GBCB5874 report: Development and application of sparse identification for nonlinear dynamic systems

Abstract

With the development of technologies in biology, a wealth of experimental data has been accumulated. Identifying the underlying mechanism from data is a challenge for complex biological systems. In this project, we applied SINDy, an algorithm to discover governing equations from data to a chaotic system, Michaelis Menten kinetics, and epidemiology-based models. Additionally, IIR filters were used to denoisenoisy data in this work. We also applied Euclidean distance and information criteria, which balanced the accuracy and complexity, to evaluate the sparse regression for model selection. IIR filters can effectively get rid of noises; model selection works well for differentiating the optimal model from a couple of candidates. This work recovered all coefficients of equations, although two redundant terms can not be screened in a SEIR model with a tiny parameter.

1 Introduction

Identifying governing equations in diverse complex dynamic systems is of predominant importance to understand, access, and improve modern technologies that make better lives. For example, system identification leads to a better understanding of many novel fields, such as studying enzymatic reactions [1], and predicting the spread of disease [1, 2]. In enzyme kinetics, the mathematical equations which indicate the interactions among enzymes, substrates, and final products can be determined by observing the change in protein half-life. Understanding the governing equations in epidemiology helps us learn the molecular basis about the virus - how it is transmitted to others, what is the replication rate, etc. Model predictions made through the governing equations in disease transmission models provide the information to the public and authorities to a judgment call as well, such as when to do self-quarantine, the decision to economic lockdown, and the risk of re-opening.

However, the central challenge is how to detect the elusive laws from true dynamic systems with abundant noise. The emergence of data-driven methods affords an alternative approach for discovering the equations governing the complex dynamics. From this perspective, there is a potential for biologists to identify the system that controls a specific biological network purely from measured time-series data. In recent years, a data-driven algorithm discovering governing equations for nonlinear dynamic systems, SINDy [3], is proposed and being developed. SINDy is based on the sparse regression that measures the most likely relationship between the term library, in which the true dynamic system is a subset, and the experimental data. The ultimate goal for this algorithm is to find out the linear combination of the fewest states in the library with the lowest error estimation.

In this project, we identified several classic biological systems, including Michaelis–Menten kinetics and epidemiology models, based on sparse identification of nonlinear dynamical systems (SINDy) [3] and regression models [4]. We also applied SINDy to an epidemiology-based system which is to study the transmission of antibiotic genes. Gaussian noise is added into generated time course data to mimic noisy data in real scenarios; efficient digital IIR filters are applied to preprocessing noisy data before identification of equations. Additionally, parsimonious models were evaluated by Pareto analysis, Euclidean distance, and information criteria [1].

2 Material and Methods

SINDy [3] is a data-driven algorithm to discover governing equations of nonlinear dynamical systems from noisy measurement data.

2.1 Data Collection and Preprocessing

Time-series data of a dynamic system are generated from existing models or collected from experiments, which are denoted as $\mathbf{x}(t_i) \in \mathbb{R}^n$. The measurement data and the corresponding derivatives are arranged into two large matrices $\mathbf{X} = [x(t_1)x(t_2)...x(t_b)]$ (Eq. 1) and $\dot{\mathbf{X}} = [\dot{x}(t_1)\dot{x}(t_2)...\dot{x}(t_b)]$ (Eq. 2). Cross validation is required in some advanced sparse identification methods, therefore data are divided into a training set $\mathbf{X}_\mathrm{T} \in \mathbb{R}^{m\times n}$ and a validation set $\mathbf{X}_\mathrm{V} \in \mathbb{R}^{v\times n}$ if needed, where $b = m + v$.

$$\mathbf{X} = \begin{bmatrix} x(t_1) \\ x(t_2) \\ \vdots \\ x(t_b) \end{bmatrix} = \begin{bmatrix} x_1(t_1) & x_2(t_1) & \dots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \dots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_b) & x_2(t_b) & \dots & x_n(t_b) \end{bmatrix} \tag{1}$$

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{x}(t_1) \\ \dot{x}(t_2) \\ \vdots \\ \dot{x}(t_b) \end{bmatrix} = \begin{bmatrix} \dot{x}_1(t_1) & \dot{x}_2(t_1) & \dots & \dot{x}_n(t_1) \\ \dot{x}_1(t_2) & \dot{x}_2(t_2) & \dots & \dot{x}_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{x}_1(t_b) & \dot{x}_2(t_b) & \dots & \dot{x}_n(t_b) \end{bmatrix} \tag{2}$$

In order to simulate real scenarios, which usually include noises, we add different levels of Gaussian noises to generated data:

$$x_{new} = x + \epsilon \times \text{Gaussian noise}$$

where ϵ represents the level of noise.

