θ^*) satisfying (9) & (10). ϵ arbitrarily small. $\delta \gg \epsilon$.

Definition 2. *: A model $\mathcal{M}$ is conditionally (ϵ, δ) sloppy with respect to an experiment space $\mathcal{Z}_{\mathcal{M}}$ at $\theta^* \in \mathcal{I} \subset \mathcal{D}_{\mathcal{M}}$, if*

$$||\theta^* - \theta_1||_2 > \delta \ \forall \theta \in \mathcal{S} \subset \mathcal{I} \tag{11}$$

$$||y(\theta^*, t) - y(\theta_1, t)||_2^2 < \epsilon \, \forall u \in \mathcal{Z} \subset \mathcal{Z}_{\mathcal{M}} \tag{12}$$

for every (θ_1, θ^*) satisfying (11) & (12). ϵ arbitrarily small. $\delta \gg \epsilon$.

Equations (8) & (9) say that a model is considered sloppy, if the (ϵ, δ) condition holds true for all possible experimental conditions $\mathcal{Z}_{\mathcal{M}}$. Similarly, (8) & (9) convey that the model is conditionally sloppy if the (ϵ, δ) condition holds true only for a subset $(\mathcal{Z})$ of all possible experimental conditions (Z).

No longer the proposed sloppiness and conditional sloppiness is used in a generic sense, however they have to be used in conjunction with ϵ and δ. By virtue of our definitions we believe, by itself the sloppiness is qualitative and where ever quantitative sloppiness is to be discussed, the (ϵ, δ) from should be used. For a large δ if the ϵ is negligibly small, then the system is considered to be (δ, ϵ) sloppy. Note that the proposed measure of sloppiness is different from the multi-scale sloppiness proposed in [12]. Multi-scale sloppiness is the ratio of maximum to minimum prediction error for a non-infinitesimal perturbation from parameter, whereas in the proposed definition, sloppiness is a function of both prediction error and parameter perturbation. Moreover, multi-scale sloppiness is only a function of δ where as, the proposed sloppiness is a function of both ϵ and δ which is a more natural way to define and understand the notion sloppiness.

Sloppiness analysis of a linear predictor

Consider a linear predictor in

$$y(x) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \tag{13}$$

$$\theta = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^T \& \quad \mathbf{X} = \begin{bmatrix} \mathbf{x_1} & \mathbf{x_2} & \cdots & \mathbf{x_n} \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\theta \tag{14}$$

The model is identifiable if $\mathbf{X}$ is full column rank. The model is sloppy if $\mathbf{X}(\theta^* - \theta_1) = \mathbf{y}$ for $||\theta^* - \theta_1||_2 > \delta$ and $||\mathbf{y}||_2 < \epsilon$.

Let $\theta^d = (\theta^* - \theta_1)$, $||\theta^d||_2 >> \epsilon$.
Using matrix norm,

$$||\mathbf{X}\theta^d||_p \leq ||\mathbf{X}||||\theta^d||_p \ \forall p \in \mathcal{R}. \tag{15}$$

Consider the extreme case for $p = 2$,

$$||\mathbf{X}|| = \frac{\epsilon^2}{\delta} \tag{16}$$

For extremely small ϵ and extremely large δ the $||\mathbf{X}|| \approx 0$ will become numerically unstable and that will result in loss of identifiability. This is the reason why in most of the cases linear least square problems are observed to be non-sloppy. Once the system is found practically unidentifiable, then a careful design of experiment will constrain the parameter uncertain making the matrix norm of data significantly large, in such case the possibility of sloppiness is almost eliminated.

Relationship between conditional sloppiness and practical identifiability

For a structurally identifiable model structure $\mathcal{M}(\theta)$ and a data set $\mathbf{z}$, precision of parameter estimate $\theta_i \in \theta$ is a measure of practical identifiability. Practical identifiability is assessed for a data set $\mathbf{z}$ given a model structure $\mathcal{M}$. A parameter θ_i is said to be practically unidentifiable for a given $\mathbf{z}$, if

$$\sigma_{\theta_i} > \hat{\delta}_i$$

$$d(\mathbf{z}, y(\hat{\theta})) < \epsilon$$

where σ_{θ_i} is the standard error of the parameter estimate θ_i and $d(\mathbf{z}, y(\hat{\theta}))$ is a prediction error. Practical identifiability can only be assessed post estimation because it is a function of the data set, whereas sloppiness can be evaluated at any given unknown space (ϕ), which includes parameter space and data set. Ideally, the infinite width of the confidence interval for a parameter estimate is considered practically unidentifiable. However, in practical scenarios, significant standard errors in the estimates are considered practically unidentifiable.

Consider the linear predictor given in (13). Choose another parameter set $\boldsymbol{\theta_1} = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}^T$ such that $\delta = \sqrt{(a_1 - b_1)^2 + \cdots + (a_n - b_n)^2}$.

$$(y(x) - y_1(x))^2 = ((a_1 - b_1)x_1 + (a_2 - b_2)x_2 + \cdots + (a_n - b_n)x_n)^2 \tag{17}$$

$$\epsilon = (\delta_1 x_1 + \delta_2 x_2 + \cdots + \delta_n x_n)^2 \tag{18}$$

If each of x_i is a vector of m observations, then (18) becomes,

$$\epsilon = \sum_{i=1}^{m} (\delta_1 x_{1_i} + \delta_2 x_{2_i} + \cdots + \delta_n x_{n_i})^2 \tag{19}$$

Inverse of (5) gives the covariance matrix for the parameter estimates. The standard error of a parameter a_i is given by

$$\sigma_{a_i} = \sqrt{\frac{1}{\sum_{i=1}^{m} x_i^2}} \tag{20}$$

Using (19), the relationship between the sloppiness and practical identifiability is derived as

$$\sigma_{a_i} = \frac{\delta_i}{(\sqrt{\epsilon - ((\sum_{k=1}^{m} \delta_k x_k)^2 + 2\sum_{i=1}^{m}(\delta_1 \delta_2 x_{1_i} x_{2_i} + \cdots))}; k \neq i \tag{21}$$

This relationship holds if the data is derived from a Gaussian distribution and the least-squares cost function. From (21), it is clear that there is a relationship between sloppiness and practical identifiability. The corresponding parameter will become unidentifiable for a small ϵ and a very large δ_i. The above result agrees with the result obtained in (16). For a generalised non-linear predictor, it is challenging to construct a relationship analytically. Hence, we propose a numerical method to analyse the non-linear predictor for sloppiness and identifiability.

