Inferring biological networks by sparse identification of nonlinear dynamics

Niall M. Mangan*[‡], Steven L. Brunton[†], Joshua L. Proctor[‡], and J. Nathan Kutz*

*Department of Applied Mathematics, University of Washington, Seattle, WA 98195 USA

[†]Department of Mechanical Engineering and the eScience Institute, University of Washington, Seattle, WA 98195 USA

[‡]Institute for Disease Modeling, 3150 139th Ave SE Bellevue, WA 98005

Abstract—Inferring the structure and dynamics of network models is critical to understanding the functionality and control of complex systems, such as metabolic and regulatory biological networks. The increasing quality and quantity of experimental data enable statistical approaches based on information theory for model selection and goodness-of-fit metrics. We propose an alternative method to infer networked nonlinear dynamical systems by using sparsity-promoting ℓ_1 optimization to select a subset of nonlinear interactions representing dynamics on a fully connected network. Our method generalizes the sparse identification of nonlinear dynamics (SINDy) algorithm to dynamical systems with rational function nonlinearities, such as biological networks. We show that dynamical systems with rational nonlinearities may be cast in an implicit form, where the equations may be identified in the null-space of a library of mixed nonlinearities including the state and derivative terms; this approach applies more generally to implicit dynamical systems beyond those containing rational nonlinearities. This method, implicit-SINDy, succeeds in inferring three canonical biological models: Michaelis-Menten enzyme kinetics, the regulatory network for competence in bacteria, and the metabolic network for yeast glycolysis.

Index Terms—dynamical systems, machine learning, sparse regression, network inference, nonlinear dynamics, biological networks

I. INTRODUCTION

Network science is of growing and critical importance across the physical, biological and engineering sciences. In biology, both the quality and quantity of modern data has inspired new mathematical techniques for inferring the complex interactions and connections between nodes in metabolic and regulatory networks. Discovering the connectivity and structure of such networks is critical in understanding the functionality and control decisions enacted by the network in tasks such as cell differentiation, cell death, or directing metabolic flux. Methods based on information theory provide rigorous statistical criteria for such model selection and network inference tasks. For example, partnering the Kullback-Leibler (KL) divergence [1], [2], a measure of information loss between empirically collected data and model generated data, and the Akaike information criteria (AIC) [3], [4], a relative estimate of information loss across models balancing model complexity and goodness-of-fit, allows for a principled model selection criteria [5]. However, for nonlinear dynamical networks, such information theoretic approaches face severe challenges since the number of connections and possible functional forms lead to a combinatorially large number of models to be evaluated with the KL/AIC framework.

We propose an alternative method to infer the dynamics and connectivity of biological networks motivated by machine learning methods including overcomplete libraries [6]–[8] and sparse regression [9], [10]. We generalize the *sparse identification of nonlinear dynamics* (SINDy) [11] algorithm to an implicit formulation that allows the dynamics to contain rational functions. We demonstrate the accuracy and robustness of the method, *implicit-SINDy*, on three representative biological models, showing that our approach gives a compelling alternative to information-theoretic methods of model selection.

A. Biological networks

Biological networks produce a diverse range of functional activities. Regulatory and metabolic networks are critical for cellular function. Breakdown of the function and control circuits of such networks can lead to cancer and other deadly diseases, motivating attempts to control of such networks using the tools of genetic engineering and pharmacology. The network dynamics in these systems can often be modeled using mass-action kinetics, producing a relatively constrained set of network motifs [12], [13]. Generally, biological regulatory and metabolic networks are considered sparse and sparsity has been used as a criteria for inferring linear network models [14]–[17]. However, accurately and robustly characterizing such biological networks, both in terms of their unique connectivity structure and functional dynamics, remains an extremely challenging task due to the underlying nonlinear dynamics.

With the emergence of large amounts of high-quality experimental data [18], [19], new opportunities exist for data-driven mathematical modeling of these biological networks. Rapid, robust and accurate model identification can greatly accelerate our understanding and control of critical biological network functions so that disease treatment and metabolic engineering protocols can be proposed. Indeed, the rich, dynamic data on such biological networks [18], [19] allows for greatly improved statistical and machine learning methods that can help characterize the fundamental dynamic interactions in complex, biological networks.

B. Model selection and information theory

Biological network models can be developed using principled dynamical systems techniques based upon conjectured relationships between variables (nodes). Such models are constructed and iteratively modified to be consistent with experiment. Historically, such trial-and-error development could take many years before a model is quantitatively predictive. Statistical methods advocate techniques where information criteria for model selection assigns a numerical value for each candidate model. The models with the lowest values correspond to the least information loss with respect to all candidate models considered. Common examples of information criteria are the AIC [3] and the Bayesian information criteria (BIC) [20]. See [21] for a review of modern usage of information criteria.

Unfortunately, information criteria rely on the practitioner positing a set of reasonable models. If the network interactions and functional nonlinearities are unknown, then it is unlikely that one can guess the correct network structure and nonlinearities in the system. Of course, one may select from a large space of trial models, but each individual model would need to be simulated, compared to the data, and given a numerical information criteria score. This can be computationally intractable as the number of nonlinear interactions between variables scales combinatorially with the number of variables.

To highlight the complexity of model selection, consider selecting a specific model from all possible polynomials of degree d in n variables. For n variables, the number of possible monomials with degree less than or equal to d is given by

$$N_m = \binom{n+d}{d}.\tag{1}$$

The number of possible polynomial structures that may be formed by assigning nonzero coefficients to these N_m monomials is given by:

$$N_p = \sum_{k=1}^{N_m} \binom{N_m}{k}.$$
 (2)

The number of possible polynomial structures N_p may be thought of as summing over all possible polynomials with only k monomials. For example, consider polynomials up to degree d=4 with n=5 variables, as is required for the metabolic network example in this paper. This leads to $N_m=126$ and $N_p\approx 10^{38}$. The number of possible rational functions given general numerator and denominator polynomials is even larger. Simulating, evaluating, and comparing this many models in an information criteria framework is prohibitive.

A practical approach to scalable model selection builds on two highly successful machine learning techniques: libraries of candidate model nonlinearities and sparse regression. In particular, the library allows us to choose from a vast range of functional interactions in the network structure. Indeed, choosing a large library of functional forms provides the basis to select the nonlinear interactions (models) among nodes of the network. By using sparsity promoting techniques like LASSO regression [9] or elastic-net regularization [10], a

sparse subset of the components in the library are selected. The information theoretic framework selects nonlinear models that are closest in distance statistically to the empirically measured data. The SINDy method builds, via the function library, an extremely large set of potential models and selects the appropriate terms via sparse regression. This method has been recently used to successfully infer nonlinear dynamical systems [11].