Infinite impulse response (IIR) filters are used to reduce noise in data preprocessing.

2.2 Basic SINDy

The general characteristics of SINDy [3] is that most dynamical systems have only a few relevant terms, making the governing equations sparse in a nonlinear function space. A set of candidate nonlinear functions are integrated into a library $\Theta(\mathbf{X})$:

$$\Theta(\mathbf{X}) = \begin{bmatrix} 1 & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \ldots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \ldots \end{bmatrix} \tag{3}$$

where $\Theta(\mathbf{X})$ may consist of constant, polynomial, and trigonometric functions. Each column of $\Theta(\mathbf{X})$ represents a candidate function for $\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t))$.

In order to decide which nonlinearities are active, a sparse regression problem is set up to calculate the vectors of coefficients $\Xi = [\xi_1 \xi_2 \cdots \xi_n]$, where

$$\dot{\mathbf{X}} = \Theta(\mathbf{X})\Xi \tag{4}$$

Therefore, a set of governing equations can be constructed as follows:

$$\dot{\mathbf{x}}_k = \mathbf{f}_k(\mathbf{x}) = \Theta(x^T)\xi_k \tag{5}$$

The general steps of SINDy is summarized in Fig. 1. If there is no rational function in the dynamic system, construct a basic library $\Theta(\mathbf{X})$ based on denoised time course data and calculate coefficients Ξ of terms within the library to satisfy Eq. 4.

2.3 Advanced SINDy

Many dynamic systems have rational functions $\dot{x_k} = \frac{f_N(x)}{f_D(x)}$, where $f_N(x)$ and $f_D(x)$ are the numerator and denominator functions of variable x respectively. Therefore, a more general library is required to describe rational functions. Rational functions can be rewritten as follows:

$$f_N(x) - f_D(x)\dot{x_k} = 0 \tag{6}$$

Then the library is updated to fit rational functions, such Michaelis–Menten kinetics, as follows:

$$\Theta(\mathbf{X}, \dot{\mathrm{X}}) = \begin{bmatrix} 1 & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \ldots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \ldots & \mathrm{X}\dot{X} & \ldots \end{bmatrix} \tag{7}$$

where $\Theta(\mathbf{X}, \dot{\mathrm{X}})\Xi = 0$. To find the sparse vector of coefficients Ξ satisfying $\Theta(\mathbf{X}, \dot{\mathrm{X}})\Xi = 0$, it is necessary to calculate the null space of Θ:

$$\mathrm{N} = null(\Theta) \tag{8}$$

where Ξ is in the null space. Qu et al. [4] proposed an algorithm based on the alternating directions method (ADM) to identify the sparsest vector in a subspace. ADM is integrated in the advanced SINDy [1] to find the sparsest vector in the null space of Θ for biological systems with rational functions.

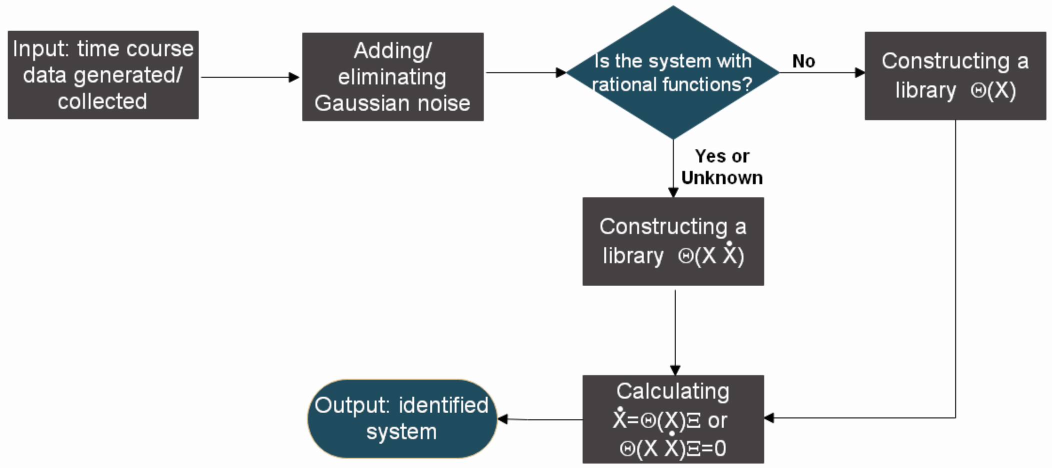

Input: time course data generated/ collected
Adding/ eliminating Gaussian noise
Is the system with rational functions?
No
Constructing a library Θ(X)
Yes or Unknown
Constructing a library Θ(X Ẋ)
Calculating Ẋ=Θ(X)Ξ or Θ(X Ẋ)Ξ=0
Output: identified system