Sloppiness analysis of non-linear predictors

We provide three numerical examples to demonstrate the sloppiness analysis of non-linear predictors. We propose a novel numerical method of analyze sloppiness based on the new definitions. The detailed working of the method is illustrated in the methods section. The first illustrative example demonstrates all the three different scenarios (i) a sloppy region, (ii) a non-sloppy region and (iii) an unidentifiable region in a toy model. The second example is a high-dimensional biochemical pathway used in [19]. This example demonstrates the working of the proposed algorithm in a high-dimensional model. The third example is one of the hallmark models used by Gutenkunst *et al.* in [6] to demonstrate sloppiness.

An illustrative example

In this example, we use a simple two-parameter state-space to demonstrate the working of our method. The proposed visual tool is generated for a linear decoupled state-space model. Three different scenarios are considered to cover the loss of structural identifiability, sloppiness, and an ideal scenario. Consider the state-space model in (22).

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{22}$$

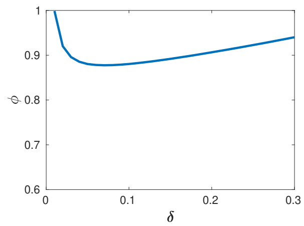
$$y(t) = x_1(t) + x_2(t)$$
$$x_1(0) = x_2(0) = 1$$

Table 3: Specifications of various scenarios

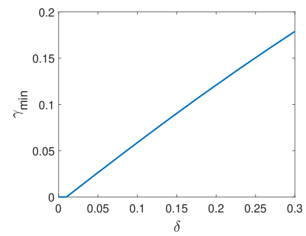
Scenario	θ^*	δ	Region
1	$\theta_1 = 0.4\ \theta_2 = 1$	0.3	non-sloppy
2	$\theta_1 = 1\ \theta_2 = 10$	0.3	sloppy
3	$\theta_1 = 0.4\ \theta_2 = 0.5$	0.3	unidentifiable

Parameters are sampled inside an n-ball with radius δ around θ^*. The model output $y(t, \theta)$ is computed for all the sampled parameters and sum-squared error γ is computed. Further, the model sensitive index is computed using (28).

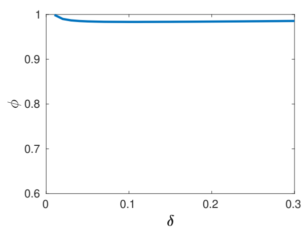
Figs. 6a and 6b show the plot for model sensitivity index and minimum deviation from reference θ^*. In case of the non-sloppy region, the curve in Fig. 6b is monotonically increasing. The model's behaviour is distinguishable from the reference point (θ^*) as δ increases. Additionally, the curve in model sensitivity plot in Fig.6a deviates away from unity value as δ increases. In case of a sloppy region, in Fig. 6d the curve has an increasing trend, but the curve is almost flat and not distinguishable as compared to the non-sloppy region from the reference point. In the third scenario, where the system is locally structurally unidentifiable, from Fig. 6d, it is seen that γ_{min} is very close to zero/ numerically zero for a non-zero δ. From Fig. 7, we can infer that both the parameters are unidentifiable in the region because at an absolute distance δ_i from the reference value, the prediction error goes to zero, which indicates that there is another parameter $\theta \neq \theta^*$ for which the prediction error is zero. In addition to that, in Fig. 6e, the curve in model sensitivity plot touches the unity value for a non-zero δ. In order to asses local structural non-identifiability, model sensitivity plot in conjunction with γ_{min} plot may be used. However, a numerically zero value of γ_{min} for a non-zero δ is sufficient to assess local structural non-identifiability. The model sensitivity index (ψ) quantifies the asymmetry between most sensitive and least sensitive parameter directions in the parameter space. The conditional sloppiness can be assessed by the pair (γ_{min}, δ). For a given δ, which can be chosen as a function of acceptable parameter range among the set of parameters, if the γ_{min} is too low, then we can say the system is conditionally sloppy with a high probability of practically unidentifiable parameters.



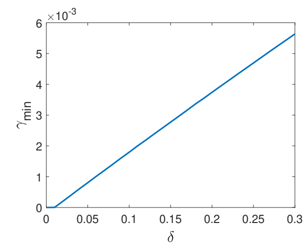
(a) Model sensitivity plot for non-sloppy region



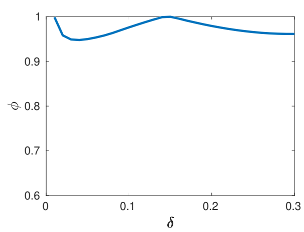
(b) γ_{min} vs δ for non-sloppy region



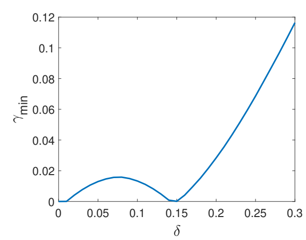
(c) Model sensitivity plot for sloppy region



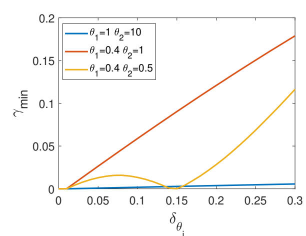
(d) γ_{min} vs δ for sloppy region



(e) Model sensitivity plot for unidentifiable region



(f) γ_{min} vs δ for unidentifiable region



(g) γ_{min} vs δ for all the three regions

Figure 6: Visual sensitivity analysis plot (a) The curve is significantly deviated from the unity value for the given δ, indicating local structural identifiability. (b) The slope of curve very close to unity value, indicating a non-sloppy region in the given delta. (c) The curve is significantly deviated from the unity value indicating local structural identifiability. (d) The slope of the curve is constant but the value of the slope is closer to the zero value indicating a sloppy region in the given δ. (e) & (f) The curve numerically hits unity and zero value at $\delta = 0.14$ indicating a local structural unidentifiability. (g) γ_{min} curve for all the three cases.