C. Contribution of this work

In this work, we extend the sparse identification of nonlinear dynamical systems (SINDy) algorithm to include rational function nonlinearities in the dynamics. Nonlinear dynamics of metabolic and regulatory networks often include rational terms, motivating this innovation. It is difficult to construct a library containing rational nonlinearities for use with the sparse regression, since a generic rational function is not simply a sparse linear combination of a small set of basis functions. Instead, we write the system as an implicit differential equation in terms of the state and derivatives, and then search for the sparsest vector in the null space of all mixed state and derivative terms. For selecting the model terms in implicit-SINDy, we make use of an optimization formulation by Wright et al. [22], and an algorithm using the alternating directions method [23], to find the sparsest vector in the null space. This selects the active terms in the dynamics. We demonstrate the algorithm to be robust, accurate and fast when applied to three canonical models of biological networks: Michaelis-Menten enzyme kinetics, the regulatory network for competence in bacteria, and the metabolic network for yeast glycolysis.

In the following sections, we provide background on the existing SINDy method, and describe the updated algorithm. We next validate the algorithm on simulated data for three important biological models: the most fundamental model for enzyme kinetics, a canonical model for regulation of cell differentiation, and a seven-node metabolic network describing glycolysis. Finally we discuss the practical application of implicit-SINDy to real biological systems, including overcoming challenges such as noise and increased system size and incorporating perturbative measurements.

II. SPARSE REGRESSION FOR DYNAMICAL SYSTEMS

Our network inference method is enabled by key innovations around the broadly applicable machine learning methods of sparse regression and overcomplete libraries. When combined with dynamical systems theory, our robust mathematical architecture is achieved.

A. LASSO and sparse approximation

The sparsity promoting ℓ_1 norm is widely used in a variety of mathematical, scientific, and engineering applications. Its popularity stems from the fundamental observation that no matter the size of the data from a complex system, there often exists an underlying low-dimensional or *sparse* representation of the patterns of interest. The ℓ_1 norm is used in fields such as video processing [24], regression [9],

machine-learning [22], [25], [26], signal processing [27]–[29], dynamical systems [30]–[36].

The least absolute shrinkage and selection operator (LASSO) [9] is one of the earliest sparsity-promoting regression techniques, and it has been successful for feature selection. The LASSO solves the linear system of equations

$$\mathbf{y} = \mathbf{\Theta}\mathbf{a},\tag{3}$$

by finding a sparse solution vector a that select features from the over-determined library Θ . This gives the sparsity criteria for variable selection.

Given the success of LASSO regression, many other sparsity promoting innovations based on the ℓ_1 norm have been developed [25], [26], [37], including the commonly used elasticnet regularization [10]. Closely related to our aims, *sparse approximation* exploits the ℓ_1 norm in order to approximate a set of measurements from a linear combination of vectors from a library [38]–[41]. The vectors in a library are called atoms. The atoms in a library can come from pre-computed training data or known bases such as Fourier or wavelets. The library is deliberately constructed to be overcomplete, thus, creating an overdetermined set of linear equations. The ℓ_1 norm overcomes this challenge in order to find a sparse representation of the measurement in the library [22], [42]. Like implicit-SINDy, the sparse approximation framework combines two critical ideas: library building and sparse regression.

B. Sparse identification of nonlinear dynamics (SINDy)

Discovering dynamical systems from data is an age old pursuit in mathematical physics. Historically, this process relied on a combination of high-quality measurements and expert intuition. With growing computational power and vast quantities of data, the *automated* discovery of dynamical systems and governing equations is a relatively recent phenomenon. Broadly speaking, these techniques may be classified as *system identification*, where methods from statistics and machine learning are used to identify dynamical systems from data.

Nearly all methods of system identification involve some form of regression of data onto dynamics, and the main distinction between the various techniques is the degree to which this regression is constrained. For example, the dynamic mode decomposition (DMD) [43] generates best-fit linear models. In some cases it is possible to extend these linear regression techniques to nonlinear systems using the Koopman operator [44], which is an infinite dimensional linear operator acting on functions of observables [45]–[47]. Recent nonlinear regression techniques have produced nonlinear dynamic models that preserve physical constraints, such as conservation of energy [48]. Genetic programming has also been used to discover dynamical systems and conservation laws from data [49], [50]. These methods are highly flexible and impose very few constraints on the form of the dynamics identified.

Here we review the recent sparse identification of nonlinear dynamics (SINDy) method, which leverages advances in machine learning and sparse regression to discover nonlinear dynamical systems from data [11]. SINDy uses sparse

regression [9] for improved numerical robustness in noisy overdetermined problems, as opposed to earlier methods [30] that use compressed sensing [27]–[29].

SINDy relies on the fact that many dynamical systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{4}$$

are sparse in a given function space. The relevant terms that are active in the dynamics are solved for using an ℓ_1 -regularized regression that penalizes the number of active terms. The general framework for SINDy is shown in Fig. 1.

Algorithmically, time-series data is collected from Eq. (4), resulting in a data matrix:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(t_1) & \mathbf{x}(t_2) & \cdots & \mathbf{x}(t_m) \end{bmatrix}^T, \tag{5}$$

where T denotes the matrix transpose. The matrix \mathbf{X} is $m \times n$, where n is the dimension of the state $\mathbf{x} \in \mathbb{R}^n$ and m is the number of measurements of the state in time. For our purposes the state variables are the measured biological components in the network (enzymes, metabolites, transcription factors etc.). Similarly, the matrix of derivatives

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{\mathbf{x}}(t_1) & \dot{\mathbf{x}}(t_2) & \cdots & \dot{\mathbf{x}}(t_m) \end{bmatrix}^T, \tag{6}$$

is collected or computed from the state data in **X**; the total-variation regularized [51] derivative [52] provides numerically robust method to compute derivatives from noisy data.

Next, a library of candidate nonlinear functions is constructed from X:

$$\Theta(\mathbf{X}) = \begin{bmatrix} \mathbf{1} & \mathbf{X} & \mathbf{X}^2 & \cdots & \mathbf{X}^d & \cdots & \sin(\mathbf{X}) & \cdots \end{bmatrix}, \quad (7)$$

where \mathbf{X}^d denotes the matrix containing all possible column vectors obtained from time-series of the d-th degree polynomials in the state vector \mathbf{x} . For example, for a system with two states $\mathbf{x} = \begin{bmatrix} x_1, & x_2 \end{bmatrix}^T$, the matrix $\mathbf{X}^2 = \begin{bmatrix} x_1^2(\mathbf{t}), & (x_1x_2)(\mathbf{t}), & x_2^2(\mathbf{t}) \end{bmatrix}$, where \mathbf{t} is a vector of times at which the state is measured.