Figure 1: **Framework of SINDy algorithm.** Input is time series data which are from existing models or experiments. If the system has no rational functions, the basic SINDy is applied, where $\Theta(\mathbf{X})$ is a collection of nonlinear functions of X; otherwise, the advanced SINDy with advanced library $\Theta(\mathbf{X}, \dot{\mathrm{X}})$ is applied, where $\Theta(\mathbf{X}, \dot{\mathrm{X}})$ is a collection of functions of X and $\dot{X}$. Identified sparse coefficients are calculated by Eq. 4 for systems without rational functions or $\Theta(\mathbf{X}, \dot{\mathrm{X}})\Xi = 0$ for rational functions.

2.4 Model Selection

Selecting the best model from a large set of candidates is challenging, especially for complicated systems. Therefore, further model selection algorithms can be applied after identifying a subset of models via SINDy [2], which balance the parsimony and relative information loss across models.

A threshold λ is required to select significant values in Ξ as the active terms, which means values less than λ are screened. As the appropriate λ is a priori unknown number, optimal λ is determined by two criteria: 1) error and 2) number of terms. The most parsimonious model at the sharp drop-off error was identified. Different candidate models are obtained from different λ.

Euclidean distance and information criteria, such as Akaike information criterion (AIC) and Bayes information criteria (BIC), were used to select models. Formulas are as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{9}$$

where $d(x, y)$ is the Euclidean distance between vector or matrix x and y.

$$AIC_j = 2k - 2\ln(L(X, \hat{\mu})) \tag{10}$$

where $L(x, \mu) = P(x|\mu)$ indicates the likelihood function; k indicates the number of free parameters [2].The AIC with correction for finite sample sizes are as follows:

$$AIC_c = AIC + 2(k+1)(k+2)/(m-k-2) \tag{11}$$

where m is the total number of data points. BIC is based on AIC as well:

$$BIC = AIC - 2k + 2k \times \ln(m) \tag{12}$$

In this project, we use a modified AIC equation [2] $AIC = m\ln(RSS/m) + 2k$ to evaluate models. RSS is the residual sum of squares given by $RSS = \sum_{i=1}^{m}(y_i - g(x_i; \mu))^2$, where y and g are the observations and the candidate model, respectively.

3 Results

3.1 Chaotic Lorenz System

The Lorenz system is used as an example for chaotic dynamics:

$$\begin{aligned} \dot{x} &= \alpha(y - x) \\ \dot{y} &= x(\rho - z) - y \\ \dot{z} &= xy - \beta z \end{aligned} \tag{13}$$

where parameters $\alpha = 10$, $\beta = 2$, and $\rho = 25$.

Based on framwork in Fig. 1, data are generated from Eq. 13 with initial values $[-8, 7, 27]$ (Fig. 2). We add noise with $\epsilon = 1$ and denoise via IIR filter (Fig. 3). A basic library is constructed for Lorenz system:

$$\Theta(\mathbf{X}) = \begin{bmatrix} 1(t) & \mathbf{z}(t) & \mathbf{y}(t) & \mathbf{x}(t) & \mathrm{z}^2(t) & \mathrm{yz}(t) & \mathrm{y}^2(t) & \dots & \mathrm{x}^2(t) & \dots & \mathrm{z}^3(t) \end{bmatrix}$$

We tried 10 different values of threshold λ ranging from 10^{-5} to 10 (Fig. 4 right), and found that most λ can identify the exact terms (Fig. 4). When the number of non-zero terms is 7, the corresponding Euclidean distance between observation and simulation drops sharply (Fig. 4 right).

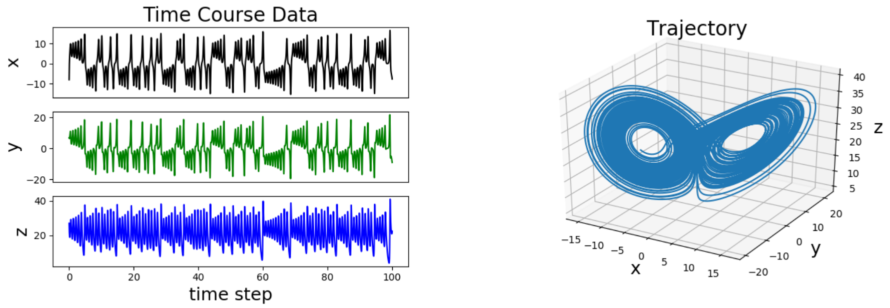


Figure 2: Time course data generated from Lorenz system (left) and the corresponding trajectory (right).