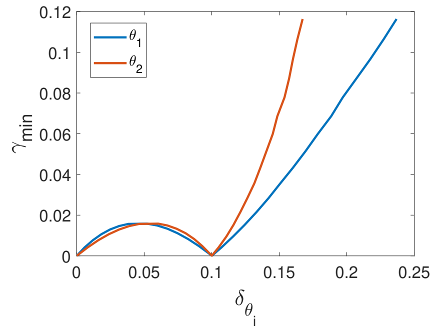
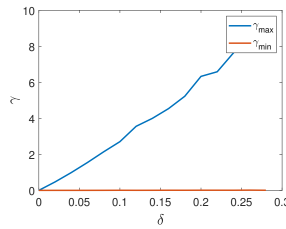


Figure 7: γ vs δ_{θ_i} for unidentifiable region. The curves touching the zero value indicate both the parameters are structurally unidentifiable

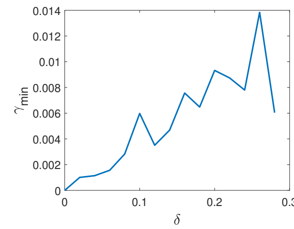
High dimensional biochemical pathway model

A linear biochemical pathway with fourteen states, sixteen parameters and one input is considered for demonstration [19]. The states $x_1(t)$ and $x_{14}(t)$ are measured. The nominal parameter set (θ^*) is taken from [19].

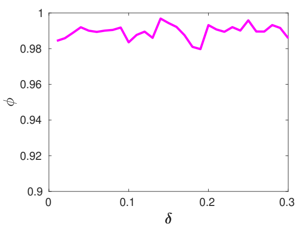
$$\mathcal{M}: \begin{cases} \dot{x}_1 = -\frac{v_m x_1}{k_m + x_1} + p_1 u \\ \dot{x}_2 = -p_1 x_1 - p_2 x_2 \\ \dot{x}_3 = -p_2 x_2 - p_3 x_3 \\ \dot{x}_4 = -p_3 x_3 - p_4 x_4 \\ \dot{x}_5 = -p_4 x_4 - p_5 x_5 \\ \dot{x}_6 = -p_5 x_5 - p_6 x_6 \\ \dot{x}_7 = -p_6 x_6 - p_7 x_7 \\ \dot{x}_8 = -p_7 x_7 - p_8 x_8 \\ \dot{x}_9 = -p_8 x_8 - p_9 x_9 \\ \dot{x}_{10} = -p_9 x_9 - p_{10} x_{10} \\ \dot{x}_{11} = -p_{10} x_{10} - p_{11} x_{11} \\ \dot{x}_{12} = -p_{11} x_{11} - p_{12} x_{12} \\ \dot{x}_{13} = -p_{12} x_{12} - p_{13} x_{13} \\ \dot{x}_{14} = -p_{13} x_{13} - p_{14} x_{14} \\ y(t) = x_1(t) + x_{14}(t) \end{cases} \tag{23}$$



(a) δ vs γ plot



(b) δ vs γ_{min} plot



(c) δ vs ψ plot

Figure 8: (a) The curves indicate uniform slope sensitivity ratio. (b) The curve has significantly small slope and numerically small ϵ which indicate the sloppiness for the given δ (c) The curve is significantly deviated form the unity value indicating local structural identifiability of the model.

From Fig. 8b, the model is locally structurally identifiable. In Fig. 8a, the ratio of values of $\frac{\gamma_{max}}{\gamma_{min}}$ increases as δ increases; however, the ratio is not numerically significant. Hence, the model is not sloppy in the traditional and multi-scale notion of sloppiness. On the other hand, in Fig. 8b, the for $\delta > 0.25$, the $\gamma_{min} \approx 0.015$. This observation implies that the system is (ϵ, δ) sloppy, and a sub-set of insensitive parameters will contribute to the sloppiness.

From Figures 9 to 12, it is seen that the parameters v_m, k_m, p_7 and p_9 are insensitive for $\delta_i > 0.03$ compared to other parameters. The sensitivity in maximum deviation direction is significantly high for the parameters p_5, p_6, p_{13} and p_{14} in the vicinity of θ_i^*. In sum, the biochemical network considered is structurally locally identifiable, and only 4 out of 16 parameters are insensitive for a particular region in δ. Further, the model has a moderate multi-scale sloppiness ratio inside the specified δ, which indicates only an acceptable isotropic sensitivity in the parameter directions.

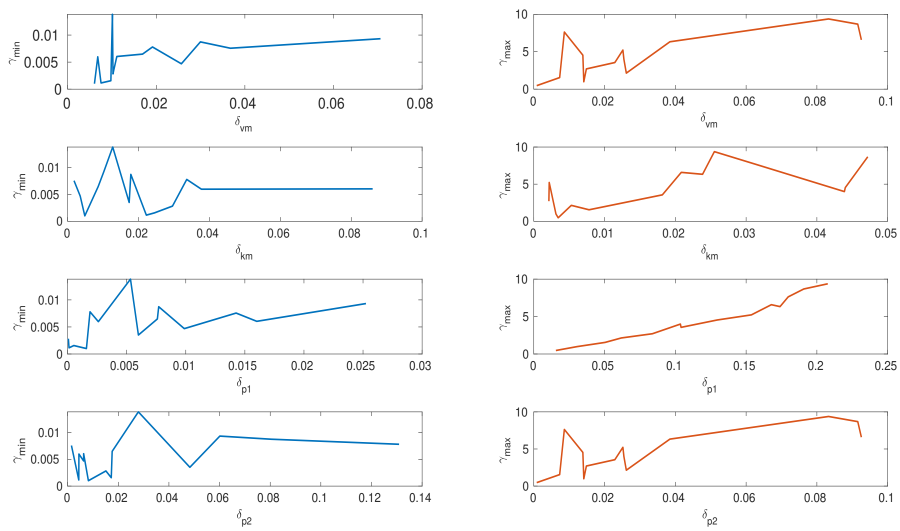


Figure 9: The x-axis depicts the relative distance of the particular parameter θ_i from its reference value θ_i^*. The y-axis on the left is the minimum sum-square deviation and on the right maximum sum-square deviation from $y^*(t)$. Parameters v_m and k_m are insensitive for $\delta_i > 0.03$. The maximum sensitivity of parameter p_1 is linearly increasing as δ_i increases.