It is now possible to relate the time derivatives in $\dot{\mathbf{X}}$ to the candidate nonlinearities in $\Theta(\mathbf{X})$ by:

$$\dot{\mathbf{X}} = \mathbf{\Theta}(\mathbf{X})\mathbf{\Xi},\tag{8}$$

where each column ξ_k in Ξ is a vector of coefficients that determines which terms are active in the k-th row equation of Eq. (4). To enforce sparsity in the dynamics, we solve for each column ξ_k using sparse regression, such as the LASSO [9]:

$$\boldsymbol{\xi}_k = \operatorname{argmin}_{\boldsymbol{\xi}_k'} \|\dot{\mathbf{X}}_k - \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\xi}_k'\|_2 + \lambda \|\boldsymbol{\xi}_k'\|_1, \tag{9}$$

where $\dot{\mathbf{X}}_k$ is the k-th column of $\dot{\mathbf{X}}$. Once the sparse coefficient vectors $\boldsymbol{\xi}_k$ are determined, a model of the nonlinear dynamical system may be constructed:

$$\dot{\mathbf{x}} = \mathbf{\Xi}^T \left(\mathbf{\Theta}(\mathbf{x}^T) \right)^T. \tag{10}$$

Using sparse regression to identify active terms in the dynamics from the candidate library $\Theta(\mathbf{X})$ is a convex optimization. The alternative is to apply a separate constrained regression on every possible subset of nonlinearities, and then to choose the model that is both accurate and sparse. This

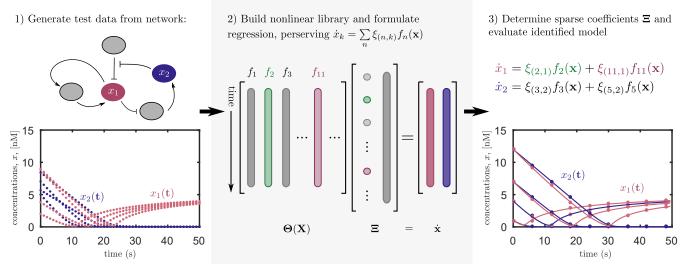


Fig. 1. Methodology for sparse identification of nonlinear dynamics (SINDy) from data. First, data is generated from a dynamical system, in this case a biological network. The time series of data is synthesized into a nonlinear function library, and the terms in this library are related to the time derivative by an overdetermined linear regression problem. Enforcing sparsity ensures that only a small number of coefficients are nonzero, identifying the few active terms in the dynamics that are needed to model the system.

brute-force search is intractable, and the SINDy method makes it possible to select the sparse model in this combinatorially large set of candidate models.

The polynomial and trigonometric nonlinearities in Eq. (7) are sufficient for a large class of dynamical systems. For example, evaluating all polynomials up to order n is equivalent to assuming that the biological network has dynamics determined by mass action kinetics up to n-mers (monomers, dimers, trimers, etc.). However, if there are time-scale separations in the mass action kinetics, fast reactions are effectively at steady state, and the remaining equations contain rational functions [53]. As we consider systems where the dynamics include rational functions, constructing a comprehensive library becomes more complicated. If we generate all rational nonlinearities:

$$f(\mathbf{x}) = \frac{f_N(\mathbf{x})}{f_D(\mathbf{x})},\tag{11}$$

where $f_N(\mathbf{x})$ and $f_D(\mathbf{x})$ are both polynomial functions, the library would be prohibitively large for computational purposes. Therefore, we develop a computationally tractable framework in the next section for functional library construction that accounts for dynamics with rational functions.

III. INFERRING NONLINEAR DYNAMICAL SYSTEMS WITH RATIONAL FUNCTIONS

Many relevant dynamical systems contain rational functions in the dynamics, motivating the need to generalize the SINDy algorithm to include more general nonlinearities than simple polynomial or trigonometric functions. The original SINDy algorithm bypasses the computation and evaluation of all N_p candidate regression models, as enumerated in Eq. (2), by performing a sparse approximation of the dynamics in a library constructed from the N_m candidate monomial features. However, it is not possible to simply apply the original SINDy procedure and include rational functions, since generic

rational nonlinearities are not sparse linear combinations of a small number of rational functions. Instead, it is necessary to modify the sparse dynamic regression problem to solve for the sparsest *implicit* ordinary differential equation according to the following procedure.

Consider a dynamical system of the form in Eq. (4), but where the dynamics of each $k = 1, 2, \dots, n$ variables may contain rational functions:

$$\dot{x}_k = \frac{f_N(\mathbf{x})}{f_D(\mathbf{x})} \tag{12}$$

where $f_N(\mathbf{x})$ and $f_D(\mathbf{x})$ represent numerator and denominator polynomials in the state variable \mathbf{x} . For each equation, it is possible to multiply both sides by the denominator polynomial, resulting in the equation:

$$f_N(\mathbf{x}) - f_D(\mathbf{x})\dot{x}_k = 0. \tag{13}$$

The implicit form of Eq. (13) motivates a generalization of the function library Θ in Eq. (7) in terms of the state \mathbf{x} and the derivative \dot{x}_k :

$$\Theta(\mathbf{X}, \dot{x}_k(\mathbf{t})) = \begin{bmatrix} \Theta_N(\mathbf{X}) & \text{diag}(\dot{x}_k(\mathbf{t})) \Theta_D(\mathbf{X}) \end{bmatrix}.$$
 (14)

The first term, $\Theta_N(\mathbf{X})$, is the library of numerator monomials in \mathbf{x} , as in Eq. (7). The second term, diag $(\dot{x}_k(\mathbf{t})) \Theta_D(\mathbf{X})$, is obtained by multiplying each column of the library of denominator polynomials $\Theta_D(\mathbf{X})$ with the vector $\dot{x}_k(\mathbf{t})$ in an element-wise fashion. For a single variable x_k , this would give the following:

$$\operatorname{diag}(\dot{x}_k(\mathbf{t}))\Theta(\mathbf{X}) = \left[\dot{x}_k(\mathbf{t}) \ (\dot{x}_k x_k)(\mathbf{t}) \ (\dot{x}_k x_k^2)(\mathbf{t}) \ \dots \right]. \tag{15}$$

A schematic of this library is shown in Fig. 2. In most cases, we will use the same polynomial degree for both the numerator and denominator library, so that $\Theta_N(\mathbf{X}) = \Theta_D(\mathbf{X})$. Thus, the augmented library in Eq. (14) is only twice the size of the original polynomial library in Eq. (7).