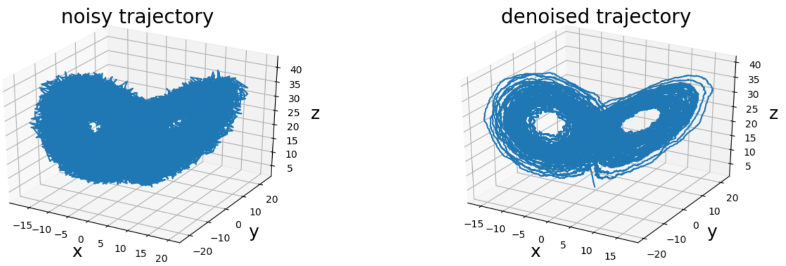
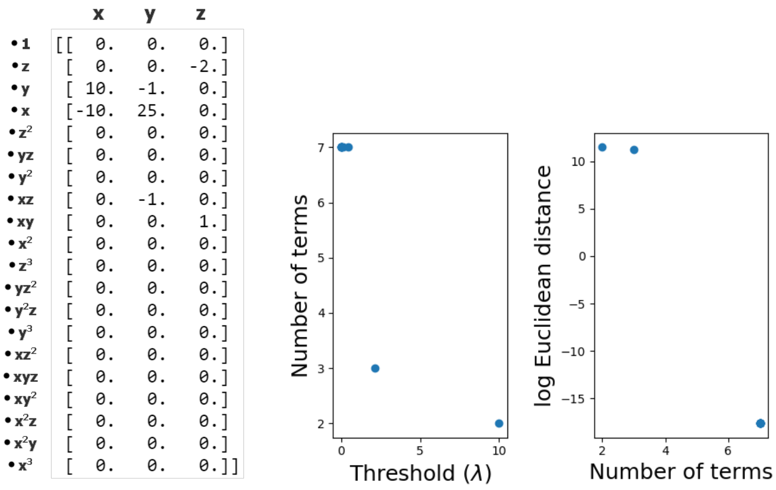


Figure 3: Trajectory of noisy Lorenz system (left) and the corresponding denoised trajectory via IIR filter (right).

3.2 Michaelis-Menten Kinetics

Michaelis-Menten (Eq. 14) is a classical mathematical model that is used to study enzyme kinetics. The following equation illustrates the dynamics of substrate([S]) based on binding affinity to enzyme. J_{in} is the source influx of the



```
[[  0.   0.   0.]
 [  0.   0.  -2.]
 [ 10.  -1.   0.]
 [-10.  25.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.  -1.   0.]
 [  0.   0.   1.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]
 [  0.   0.   0.]]
```

Figure 4: Sparse Coefficients (Ξ) of noisy Lorenz dynamics with $\lambda = 1$ (left) and The Pareto front (right). 'Number of terms' is the total number of non-zero terms in Ξ; 'log Euclidean distance' is the logarithm value of Euclidean distance between observation and simulation.

substrate, V_{max} indicates the maximum enzyme reaction rate when saturation of substrate occurs, K_m is the Michaelis constant indicating the half-life concentration of substrate.

Hyper-parameter, λ, in Eq. 15 are tuned to discover the linear combination of fewest terms in sparse library. As a result, the fewest term number of four has been successfully identified with lowest error value (Fig. 5, top right).

Additionally, Gaussian noise with 0.1 standard deviation are applied in Michaelis-Menten kinetics. This noisy data set is filtered by a Python package, Signal. After filtering, SINDy (Fig. 5, bottom left) also discovers the governing equation of enzyme kinetic (Fig. 5, bottom right), which indicates the potential application of SINDy on noisy biological data.

$$\frac{d[S]}{dt} = J_{in} - \frac{V_{max}[S]}{K_m + [S]} \tag{14}$$

where $J_{in} = 0.6$, $V_{max} = 1.5$, $K_m = 0.3$.

$$L_{lasso}(\Xi) = min_{\Xi}(\sum_{i-1}^{n}(X - \Theta(X)\Xi)_2^2 + \lambda R(\Xi)) \tag{15}$$