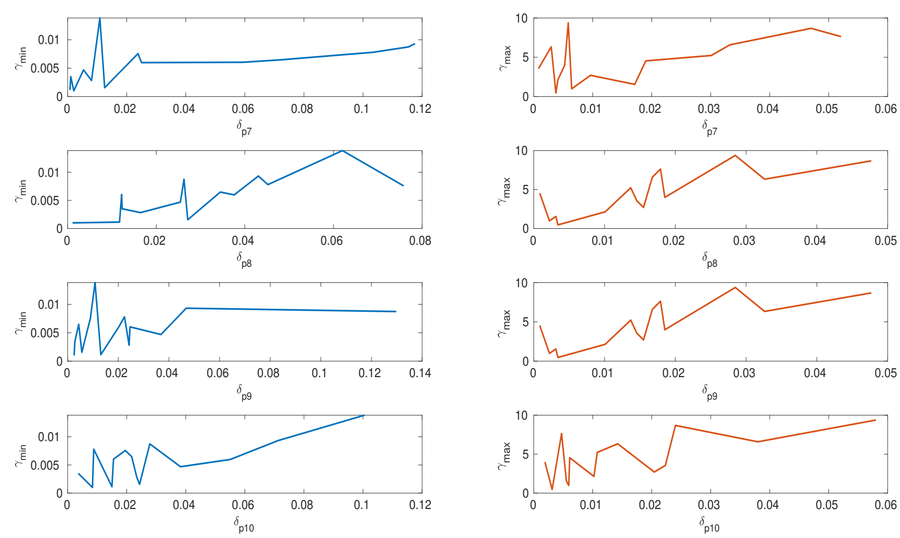


Figure 11: The x-axis depicts the relative distance of the particular parameter θ_i from its reference value θ_i^*. The y-axis on the left is the minimum sum-square deviation and on the right maximum sum-square deviation from $y^*(t)$. The parameters p_7 is insensitive for $\delta_i > 0.02$ and p_9 for $\delta_i > 0.05$

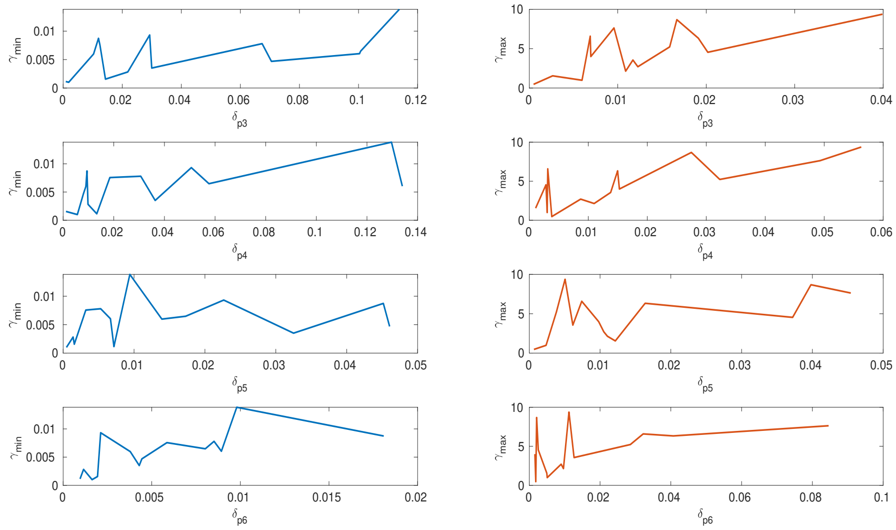


Figure 10: The x-axis depicts the relative distance of the particular parameter θ_i from its reference value θ_i^*. The y-axis on the left is the minimum sum-square deviation and on the right maximum sum-square deviation from $y^*(t)$. While non of the parameters is insensitive, parameters p_5 and p_6 are highly sensitive in the vicinity of the $\theta_i^*(\delta_i^* < 0.01)$

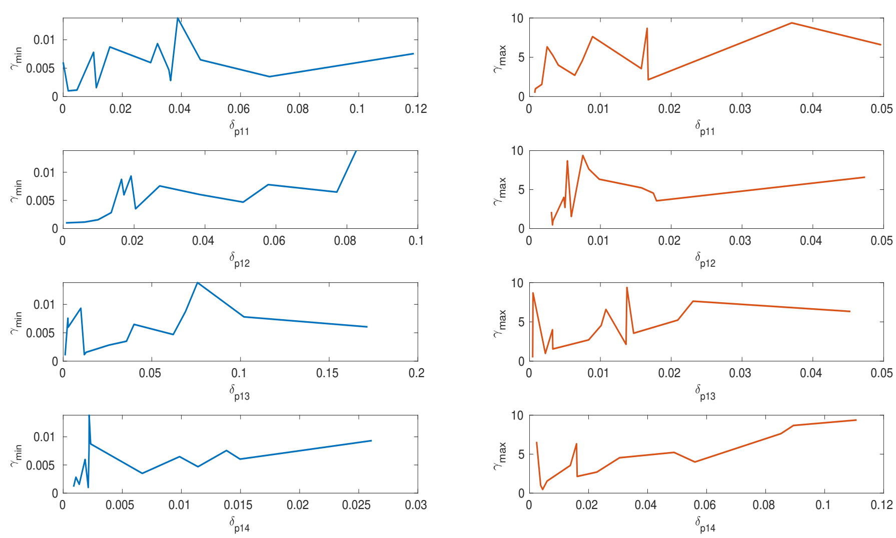


Figure 12: The x-axis depicts the relative distance of the particular parameter θ_i from its reference value θ_i^*. The y-axis on the left is the minimum sum-square deviation and on the right maximum sum-square deviation from $y^*(t)$. The parameters p_{13} and p_{14} are highly sensitive in the vicinity of θ_i^*.