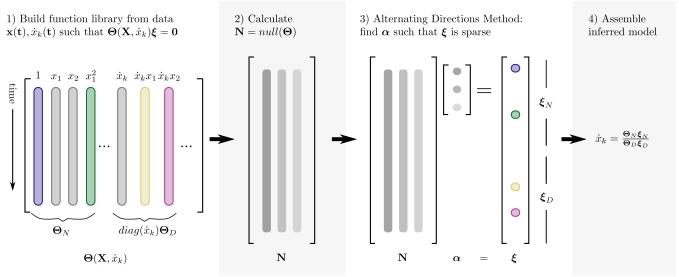


Fig. 2. SINDy algorithm for rational functions. Assemble a matrix, $\Theta(\mathbf{X}, \dot{x}_k)$ where each column is a nonlinear function evaluated for time series data x_1, x_2, x_3 .. and one species derivative $\dot{x}_k(\mathbf{t})$. Next, calculate \mathbf{N} a orthonormal basis for the null space of Θ . Then, use an alternating directions method [23] to find a sparse vector, $\boldsymbol{\xi}$, in the null space. The sparse vector $\boldsymbol{\xi}$, then satisfies $\Theta\boldsymbol{\xi}=0$. Using the sparse coefficients from $\boldsymbol{\xi}$ and the functional library $\boldsymbol{\Theta}$ assemble the inferred model. This algorithm must be performed for the derivative \dot{x}_k of each species.

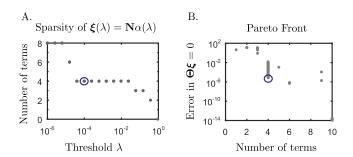


Fig. 3. A. Increasing the sparsity threshold λ during ADM creates coefficient vectors, $\boldsymbol{\xi}$, with monotonically decreasing number of terms. B. For each $\boldsymbol{\xi}(\lambda)$ we calculate an error as $|\Theta\boldsymbol{\xi}|$, and produce the Pareto Front. For the cases tested here, a large cliff in the error indicates the best choice of $\boldsymbol{\xi}(\lambda)$ (circled on A. and B.) and the most parsimonious model.

We may now write the dynamics in Eq. (13) in terms of the augmented library in Eq. (14):

$$\Theta(\mathbf{X}, \dot{x}_k(\mathbf{t}))\boldsymbol{\xi}_k = \mathbf{0}. \tag{16}$$

The sparse vector of coefficients $\boldsymbol{\xi}_k$ will have non-zero entries for the terms active in the nonlinear dynamics. However, it is not possible to use the same method of sparse regression as in the original SINDy algorithm, i.e. to find the sparsest vector $\boldsymbol{\xi}_k$ that satisfies Eq. (16), since the sparsest vector would be identically zero.

To find the sparsest non-zero vector $\boldsymbol{\xi}_k$ that satisfies Eq. (16), we note that any such vector will be in the null space of $\boldsymbol{\Theta}$. After identifying the null space of $\boldsymbol{\Theta}$, we need only find the sparsest vector in this subspace. Although this is a non-convex problem, there are straightforward algorithms based on the alternating directions method (ADM) developed by Qu *et al.* [23] to identify the sparsest vector in a subspace.

A. Algorithm for sparse selection of rational functions.

The algorithm for finding ξ_k is as follows. First, we build our functional library $\Theta(\mathbf{X}, \dot{x}_k(\mathbf{t}))$ using both the time series

data of the state variables and derivative, as discussed above. Second, we calculate a matrix, N, with columns spanning the null space of Θ . We wish to find the linear combinations of columns in N that produces a sparse vector $\boldsymbol{\xi}$. For this third step, we use the alternating directions method developed by Qu *et al.* [23] that finds the sparsest vector in a subspace. We enforce some magnitude of sparsity using a threshold, λ . For the fourth and final step, we select the active nonlinear functions using $\boldsymbol{\xi}$ and $\boldsymbol{\Theta}$, and assemble the inferred model.

As the appropriate λ is unknown *a priori*, we repeat the third and fourth steps for varying λ . Increasing λ increases the sparsity (decreasing the number of terms) in ξ , as shown in Fig. 3A. Each $\xi(\lambda)$ produces an inferred model of varying accuracy and sparsity. From these models we calculate a Pareto front and select the most parsimonious model, as shown in Fig. 3B. The most parsimonious model is readily identifiable at the sharp drop-off on the Pareto plot. As we will show, this method succeeds at identifying the correct rational terms and coefficients.

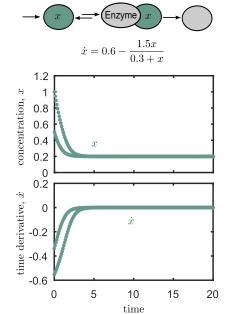
B. General formulation for implicit ODEs

The procedure above may be applied to identify more general implicit ordinary differential equations, beyond those just containing rational function nonlinearities. The library $\Theta(\mathbf{X}, \dot{x}_k(\mathbf{t}))$ contains a subset of the columns of the library $\Theta([\mathbf{X} \ \dot{\mathbf{X}}])$, which is obtained by building nonlinear functions of the state \mathbf{x} and derivative $\dot{\mathbf{x}}$.

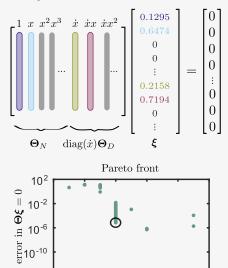
Identifying the sparsest vector in the null space of $\Theta([\mathbf{X}\ \dot{\mathbf{X}}])$ provides more flexibility in identifying nonlinear equations with mixed terms containing various powers of any combination of derivatives and states. For example, the system given by

$$\dot{x}^3 x - \dot{x}x^2 - x^3 = 0 \tag{17}$$

1) Generate test data from system:



2) Build functional library. Sparsely select terms and find λ where error drops on Pareto front:



3) Construct inferred model and compare with data from new initial conditions:

$$\dot{x} = \frac{0.1295 - 0.6474x}{0.2158 + 0.7194x} = \frac{0.6 - 3x}{1 + 3.33x}$$
$$= \frac{0.6(1 + 3.33x)}{1 + 3.33x} - \frac{1.999x + 3x}{1 + 3.33x}$$
$$= 0.6 - \frac{1.5x}{0.3 + x}$$

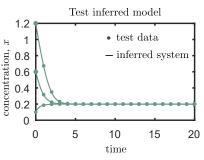


Fig. 4. Algorithm applied to the Michaelis-Menten kinetics for an enzymatic reaction. Step 1) Generate two time series of the single state variable, x(t), and time derivative, $\dot{x}(t)$. Step 2) Discovered active functions and their corresponding coefficients are indicated by color. The error drops sharply at 4 terms on the Pareto front (circled). The most parsimonious model has four active functions: two in the numerator and two in the denominator (indicated by color). Step 3) Allowing for rational function factorization, the inferred model is equivalent to the original model.