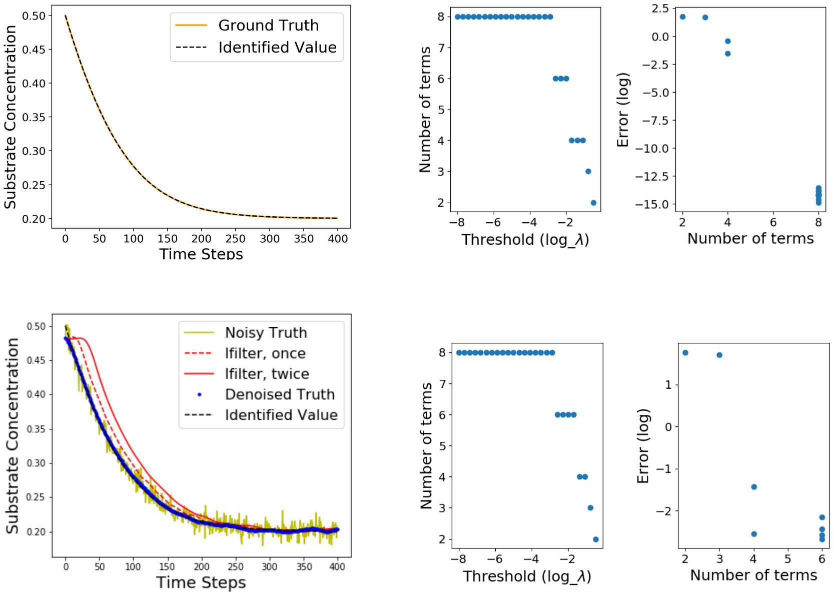


Figure 5: Time course data generated from Michaelis-Menten model(top left); evaluation on parsimonious models (top right); Noisy, denoised, and identified time course data (bottom left); evaluation on parsimonious models (bottom right). Noise level $\epsilon = 0.01$.

3.3 Epidemiology-based Models

Epidemiology models are often used to infer disease transmission rates and investigate outbreaks. This type of models can also be applied to study bacterial resistance [5]. A typical susceptible-exposed-infectious-recovered (SEIR) model and a antibiotic resistance transmission (RSM) model are studied in this section [6] (Fig. 6).

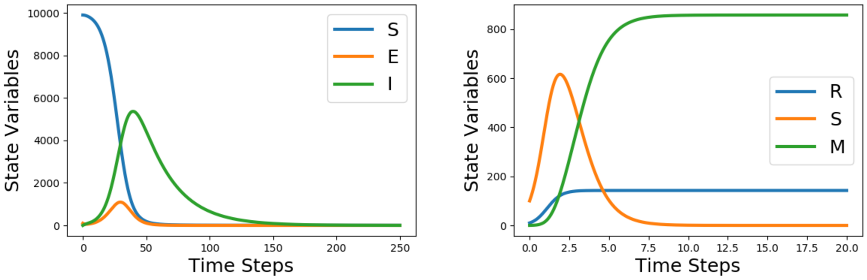


Figure 6: Time course data generated via SEI(R) model, where resistant state is not shown (left) and time course data generated via RSM model (right).

3.3.1 SEI(R) Model

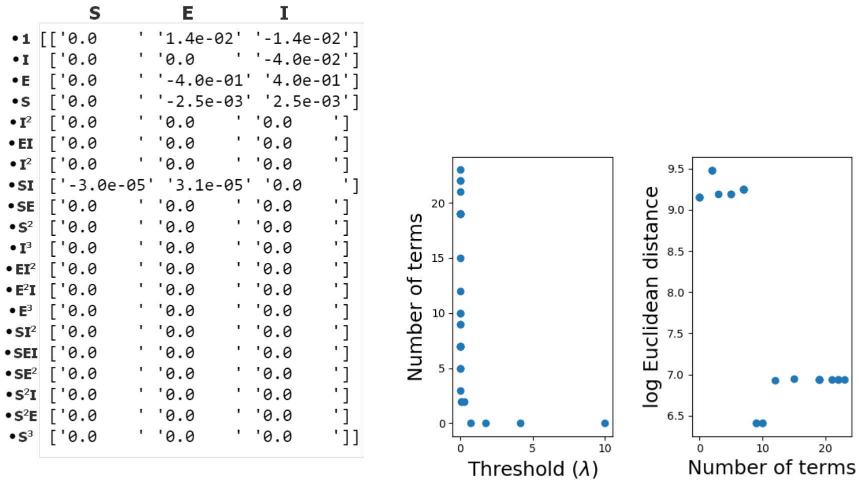
Ordinary differential equations of SEIR model are as follows:

$$\begin{array}{rcl} \frac{d\mathrm{S}}{dt} & = & \mu - \frac{\beta}{\mathrm{N}}\mathrm{SI} - \nu\mathrm{S} \\ \frac{d\mathrm{E}}{dt} & = & \frac{\beta}{\mathrm{N}}\mathrm{SI} - (\mu+\alpha)\mathrm{E} \\ \frac{d\mathrm{I}}{dt} & = & \alpha\mathrm{E} - (\mu+\gamma)\mathrm{I} \\ \frac{d\mathrm{R}}{dt} & = & \gamma\mathrm{I} - \mu\mathrm{R} + \nu\mathrm{S} \end{array} \tag{16}$$

where parameters $\mu = 0$, $\beta = 0.3$, $\mathrm{N} = 10^4$, $\nu = 0$, and $\alpha = 0.4$, $\gamma = 0.04$. As S, E, and I are not dependent on R, we ignored R state in equation identification for the moment. The first set of initial values is $[\mathrm{S} = 0.99 \times 10^4, \mathrm{E} = 0.01 \times 10^4, \mathrm{I} = 0]$. 30 values of λ ranging from 10^{-10} to 10 were used as the threshold to identify significant terms (Fig. 7).