.

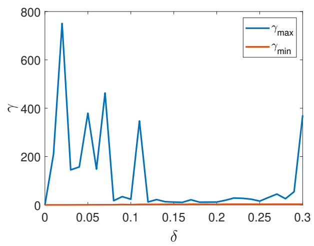
Table 4: Summary of model features analysed in this study

Model	States	Parameters	Identifiable	Insensitive parameters
State-space: Case 1	2	2	Yes	-
State-space: Case 2	2	2	Yes	θ_2
State-space: Case 3	2	2	No	-
Minimal Cascade	2	10	Yes	k_d, v_d, K_1, K_2, V_4
Biochemical pathway	14	16	Yes	v_m, k_m, p_7, p_9

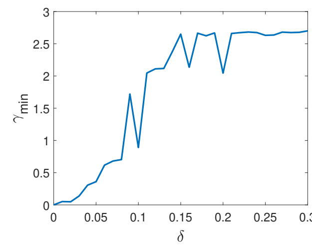
Mitotic oscillator

Minimal cascade model for mitotic oscillator is a ordinary differential equation model with 3 states and 10 parameters. This model is one of the 16 system biology models that were shown to be sloppy [6]. The original model and the nominal parameters were obtained from [28]. Here, we analyze the behavior of the model by constructing the visual plot. The model equations are given in (24)

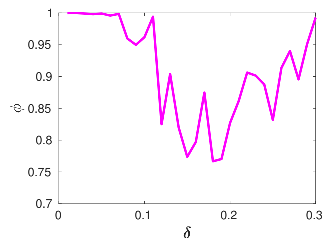
$$\mathcal{M}: \begin{cases} \frac{dC}{dt} = v_i - k_d C - v_d X \frac{C}{K_d + C} \\ \frac{dM}{dt} = \frac{V1(1-M)}{(K1+(1-M))} - \frac{V_2 M}{K_2 + M} \\ \frac{dX}{dt} = \frac{V3(1-X)}{(K3+(1-X))} - \frac{V_4 X}{X + K_4} \\ y(t) = C(t) + M(t) + X(t) \\ X(0) = M(0) = C(0) = 0.01 \end{cases} \tag{24}$$



(a) δ vs γ plot



(b) δ vs γ_{min} plot



(c) δ vs ϕ plot

Figure 13: (a) The curves indicate non-uniform sensitivity in the given δ (b) The system is initially sloppy till $0 < \delta < 0.03$ and becomes non-sloppy for $0.05 < \delta < 0.15$ and again sloppy for $0.15 < \delta < 0.3$. The system is locally structurally identifiable

The true parameters around which the model's behaviour are analysed is taken from [28]. It can be observed that from Fig. 13a that for $\delta < 0.1$, there is a huge asymmetry between minimum and maximum deviation, which implies that the system will be extremely sloppy with respect to traditional definition of sloppiness (2) and multi-scale sloppiness [12]. From Fig. 13b, it can be seen that as δ increases, γ_{min} also increases but goes nearly flat after $\delta > 0.12$, indicating negligible change in the γ_{min} as δ changes, which implies that the model is (δ, ϵ) sloppy for $\delta > 0.12$.

Fig. 14 and Fig. 15 show how γ_{min} and γ_{max} changes for each parameter as the absolute value changes. It can be clearly seen that parameters $k_d, v_d, K1, K_2$ and V_4 are insensitive as there is no significant change in γ_{min} for the relative change $\delta_i > 0.01$. On the other hand, the parameters K_3, V_i, V_2 and V_4 are highly sensitive for $\delta_i < 0.01$. This gives us a good idea of how the system behave with respect to the changes in specific parameter intervals. This can be used to fix initial values of parameter during an estimation exercise to avoid the sloppy region; which is one of the crucial challenges in a sloppy model [13]. Our method can give a good range of initial parameter values.

In summary, we found that (i) the model is locally structurally identifiable (ii) the model is sloppy in the traditional sense of sloppiness also from the proposed definition of sloppiness (iii) our method have identified insensitive parameters and most sensitive parameters (iv) the proposed method have also identified an interval for each parameter (θ_i) over which the model is most and least sensitive.

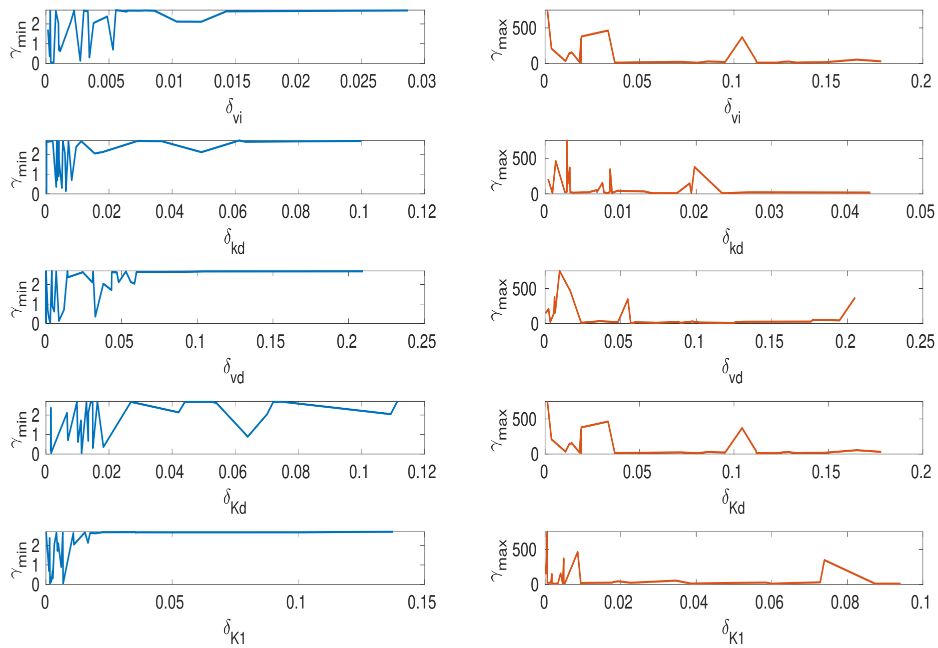


Figure 14: γ_{min} and γ_{max} changes with respect to the absolute change in each parameter v_i, k_d, v_d, K_d, K_1 from reference values. The parameters $K1$ and v_d are highly insensitive within the given δ