4

6

number of terms

8

10

2

may be encoded as a sparse vector in the null space of $\Theta([\mathbf{X} \ \dot{\mathbf{X}}])$. It is also straightforward to extend the formulation to include higher order derivatives, by increasing the features in the Θ library. For example, second-order implicit dynamical systems may be formulated in the following library:

$$\Theta\left(\begin{bmatrix} \mathbf{X} & \dot{\mathbf{X}} & \ddot{\mathbf{X}} \end{bmatrix}\right). \tag{18}$$

 10^{-14}

0

The generality of this approach enables the identification of many more systems of interest, in addition to those systems with rational function nonlinearities explored below.

IV. RESULTS

The implicit-SINDy architecture is tested on a number of canonical models of biological networked dynamical systems. Validation of the method on these models allows for potential broader application. We demonstrate that the method is fast, accurate and robust for inferring Michaelis-Menten enzyme kinetics, the regulatory network for competence in bacteria, and the metabolic network for yeast glycolysis.

A. Simple example: Michaelis-Menten kinetics

Perhaps the most well known model for enzyme kinetics is the Michaelis-Menten model [54], [55]. This model captures the dynamics of an enzyme binding and unbinding with a substrate (x), and then reacting irreversibly to produce a product, as shown in Fig. 4. A separation of time-scales argument, where binding and unbinding dynamics are fast, or a more general steady state assumption [56], reduces the

dynamics to a single state-variable equation with a rational function in the dynamics. Traditionally, biochemists vary the initial concentration of x in a titration experiment to fit the Michaelis-Menten equation to the data.

Using time series data from only two initial concentrations, our algorithm extracts the correct functional form from a larger subset of possible functions and fits the coefficients accurately (Fig. 4). First we generate data from the single dynamic equation

$$\dot{x} = j_x - \frac{V_{max}x}{K_m + x},\tag{19}$$

with some flux source of x, j_x , and an enzymatic reaction of the Michaelis-Menten form consuming x. Here, V_{max} is the maximum rate of the reaction and K_m is the concentration of half-maximal reaction rate. Generally the time series data of the concentration, $x(\mathbf{t})$, is measurable, while the time series data for the derivative can be calculated from $x(\mathbf{t})$.

Next, we apply implicit-SINDy to determine the coefficient vector $\boldsymbol{\xi}$ and sparsely select the active functions in the dynamics. The library contains polynomial terms up to degree four and has 10 columns. The Pareto front for this system has a sharp drop off in error from around 0.01 to 10^{-5} at four terms, indicating the λ for the most parsimonious model. The associated $\boldsymbol{\xi}$ selects 4 active terms from the function library. Finally, the rational function constructed from $\boldsymbol{\Theta}$ and $\boldsymbol{\xi}$ needs to be factored to be interpreted as the source flux and Michaelis-Menten terms. When rearranged, the coefficients match the original system. Unsurprisingly, the inferred model

1) Generate test data from system: 2) Build functional library. Sparsely select terms and find λ where error drops on Pareto front: $1 \ x_1 \ x_2 \ x_1^2 \ x_2^2 \ \ x_1^5 x_2 \quad \dot{x}_2 \ \dot{x}_2 x_1 \ \dot{x}_2 x_2 \ \dot{x}_2 x_1^5 \ \dot{x}_2 x_1^6 \ \dot{x}_2 x_2 x_1^5$ 0 $0.5\dot{0}000$ 0.00027 ġ 0.50000derivative, \dot{x} , [nM/s] concentration, x, [nM]. 0 0 : $0.50000 \\ 0.50000$ 10 Θ_N $\operatorname{diag}(\dot{x}_2)\boldsymbol{\Theta}_D$ ξ 3) Compare constructed model with data from new initial conditions: 0 Pareto front 10⁰ [n]• test data error in $\Theta \xi = 0$ 10⁻¹ concentration, x, inferred system 10 10^{-2} 5 • 10⁻³ 10 0 30 40 50 5 10 15 0 10 20 30 40 50 10 20 0 20 time, [s] number of terms time, [s]

Fig. 5. Algorithm applied to a regulatory network with two measured state-variables: x_1 (blue) and x_2 (pink). Pointed arrows indicate activation and blunted arrows indicate repression. The functional library Θ , sparse vector $\boldsymbol{\xi}$, and Pareto Front for x_2 are shown in Step 2. 10 of 56 terms are active in the library for the most parsimonious model: 4 in the numerator and 6 in the denominator. The inferred model is nearly equivalent to the original model for test data.

matches the original model for time series generated from new initial conditions that were not used in the training data.

B. Regulatory network: B. subtilis competence

Having shown that our method works for the simplest rational model relevant to biological networks, we next test it on a regulatory model with two state variables [57]. Süel et al. [57] demonstrated that a dynamic gene network enables cells to switch between multiple behaviors - in this case B. subtilis bacteria switch between taking up DNA from the environment (competence) and vegetative growth. Other regulatory networks such as the circadian clock [58], [59] and cell cycle oscillators have similar structure and dynamics. In particular, similar dynamics may drive cancer-relevant systems like the tumor suppressor p53 [60].

The dynamics of regulatory system with two states can be described by the following two non-dimensional equations:

$$\dot{x}_1 = a_1 + \frac{a_2 x_1^2}{a_3 + x_1^2} - \frac{x_1}{1 + x_1 + x_2}, \qquad (20a)$$

$$\dot{x}_2 = \frac{b_1}{1 + b_2 x_1^5} - \frac{x_2}{1 + x_1 + x_2}. \qquad (20b)$$

$$\dot{x}_2 = \frac{b_1}{1 + b_2 x_1^5} - \frac{x_2}{1 + x_1 + x_2}. (20b)$$

These two equations are a reduction of dynamical system with six states. Each rational function arises from a steady state (or time-scale separation) assumption about the regulatory processes. The second term (scaled by a_2) in Eq. (20a) represents protein x_1 , ComK, activating its own production in an autoregulatory, positive feedback loop. The first term (scaled by b_1) in Eq. (20b), describes x_1 mediated repression of x_2 , ComS, in a negative feedback loop. Both of these terms have a Hill-function form, where the power indicates the number of x_1 proteins involved cooperatively in the regulatory complex [12]. The combination of positive and negative feedback results in the network's functional capabilities. The last term in Eqs. (20a) and (20b) describes degradation of x_1 and x_2 , mediated by a third unmeasured protein, MecA.