When $\lambda = 2.04 \times 10^{-3}$, the Euclidean distance drops sharply; the corresponding coefficients Ξ are shown in Fig. 7 (top). All terms in Eq. 16 were recovered correctly by this Ξ (Fig. 7 top); however, two redundant terms arise: '1.4×10^{-2}' and '$-3 \times 10^{-3}[\mathrm{S}]$'. As $\frac{\beta}{\mathrm{N}} = 3 \times 10^{-5}$ is way smaller than these two redundant terms, it is hard to screen these redundant terms without ignoring $\frac{\beta}{\mathrm{N}}$.

We also applied information criteria, including AIC, $\mathrm{AIC_c}$, and BIC to evaluate and select identified models (Fig. 8). All evaluation Indices in Fig. 8 select the optimal model with the number of terms being 9, and we got a very similar coefficients space with Fig. 7 which is evaluated by Euclidean distance.



```
           S           E           I
•1    [['0.0      ' '1.4e-02' '-1.4e-02']
•I     ['0.0      ' '0.0     ' '-4.0e-02']
•E     ['0.0      ' '-4.0e-01' '4.0e-01']
•S     ['0.0      ' '-2.5e-03' '2.5e-03']
•I^2   ['0.0      ' '0.0     ' '0.0     ']
•EI    ['0.0      ' '0.0     ' '0.0     ']
•I^2   ['0.0      ' '0.0     ' '0.0     ']
•SI    ['-3.0e-05' '3.1e-05' '0.0      ']
•SE    ['0.0      ' '0.0     ' '0.0     ']
•S^2   ['0.0      ' '0.0     ' '0.0     ']
•I^3   ['0.0      ' '0.0     ' '0.0     ']
•EI^2  ['0.0      ' '0.0     ' '0.0     ']
•E^2I  ['0.0      ' '0.0     ' '0.0     ']
•E^3   ['0.0      ' '0.0     ' '0.0     ']
•SI^2  ['0.0      ' '0.0     ' '0.0     ']
•SEI   ['0.0      ' '0.0     ' '0.0     ']
•SE^2  ['0.0      ' '0.0     ' '0.0     ']
•S^2I  ['0.0      ' '0.0     ' '0.0     ']
•S^2E  ['0.0      ' '0.0     ' '0.0     ']
•S^3   ['0.0      ' '0.0     ' '0.0     ']]
```

Figure 7: Sparse Coefficients (Ξ) of SEI(R) model with $\lambda = 2.04 \times 10^{-3}$ (top) and the Pareto front (bottom).

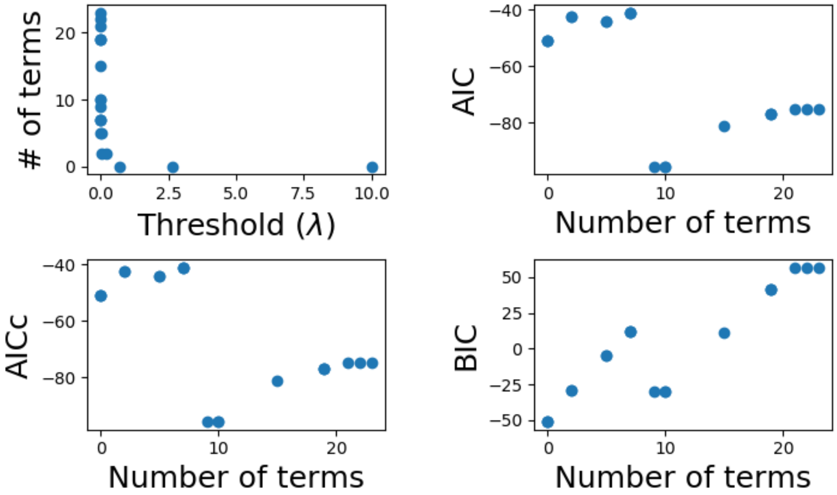


Figure 8: Model selection for SEIR model based on information criteria.