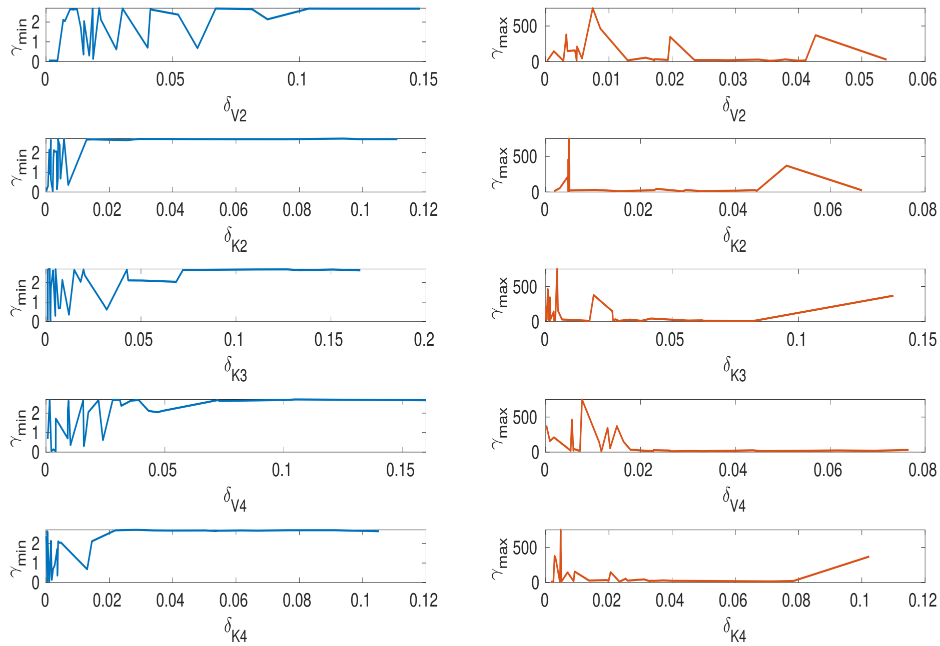


Figure 15: The figure shows how the γ_{min} and γ_{max} changes with respect to the absolute change in each parameters V_2, K_2, K_3, V_4, K_4 from reference values. The parameters $K_4, V_4 \& K_2$ are insensitive and contributes to the sloppy direction.

Methodology

In this section, we propose a visual tool to detect and quantify conditional sloppiness and detect the loss of structural identifiability for generic ordinary differential equation models.

Mathematical Model Formulation

We consider the model of the form,

$$\dot{\mathbf{x}}(\mathbf{t}) = f(\mathbf{x}(\mathbf{t}), \theta, \mathbf{u}(\mathbf{t})) \tag{25}$$

$$\mathbf{y}(\mathbf{t}) = h(\mathbf{x}(\mathbf{t}), \theta)$$

$$\mathbf{x}(\mathbf{t_0}) = \mathbf{x_0}(\theta)$$

where $\mathbf{x} = (x_1, x_2, \ldots x_{n_x}) \in R^{n_x}$ is a state vector. $\mathbf{u} = (u_1, u_2, \ldots u_{n_u}) \in R^{n_u}$ is an n_u-dimensional input vector, and $\mathbf{y} = (y_1, y_2, \ldots y_{n_y}) \in R^{n_y}$ is an n_y-dimensional output vector. The vector $\theta = (\theta_1, \theta_2, \ldots \theta_{n_\theta} \in R^{n_\theta})$ is the vector of parameters. The system has state function ($\mathbf{f}$) and the observation function ($\mathbf{h}$). The observation function can be modified by an experiment scheme, whereas the state function is fixed.

Proposed method

The visual tool is based on the definition of sloppiness proposed in the previous section. The primary idea is to study the behaviour of the model structure around the point of interest in the parameter space. A Euclidean ball of radius δ is sampled around the parameter of interest θ^* using a multivariate Gaussian distribution. The radius δ is subdivided into l equally spaced segments. The behaviour of the model is evaluated in each of the sub-Euclidean balls. The maximum and minimum deviation from the chosen parameter vector is plotted against the radius vector.

Procedure

1. Divide the radius δ into l equal segments, δ as δ_k, $k = 1, 2.....l$

2. Fix the sample size N for $\delta_k = 1$ and sample parameters from an Euclidean ball B of radius δ around θ^* in the parameter space using uncorrelated multivariate Gaussian with standard deviation as δ_k.

$$B(\theta^*, \delta) = \{\theta | \, \|\theta^* - \theta\|_2 \leq \delta\} \tag{26}$$

3. Simulate the model output $y^*(t)$ at the optimal/true parameter θ^*

4. Simulate the model output for all the parameters in the Euclidean ball and compute the sum square error γ with the optimal output $y^*(t)$.

$$\gamma = \sum_{t=1}^{N} (y^*(t, \theta^*) - y(t, \theta))^2$$

5. Compute γ_{min} and γ_{max} from each increment δ_k

$$\gamma_{min_k} = \min \sum_{t=1}^{N} (y^*(t, \theta^*) - y(t, \theta))^2$$

,

$$\gamma_{max_k} = \max \sum_{t=1}^{N} (y^*(t, \theta^*) - y(t, \theta))^2$$

6. Update sample size

$$N(k+1) = N(k) + \alpha \left(\frac{\delta_{k+1}}{\delta_k} \right)^n \tag{27}$$

7. Repeat steps 4 to 6 while $k \leq l$.
8. Compute the model sensitivity index

$$\psi = 1 - \frac{\gamma_{min}}{\gamma_{max}} \tag{28}$$

9. Plot (δ,γ_{min}) and (δ,ψ)

Sampling n-ball using multivariate Gaussian distribution is one of the efficient methods. However, for a sufficiently large dimension, with a high probability, the distance between all points will be the same, and the volume of the n-ball goes to zero [29]. To overcome this issue, we need to generate points from independent Gaussian distribution and normalise the each vector [29]. While increasing the radius δ_k, the sample size N has to be increased to avoid missing the regions of unidentifiability. For each increment in δ_k, we derived Eq. (27) to update the sample size. The parameter α can be used as a turning parameter to attain certain accuracy in the sampling and n is the dimension of the parameter space. Figure 16 illustrates the procedure to construct the visual tool.