Using this model we generate 40 time series for the regulatory system, as shown in Fig. 5. This model challenges our method in two ways. First, the method must correctly identify the dynamic dependence on two state variables. Second, the model contains polynomial functions up to the 5th degree in the numerator of one term in Eq. (20b). To include this term, the library must contain polynomials up to degree six. Even without knowing the highest polynomial power ahead of time, it is possible to use the implicit-SINDy by trying libraries of increasing polynomial degree. If the library does not have all the required terms, there will be no clear drop off in the Pareto front as there is in Fig. 5. The library with degree six polynomials in two state variables contains 56 columns, of which 10 are active in the most parsimonious model for x_2 dynamics. The constructed models for x_1 and x_2 match almost

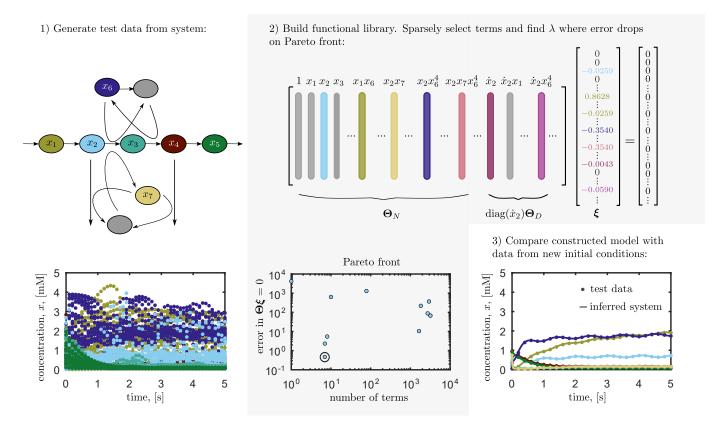


Fig. 6. Algorithm applied to a metabolic network for yeast glycolysis. Step 1). The 7 measured state variables are indicated with separate colors consistent with time series. Arrows indicate reactions between components. Not all data required to infer the network is plotted. Step 2) shows functional library Θ and corresponding coefficient vector $\boldsymbol{\xi}$ for x_2 . Seven functions are active: 5 in the numerator and 2 in the denominator. Step 3) The inferred model is nearly equivalent to the original system for test data. One time series is shown for all 7 state variables.

exactly with test data generated from the original model. As with our first example, the extracted rational function can be factored to recover exactly the form of Eq. (20b). Additionally, the coefficients identified are within 2% error of the true coefficients shown in Table I.

 $\label{table I} \textbf{TABLE I}$ Parameter identification for regulatory network

Parameter	units	True Value	Extracted value
a_1	[nM/s]	0.004	0.00393
a_2	[nM/s]	0.07	0.07006
a_3	[nM]	0.04	0.04000
b_1	[nM/s]	0.82	0.8148
b_2	[nM]	1854.5	1851.9

C. Metabolic network: yeast glycolysis

As a final example, we test our method on a metabolic network. Glycolysis, the process of breaking down glucose to extract energy (ATP and NADPH), is part of central metabolism for all cells. Uncovering the metabolic network for glycolysis took over 100 years from its initial discovery by Pasteur [61]. Accelerated inference of metabolic networks would aid metabolic disease intervention [62]. Bacteria perform a wide range of yet-to-be discovered chemistry which could be harnessed through metabolic engineering to produce high-value products such as drugs and biofuels [63] [64].

Not only does the yeast glycolysis model we analyze have a larger number of interacting state variables, but it is oscillatory [65]. This network has also been previously analyzed as a test case for model inference [66]. The network shown in Fig. 6 has three equations with rational functions and four with polynomials:

$$\dot{x}_1 = c_1 + \frac{c_2 x_1 x_6}{1 + c_3 x_6^4} \tag{21a}$$

$$\dot{x}_2 = \frac{d_1 x_1 x_6}{1 + d_2 x_6^4} + d_3 x_2 - d_4 x_2 x_7 \tag{21b}$$

$$\dot{x}_3 = e_1 x_2 + e_2 x_3 + e_3 x_2 x_7 + e_4 x_3 x_6$$
 (21c)

$$\dot{x}_4 = f_1x_3 + f_2x_4 + f_3x_5 + f_4x_3x_6 + f_5x_4x_7$$
 (21d)

$$\dot{x}_5 = g_1 x_4 + g_2 x_5 \tag{21e}$$

$$\dot{x}_6 = \frac{h_1 x_1 x_6}{1 + h_2 x_6^4} + h_3 x_3 + h_5 x_6 + h_4 x_3 x_7 \tag{21f}$$

$$\dot{x}_7 = j_1 x_2 + j_2 x_2 x_7 + j_3 x_4 x_7. \tag{21g}$$

Given sufficient data, implicit-SINDy correctly infers the network structure and coefficients. In Fig. 6, we show the sparsely selected terms and Pareto front for Eq. (21c). The method correctly selects 7 terms from a library of 3432 functions: 5 in the numerator and 2 in the denominator. Table II shows the true and extracted coefficient values for the model. Some of the parameters, c_1 for example, had incorrect functional

TABLE II
PARAMETER IDENTIFICATION FOR METABOLIC NETWORK

Parameter	units	True Value	Extracted value
c_1	[mM/min]	2.5	2.5002*
c_2	[1/(mM min)]	-100	-99.7979
c_3	$[mM^{-1/4}]$	13.6769	13.6489
d_1	[1/(mM min)]	200	200.6512
d_2	$[mM^{-1/4}]$	13.6769	13.7209
d_3	[1/min]	-6	-5.9998*
d_4	[1/(mM min)]	-6	-5.9998*
e_1	[1/min]	6	6.0133
e_2	[1/min]	-64	-64.140
e_3	[1/(mM min)]	6	-6.0133
e_4	[1/(mM min)]	16	16.0333
f_1	[1/min]	64	63.6145
f_2	[1/min]	-13	-12.9277
f_3	[1/min]	13	12.9277
f_4	[1/(mM min)]	-16	-15.9036
f_5	[1/(mM min)]	-100	-99.3976
g_1	[1/min]	1.3	1.3002
g_2	[1/min]	-3.1	-3.1003
h_1	[1/(mM min)]	-200	-200.0^{\dagger}
h_2	$[mM^{-1/4}]$	13.6769	13.6769 [†]
h_3	[1/min]	128	128.0 [†]
h_4	[1/(mM min)]	-1.28	-1.280^{\dagger}
h_5	[1/min]	-32	-32.00^{\dagger}
j_1	[1/min]	6	6.0102
j_2	[1/(mM min)]	-18	-18.0408
j_3	[1/(mM min)]	-100	-100.2449

Values with a * have errors with functional dependence on x_6 . We show the leading order term assuming a Taylor series expansion. † Eq. (21g) required more data to identify, and this allowed extraction of the coefficients to a higher precision.

dependence on x_6 after factoring the discovered polynomial. However these dependencies were very small (< 0.1% error).