3.3.2 Antibiotic Resistance Gene (ARG) Model - RSM

We simplified RSM model in [6] by ignoring the drug, because we want to develop a model for ARG transmission outside organism. The ordinary differential equations of RSM model are as follows:

$$\begin{aligned} \frac{d\mathrm{R}}{dt} &= \alpha_{\mathrm{R}}\mathrm{R}(1 - \tfrac{\mathrm{R+S+M}}{K}) \\ \frac{d\mathrm{S}}{dt} &= \alpha_{\mathrm{S}}\mathrm{S}(1 - \tfrac{\mathrm{R+S+M}}{K}) - e\mathrm{SR} - f\mathrm{SM} \\ \frac{d\mathrm{M}}{dt} &= \alpha_{\mathrm{M}}\mathrm{M}(1 - \tfrac{\mathrm{R+S+M}}{K}) + e\mathrm{SR} + f\mathrm{SM} \end{aligned} \tag{17}$$

where R, S, M are for resistant, susceptible, and transconjugant bacteria, respectively; α_i is the growth rate of bacterial subpopulation, i =R,S,or M; K is the carrying capacity of system; e and f are the transfer rate of resistance elements from resistant or transconjugant bacteria. Parameters are shown in Table. 1. One set of initial values of R, S, M is [10,100,0].

Table 1: Parameters.

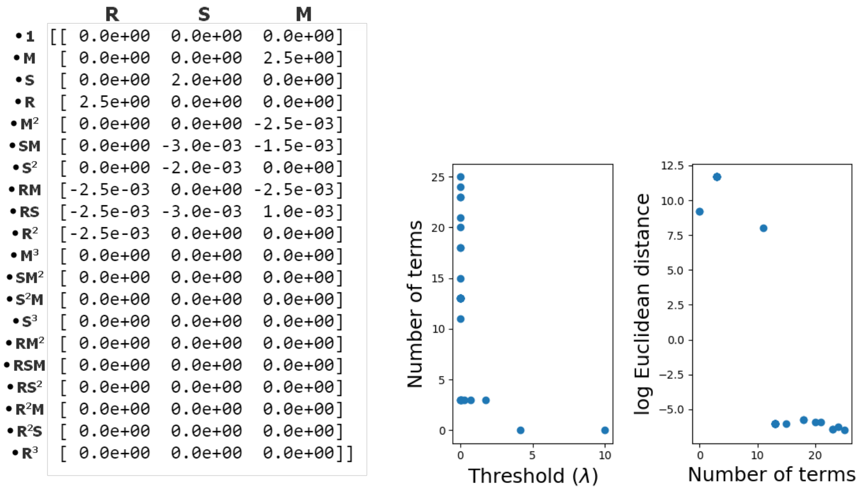
Parameters	Values
α_{R}	2.5
α_{S}	2
α_{M}	2.5
K	10^3
e, f	0.001

A set of λ ranging from 10^{-10} to 10 were used for RSM model (Fig. 9 bottom). When $\lambda = 6.7 \times 10^{-4}$, the number of identified non-zero terms drops sharply, with Ξ in Fig. 9. Governing equations of RSM were identified perfectly. Model selections based on information criteria are shown in Fig. 10. Compared with Euclidean Distance (Fig. 7, 9), Information criteria (Fig. 8,10), considering the number of non-zero terms of Ξ, can balance the accuracy and complexity to avoid over-fitting.

3.4 Discussion

Discovering underlying mechanisms of complex dynamic systems is challenging. Additionally, estimating a large number of parameters is difficult. Therefore, a reverse algorithm, SINDy, without looking into specific networks and parameters was proposed to identify governing equations only from datasets. In this work we applied SINDy into Lorenz system, Michaelis-Menten kinetics, SEIR model, and an epidemiology-based ARG transmission model.

IIR filters were introduced to denoise noisy time course data. Most models except for the SIER model, which contains very different magnitudes of parameters, are perfectly recovered. We used Euclidean distance and information criteria to rank candidate identified models and selected the optimal one based on both errors and number of terms.



```
           R         S          M
•1   [[ 0.0e+00  0.0e+00  0.0e+00]
•M    [ 0.0e+00  0.0e+00  2.5e+00]
•S    [ 0.0e+00  2.0e+00  0.0e+00]
•R    [ 2.5e+00  0.0e+00  0.0e+00]
•M^2  [ 0.0e+00  0.0e+00 -2.5e-03]
•SM   [ 0.0e+00 -3.0e-03 -1.5e-03]
•S^2  [ 0.0e+00 -2.0e-03  0.0e+00]
•RM   [-2.5e-03  0.0e+00 -2.5e-03]
•RS   [-2.5e-03 -3.0e-03  1.0e-03]
•R^2  [-2.5e-03  0.0e+00  0.0e+00]
•M^3  [ 0.0e+00  0.0e+00  0.0e+00]
•SM^2 [ 0.0e+00  0.0e+00  0.0e+00]
•S^2M [ 0.0e+00  0.0e+00  0.0e+00]
•S^3  [ 0.0e+00  0.0e+00  0.0e+00]
•RM^2 [ 0.0e+00  0.0e+00  0.0e+00]
•RSM  [ 0.0e+00  0.0e+00  0.0e+00]
•RS^2 [ 0.0e+00  0.0e+00  0.0e+00]
•R^2M [ 0.0e+00  0.0e+00  0.0e+00]
•R^2S [ 0.0e+00  0.0e+00  0.0e+00]
•R^3  [ 0.0e+00  0.0e+00  0.0e+00]]
```