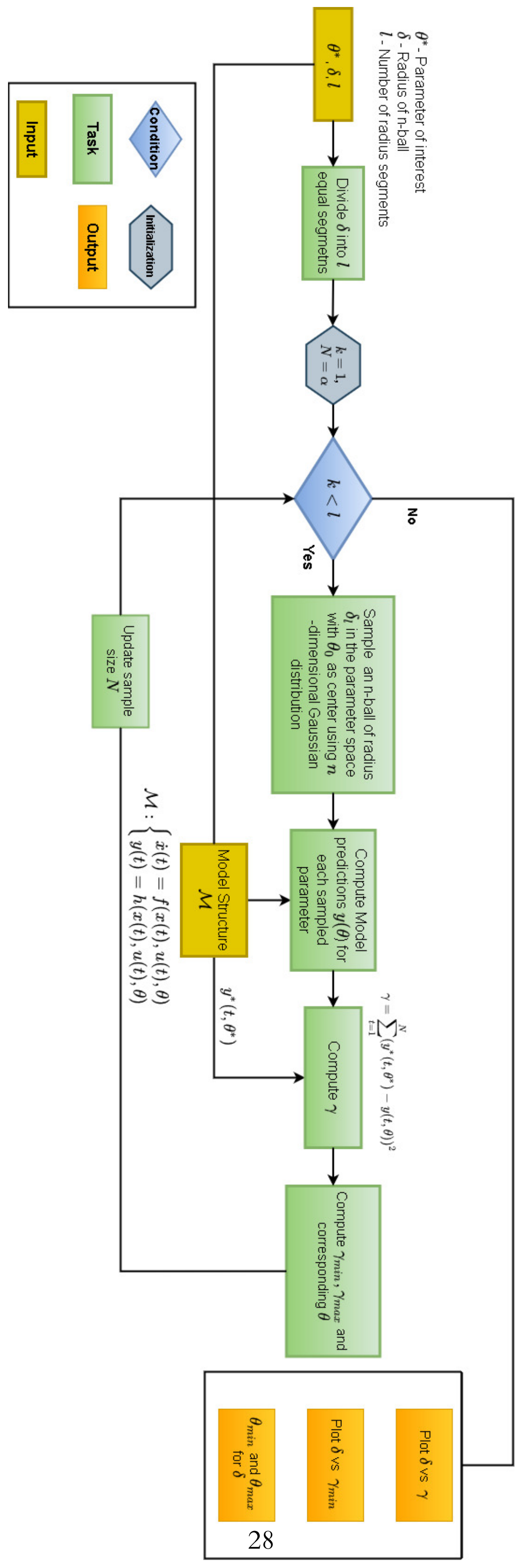


Figure 16: Workflow to construct the visual tool

Discussion

Sloppiness, practical identifiability, and structural identifiability are the most frequently encountered challenges in computational modelling, particularly on complex dynamical systems. Assessing the model structure for sloppiness and parameter unidentifiability becomes imperative for successful parameter estimation. The concept of structural identifiability is well-established in the domain of system identification, and hence a great deal of work has been done on assessing the structural identifiability of a model structure. However, most analytical methods are limited to a specific type of non-linearity and miniature models. A few numerical methods also have been developed to assess structural identifiability. On the other hand, sloppiness has not been investigated with similar rigour.

Though there has been considerable discussion on sloppiness in the literature, there were still a few crucial unanswered questions: (i) what the source of sloppiness is? (ii) what is the relationship of sloppiness with parameter uncertainty / practical identifiability? (iii) does the measure of sloppiness indicate the goodness of the estimated parameters? This work provides definitive answers to all of these questions.

We position the notion of sloppiness relative to well-established concepts such as structural identifiability and practical identifiability. Further, we show by simulation studies the challenges in applying the current measure of sloppiness in a modeling exercise and propose to argue that the ambiguity in understanding sloppiness and related questions are due to the lack of mathematical formalism. We developed two new theoretical definitions of sloppiness, namely sloppiness and conditional sloppiness. Conditional sloppiness is conditioned on the experiment space. Using the proposed definition of sloppiness, we showed that the linear predictors cannot be sloppy but eventually become unidentifiable. A mathematical relationship between practical identifiability and conditional sloppiness has been derived for generalized linear predictors.

A numerical method is proposed to assess the conditional sloppiness for generalized non-linear predictors. The proposed method helps determine the model's behaviour around a point of interest in the parameter space. An n-ball of radius δ is constructed with the centre as a reference parameter. Deviation in the prediction from the reference parameter is computed for all the parameters sampled from the n-ball. The plot shows the minimum and maximum deviations in the prediction of the delta increases. The proposed tool also helps find the most sensitive and least sensitive parameters with an interval. The method can also detect structural unidentifiability

The proposed tool is applied to three different models, including a hallmark system biology model. The proposed method detected sloppiness, structural unidentifiabilities, sensitive and insensitive parameters. The analysis gives a holistic picture of the system's behaviour in a subset of a region in the parameters space. The parameter interval obtained from the proposed method can be used to fix the initial parameter values for the parameter optimization in sloppy models, primarily to avoid flat regions where optimization algorithms may get stuck.

In summary, we see four crucial contributions in this study (i) rightly positioning the concept of sloppiness in relationship with identifiability (ii) finding the challenges in the current measure of sloppiness (iii) proposed a new mathematical definition of sloppiness (iv) a unified framework to assess sloppiness, structural identifiability, and parameter sensitivity.