The equations for x_3 , x_4 , x_5 , and x_7 did not require a library with polynomials up to degree six, and could be inferred more quickly. On the other hand, Eq. (21g), required over twice as many measurements as the other equations with rational functions (Eqs. (21c) and (21b)).

V. CONCLUSIONS

In this work we developed an implicitly formulated method for sparse identification of nonlinear dynamics: implicit-SINDy. The method allows for constructing nonlinear dynamics with rational functions using a library of functional forms that is still computationally manageable for reasonably-sized biological networks. An alternating directions method for selection of a sparse vector in the null space of the library [23] enables us to find and construct a parsimonious model from the full library. Using implicit-SINDy on data generated from three biological models (enzyme kinetics, regulation, and metabolism), we are able to accurately reconstruct the underlying system in each case. Indeed, we correctly recover the coefficients to within 2% of the original values. These results make implicit-SINDy a promising method for model discovery of biological networks.

SINDy is a *data-driven* methodology, meaning it selects the connectivity and dynamics based on the information content of the data alone. It has many advantages, including the fact

that there is no parameter tuning in the inferred models aside from a sparsity threshold which is determined by a Pareto front. Moreover, implicit-SINDy greatly expands our ability to rapidly select a model from a large class of candidate dynamical systems, even when nonlinear derivative terms are present. In practice, the method functions much like a highly efficient unsupervised learning algorithm, sparsely selecting dynamics from a large library of possibilities. It differs from information theoretic techniques where a number of viable models are posited and selection of the best model is based upon the minimization of information loss. Such alternative techniques rely on physical insight (supervised learning) to generate individual models, thus potentially limiting the dynamical systems considered. Given the large number of biological models driven by mass-action kinetics, the implicit-SINDy method can be a critically enabling method for datadriven discovery of underlying biological principals.

There are many intriguing future directions for the method, both in theory and practice. In theory, the connection between the implict-SINDy selection process and information criteria such as AIC and BIC remains an open question. We hope to pursue this further in order to establish a rigorous statistical connection between information theory metrics and sparse selection. In practice, two remaining challenges exist for practical implementation: (i) improving robustness to noise and (ii) reducing the number of time-series measurements. As one would expect, noise compromises the calculation of the null space of a matrix. Recent work by Gavish and Donoho provided a general method for recovering a low-rank matrix (and therefore null space) from noisy data [67]. Such a thresholding technique may be used to make implicit-SINDy more robust to noise. It remains an open question how long a time-series must be sampled in order to accurately produce the underlying dynamics. This will be also considered in future work.

ACKNOWLEDGMENT

SLB acknowledges support from the U.S. Air Force Center of Excellence on Nature Inspired Flight Technologies and Ideas (FA9550-14-1-0398). JLP and NMM would like to thank Bill and Melinda Gates for their active support of the Institute for Disease Modeling and their sponsorship through the Global Good Fund.

REFERENCES

- S. Kullback and R. A. Leibler, "On Information and Sufficiency," The Annals of Mathematical Statistics, vol. 22, no. 1, pp. 79–86, mar 1951.
- [2] S. Kullback, Information Theory and Statistics, J. W. &. Sons, Ed., 1959.
- [3] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Petrov, B.N.; Csáki, F., 2nd International Symposium* on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, 1973, pp. 267–281.
- [4] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [5] K. Burnham and D. Anderson, Model Selection and Multi-Model Inference, 2nd ed. Springer, 2002.
- [6] C. M. Bishop and Others, Pattern recognition and machine learning. Springer New York, 2006, vol. 1.
- [7] K. P. Murphy, Machine learning: a probabilistic perspective. MIT press, 2012.

- [8] G. James, D. Witten, T. Hastie, and R. Tibshirani, An introduction to statistical learning. Springer, 2013.
- [9] R. Tibshirani, "Regression Shrinkage and Selection via the LASSO," J. of Roy. Statistical Soc., vol. 58, no. 1, pp. 267–288, 1996.
- [10] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," pp. 301–320, 2005.
- [11] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data: Sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, 2016.
- [12] U. Alon, An Introduction to Systems Biology: Design Principles of Biological Circuits, 2007.
- [13] U. Alon, "Network motifs: theory and experimental approaches." Nature reviews. Genetics, vol. 8, no. 6, pp. 450–61, 2007.
- [14] M. G. Grigorov, "Global properties of biological networks," *Drug Discovery Today*, vol. 10, no. 5, pp. 365–372, 2005.
- [15] M. Andrecut and S. A. Kauffman, "On the sparse reconstruction of gene networks," *Journal of Computational Biology*, vol. 15, no. 1, pp. 21–30, 2008.
- [16] M. Öksüz, H. Sadikolu, and T. Çakir, "Sparsity as cellular objective to infer directed metabolic networks from steady-state metabolome data: A theoretical analysis," *PLoS ONE*, vol. 8, no. 12, pp. 1–7, 2013.
- [17] T. E. M. Nordling and E. W. Jacobsen, "On Sparsity As a Criterion in Reconstructing Biochemical Networks," *Proceedings of the 18th International Federation of Automatic Control (IFAC) World Congress*, 2011, no. 2007, pp. 11672–11678, 2011.
- [18] D. G. Spiller, C. D. Wood, D. A. Rand, and M. R. H. White, "Measurement of single-cell dynamics," *Nature*, vol. 465, no. June, pp. 736–745, 2010.
- [19] M. Wu and A. K. Singh, "Single-cell protein analysis," Current Opinion in Biotechnology, vol. 23, no. 1, pp. 83–88, 2012.
- [20] G. Schwarz, "Estimating the dimension of a model," The Annals of Statistics, no. 6, pp. 461–464, 1978.
- [21] G. Claeskens and N. L. Hjorth, Model Selection and Model Averaging. Cambridge University Press, 2008.
- [22] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
- [23] Q. Qu, J. Sun, and J. Wright, "Finding a sparse vector in a subspace: Linear sparsity using alternating directions," in *Advances in Neural Information Processing Systems* 27, 2014, pp. 3401—3409.
- [24] J. V. Shi, W. Yin, A. C. Sankaranarayanan, and R. G. Baraniuk, "Video compressive sensing for dynamic MRI," submitted for publication, 2013.
- [25] R. Tibshirani, T. Hastie, B. Narasimhan, and G. Chu, "Diagnosis of multiple cancer types by shrunken centroids of gene expression," *Proceedings of the National Academy of the Sciences*, vol. 99, no. 10, pp. 6567–6572, 2002.
- [26] H. Zou, T. Hastie, and R. Tibshirani, "Sparse Principal Component Analysis," J. Comp. and Graphical Statistics, vol. 15, no. 2, pp. 262– 286, 2006.
- [27] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [28] E. J. Candès, "Compressive Sensing," Proceedings of the International Congress of Mathematics, 2006.
- [29] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–120, 2007.
- [30] W. X. Wang, R. Yang, Y. C. Lai, V. Kovanis, and C. Grebogi, "Predicting catastrophes in nonlinear dynamical systems by compressive sensing," *Physical Review Letters*, vol. 106, no. 15, pp. 1–4, 2011.
- [31] V. Ozolicnš, R. Lai, R. Caflisch, and S. Osher, "Compressed modes for variational problems in mathematics and physics," *Proceedings of* the National Academy of Sciences, vol. 110, no. 46, pp. 18368–18373, 2013.
- [32] H. Schaeffer, R. Caflisch, C. D. Hauck, and S. Osher, "Sparse dynamics for partial differential equations," *Proceedings of the National Academy* of Sciences USA, vol. 110, no. 17, pp. 6634–6639, 2013.
- [33] I. Bright, G. Lin, and J. N. Kutz, "Compressive sensing and machine learning strategies for characterizing the flow around a cylinder with limited pressure measurements," *Physics of Fluids*, vol. 25, pp. 127 102– 127 115, 2013.
- [34] S. L. Brunton, J. H. Tu, I. Bright, and J. N. Kutz, "Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems," SIAM Journal on Applied Dynamical Systems, vol. 13, no. 4, pp. 1716–1732, 2014.