Figure 9: Sparse Coefficients (Ξ) of RSM model with $\lambda = 6.72 \times 10^{-4}$ (top) and the Pareto front (bottom).

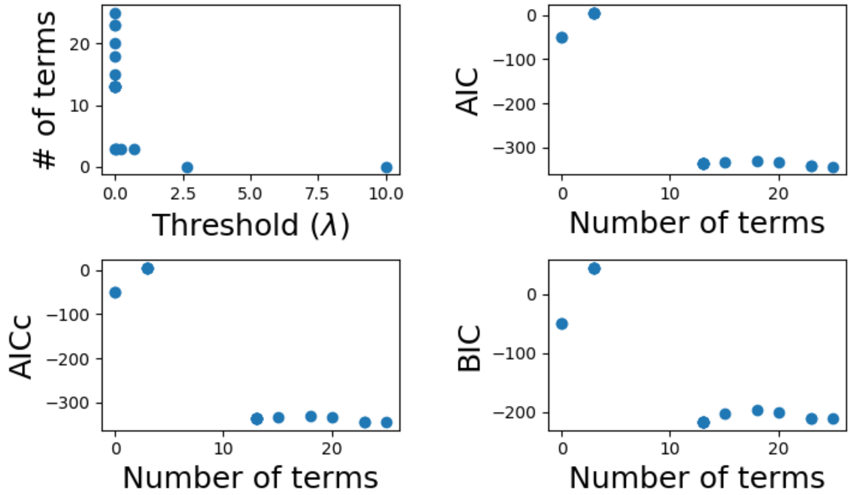


Figure 10: Model selection for RSM model based on information criteria.

Based on the application of SINDy to the SIER model, we found a limitation of current SINDy. it is hard to discover the governing equations including both tiny and huge parameters. For example, in a disease or ARG transmission model, one parameter could be 0.01, and the other could be 1000. Therefore, some redundant terms between 0.01 and 1000 are hard to screen. We thought further model selections may solve this problem; but after trying Euclidean distance and information criteria, the problem is not fixed. The sparse regression method is still being developed. One possible solution may be considering time course data locally, which means identifying equations within a short period separately, just as a hybrid dynamic system. Also, if some parameters are very tiny compared with other parameters, it might mean the corresponding terms are not significant. Increasing the threshold may improve accuracy further, though identified models would lose some terms.

Another limitation of SINDy is the biological meaning. For example, the second and third equations in RSM model contain more than one ‘SM’ term with different biological explanations. Equations identified by SINDy would lump all ‘SM‘ terms into one ‘SM’. Therefore, it is hard to biologically explain some terms obtained via SINDy.

Availability of data and materials

The datasets generated and analysed during the current study are available at
https://github.com/mircobial-evolution-dynamic/governing-equations
Please pay attention to README.md, for there are several redundant files or drafts in this repository.

References

[1] Niall M. Mangan, Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Inferring biological networks by Sparse Identification of Nonlinear Dynamics. *IEEE TRANSACTIONS ON MOLECULAR, BIOLOGICAL, AND MULTI-SCALE COMMUNICATIONS* 2.1 (2016): 52-63.

[2] Niall M. Mangan, J. Nathan Kutz, Steven L. Brunton, and Joshua L. Proctor. Model selection for dynamical systems via sparse regression and information criteria. *Proc. R. Soc.* A.473 (2017): 20170009.

[3] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *PNAS* 113.15 (2016): 3932-3937.

[4] Q. Qu, J. Sun, and J. Wright. Finding a sparse vector in a subspace: Linear sparsity using alternating directions. *Proc. Adv. Neural Inf. Process. Syst.* Montreal, QC, USA (2014): 3401-3409.

[5] Bahatdin Dasbasi, and İlhan Öztürk. Mathematical modelling of bacterial resistance to multiple antibiotics and immune system response. *Springer-Plus* 5.408 (2016): DOI 10.1186/s40064-016-2017-8.

[6] Ronette Gehring, Phillip Schumm, Mina Youssef, and Caterina Scoglio. A network-based approach for resistance transimission in bacterial populations. *Journal of Theoretical Biology* 262 (2010): 97-106.