References

[1] Jonathan R Karr, Jayodita C Sanghvi, Derek N Macklin, Miriam V Gutschow, Jared M Jacobs, Benjamin Bolival, Jr, Nacyra Assad-Garcia, John I Glass, and Markus W Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401, July 2012.

[2] Guan-Sheng Liu, Richard Ballweg, Alan Ashbaugh, Yin Zhang, Joseph Facciolo, Melanie Cushion, and Tongli Zhang. A quantitative systems pharmacology (qsp) model for pneumocystis treatment in mice. *BMC Systems Biology*, 12, 12 2018.

[3] Rukmini Kumar, Kannan Thiagarajan, Lakshmanan Jagannathan, Liming Liu, Kapil Mayawala, Dinesh de Alwis, and Brian Topp. Beyond the single average tumor: Understanding io combinations using a clinical qsp model that incorporates heterogeneity in patient response. *CPT: pharmacometrics & systems pharmacology*, 10(7):684–695, Jul 2021. 33938166[pmid].

[4] Arun Tangirala. *Principles of System Identification: Theory and Practice*. 10 2018.

[5] R. Bellman and K.J. Åström. On structural identifiability. *Mathematical Biosciences*, 7(3):329–339, 1970.

[6] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Computational Biology*, 3(10), 2007.

[7] Oana Teodora Chis, Julio R. Banga, and Eva Balsa-Canto. Structural identifiability of systems biology models: A critical comparison of methods. *PLoS ONE*, 6(11), 2011.

[8] Alejandro F. Villaverde, Neil D. Evans, Michael J. Chappell, and Julio R. Banga. Input-dependent structural identifiability of nonlinear systems. *IEEE Control Systems Letters*, 3(2):272–277, 2019.

[9] Ricky Chachra, Mark Transtrum, and James Sethna. Comment on "sloppy models, parameter uncertainty, and the role of experimental design". *Molecular BioSystems*, 7:2522; author reply 2523–4, 04 2011.

[10] Mario Castro and Rob J. de Boer. Testing structural identifiability by a simple scaling method. *PLoS Computational Biology*, 16(11):1–15, 2020.

[11] Alejandro F. Villaverde, Antonio Barreiro, and Antonis Papachristodoulou. Structural identifiability of dynamic systems biology models. *PLOS Computational Biology*, 12(10):1–22, 10 2016.

[12] Dhruva V. Raman, James Anderson, and Antonis Papachristodoulou. Delineating parameter unidentifiabilities in complex models. *Physical Review E*, 95(3), mar 2017.

[13] Mark K. Transtrum, Benjamin B. MacHta, and James P. Sethna. Why are nonlinear fits to data so challenging? *Physical Review Letters*, 104(6):2–5, 2010.

[14] Ryan N. Gutenkunst, Fergal P. Casey, Joshua J. Waterfall, Christopher R. Myers, and James P. Sethna. Extracting falsifiable predictions from sloppy models. *Annals of the New York Academy of Sciences*, 1115:203–211, 2007.

[15] Brian K. Mannakee, Aaron P. Ragsdale, Mark K. Transtrum, and Ryan N. Gutenkunst. *Sloppiness and the Geometry of Parameter Space*, pages 271–299. Springer International Publishing, Cham, 2016.

[16] Bryan C. Daniels, Yan Jiun Chen, James P. Sethna, Ryan N. Gutenkunst, and Christopher R. Myers. Sloppiness, robustness, and evolvability in systems biology. *Current Opinion in Biotechnology*, 19(4):389–395, 2008.

[17] J Waterfall. *Universality in Multiparameter Fitting: Sloppy Models*. PhD thesis, 01 2006.

[18] Andrew White, Malachi Tolman, Howard D Thames, Hubert Rodney Withers, A Mason, and Mark K Transtrum. The Limitations of Model-Based Experimental Design and Parameter Estimation in Sloppy Systems. *PLoS Computational Biology*, pages 1–26, 2016.

[19] Oana-teodora Chis, Alejandro F Villaverde, Julio R Banga, and Eva Balsa-canto. Mathematical Biosciences On the relationship between sloppiness and identifiability. *Mathematical Biosciences*, 282:147–161, 2016.

[20] Joseph distefano iii, dynamic systems biology modeling and simulation. *Science Progress*, 102(4):378–378, 2019.

[21] Attila Gábor, Alejandro F. Villaverde, and Julio R. Banga. Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems. *BMC Systems Biology*, 11(1):1–16, 2017.

[22] Ivan Borisov and Evgeny Metelkin. Confidence intervals by constrained optimization—an algorithm and software package for practical identifiability analysis in systems biology. *PLOS Computational Biology*, 16(12):1–13, 12 2020.

[23] Christian Tönsing, Jens Timmer, and Clemens Kreutz. Cause and cure of sloppiness in ordinary differential equation models. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(2):1–16, 2014.

[24] Joshua F. Apgar, David K. Witmer, Forest M. White, and Bruce Tidor. Sloppy models, parameter uncertainty, and the role of experimental design. *Molecular BioSystems*, 6(10):1890–1900, Oct 2010. 20556289[pmid].

[25] Joshua Waterfall, Fergal Casey, Ryan Gutenkunst, Kevin Brown, Chris Myers, Piet Brouwer, Veit Elser, and James Sethna. Sloppy-model universality class and the vandermonde matrix. *Physical review letters*, 97:150601, 11 2006.

[26] Alejandro Villaverde and Julio Banga. Reverse engineering and identification in systems biology: Strategies, perspectives and challenges. *Journal of the Royal Society, Interface / the Royal Society*, 11:20130505, 02 2014.

[27] Prem Jagadeesan, Karthik Raman, and Arun K. Tangirala. A new index for information gain in the bayesian framework. *IFAC-PapersOnLine*, 53(1):634–639, 2020. 6th Conference on Advances in Control and Optimization of Dynamical Systems ACODS 2020.

[28] A. Goldbeter. A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. *Proceedings of the National Academy of Sciences of the United States of America*, 88(20):9107–9111, Oct 1991. 1833774[pmid].

[29] Avrim Blum, John Hopcroft, and Ravindran Kannan. *Foundations of Data Science*. Cambridge University Press, 2020.