- [35] A. Mackey, H. Schaeffer, and S. Osher, "On the Compressive Spectral Method," Multiscale Modeling & Simulation, vol. 12, no. 4, pp. 1800– 1827, 2014
- [36] S. L. Brunton and B. R. Noack, "Closed-loop turbulence control: Progress and challenges," *Applied Mechanics Reviews*, vol. 67, no. 050801, pp. 1–48, 2015.
- [37] C. Line, T. Hastie, D. Witten, and B. Ersbøll, "Sparse discriminant analysis," *Technometrics*, vol. 53, no. 4, pp. 406–413, 2011.
- [38] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Info. Theory*, vol. 47, no. 7, pp. 2845– 2862, 2001.
- [39] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via 1¹ minimization," *Proc. Natl. Acad. Sci.*, vol. 5, no. 100, pp. 2197–2202, 2003.
- [40] A. Tropp, A. C. Gilbert, S. Muthukrishnan, and M. J. Strauss, "Improved sparse approximation over quasi-incoherent dictionaries,," *IEEE International conference on image processing ICIP*, pp. 37–40, 2003.
- [41] A. C. Gilbert, S. Muthukrishnan, and M. J. Strauss, "Approximation of Functions over Redundant Dictionaries Using Coherence," Proc. 2003 SIAM Symposium on Discrete Algorithms SODA, pp. 243–252, 2003.
- [42] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.
- [43] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Characterization of Complex Systems*. SIAM, 2016.
- [44] B. O. Koopman, "Hamiltonian Systems and Transformation in Hilbert Space," *Proceedings of the National Academy of Sciences*, vol. 17, no. 5, pp. 315–318, 1931.
- [45] I. Mezic, "Analysis of fluid flows via spectral properties of the Koopman operator," Annual Review of Fluid Mechanics, vol. 45, pp. 357–378, 2013.
- [46] M. Budišić, R. Mohr, and I. Mezić, "Applied Koopmanism a)," Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 22, no. 4, p. 47510, 2012.
- [47] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics*, vol. 41, no. 1-3, pp. 309– 325, 2005.
- [48] A. J. Majda and J. Harlim, "Physics constrained nonlinear regression models for time series," *Nonlinearity*, vol. 26, no. 1, p. 201, 2012.
- [49] J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 9943–9948, 2007.
- [50] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science*, vol. 324, no. 5923, pp. 81–85, 2009.
- [51] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.
- [52] R. Chartrand, "Numerical differentiation of noisy, nonsmooth data," ISRN Applied Mathematics, vol. 2011, 2011.
- [53] J. Gunawardena, "Time-scale separation Michaelis and Menten's old idea, still bearing fruit," FEBS Journal, vol. 281, no. 2, pp. 473–488, 2014.
- [54] Michaelis L & Menten M, "Die kinetik der Invertinwirkung." Biochem Z, vol. 49, pp. 333–369, 1913.
- [55] K. A. Johnson and R. S. Goody, "The original Michaelis constant: Translation of the 1913 Michaelis-Menten Paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011.
- [56] G. E. Briggs, "A Further Note on the Kinetics of Enzyme Action." The Biochemical journal, vol. 19, no. 6, pp. 1037–1038, 1925.
- [57] G. M. Süel, J. Garcia-Ojalvo, L. M. Liberman, and M. B. Elowitz, "An excitable gene regulatory circuit induces transient cellular differentiation." *Nature*, vol. 440, no. 7083, pp. 545–550, 2006.
- [58] P. E. Hardin, "The circadian timekeeping system of Drosophila." Current biology: CB, vol. 15, no. 17, pp. R714–22, 2005.
- [59] J.-C. Leloup and A. Goldbeter, "Toward a detailed computational model for the mammalian circadian clock." *Proceedings of the National Academy of Sciences*, vol. 100, no. 12, pp. 7051–6, 2003.
- [60] E. Batchelor, A. Loewer, and G. Lahav, "The ups and downs of p53: understanding protein dynamics in single cells." *Nature reviews. Cancer*, vol. 9, no. 5, pp. 371–377, 2009.
- [61] E. Racker, "From Pasteur to Mitchell: a hundred years of bioenergetics." Federation proceedings, vol. 39, no. 2, pp. 210–5, feb 1980.

- [62] M. L. Yarmush and S. Banta, "Metabolic engineering: advances in modeling and intervention in health and disease." *Annual review of biomedical engineering*, vol. 5, pp. 349–381, 2003.
- [63] D. I. Ellis and R. Goodacre, "Metabolomics-assisted synthetic biology," Current Opinion in Biotechnology, vol. 23, no. 1, pp. 22–28, 2012.
- [64] D. C. Ducat, J. C. Way, and P. A. Silver, "Engineering cyanobacteria to generate high-value products." *Trends in biotechnology*, vol. 29, no. 2, pp. 95–103, feb 2011.
- [65] J. Wolf and R. Heinrich, "Effect of cellular interaction on glycolytic oscillations in yeast: a theoretical investigation." *The Biochemical journal*, vol. 345 Pt 2, pp. 321–334, 2000.
- [66] M. D. Schmidt, R. R. Vallabhajosyula, J. W. Jenkins, J. E. Hood, A. S. Soni, J. P. Wikswo, and H. Lipson, "Automated refinement and inference of analytical models for metabolic networks." *Physical biology*, vol. 8, no. 5, p. 055011, 2011.
- [67] M. Gavish and D. L. Donoho, "Optimal Shrinkage of Singular Values," 2014