

Machine Discovery of Partial Differential Equations from Spatiotemporal Data

Ye Yuan

School of Artificial Intelligence and Automation, State Key Laboratory of Digital Manufacturing Equipments and Technology, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

Junlin Li

School of Artificial Intelligence and Automation, Key Laboratory of Image Processing and Intelligent Control, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

Liang Li

School of Artificial Intelligence and Automation, Key Laboratory of Image Processing and Intelligent Control, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

Frank Jiang

School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm 11428, Sweden

Xiuchuan Tang

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

Fumin Zhang

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, 30332, USA

Sheng Liu

School of Power and Mechanical Engineering, Wuhan University, Wuhan 430074, P.R. China

Jorge Goncalves

Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, UK and the Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux, L-4367, Luxembourg

Henning U.Voss

Citigroup Biomedical Imaging Center, Weill Cornell Medical College, New York, 10021, USA

Xiuting Li

XTINGLI@HUST.EDU.CN

School of Artificial Intelligence and Automation, Key Laboratory of Image Processing and Intelligent Control, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

Jürgen Kurths

KURTHS@PIK-POTSDAM.DE

Research Domain IV - Transdisciplinary Concepts & Methods, Potsdam Institute for Climate Impact Research, Potsdam, 14473, Germany

Han Ding

School of Mechanical Science and Engineering, State Key Laboratory of Digital Manufacturing Equipments and Technology, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

Abstract

The study presents a general framework for discovering underlying Partial Differential Equations (PDEs) using measured spatiotemporal data. The method, called Sparse Spatiotemporal System Discovery (S^3d), decides which physical terms are necessary and which can be removed (because they are physically negligible in the sense that they do not affect the dynamics too much) from a pool of candidate functions. The method is built on the recent development of Sparse Bayesian Learning; which enforces the sparsity in the to-be-identified PDEs, and therefore can balance the model complexity and fitting error with theoretical guarantees. Without leveraging prior knowledge or assumptions in the discovery process, we use an automated approach to discover ten types of PDEs, including the famous Navier-Stokes and sine-Gordon equations, from simulation data alone. Moreover, we demonstrate our data-driven discovery process with the Complex Ginzburg-Landau Equation (CGLE) using data measured from a traveling-wave convection experiment. Our machine discovery approach presents solutions that has the potential to inspire, support and assist physicists for the establishment of physical laws from measured spatiotemporal data, especially in notorious fields that are often too complex to allow a straightforward establishment of physical law, such as biophysics, fluid dynamics, neuroscience or nonlinear optics.

Keywords: Data-driven Method, Partial Differential Equation, Machine Learning Application.

1. Introduction

The typical process of discovering physical laws involves collecting measurement data on physical phenomena and then using them to reconstruct and validate the synthesis of related, previously discovered physical laws. This foundational discovery approach is fueled by and leverages the ingenuity of physicists researching the phenomena. The physics community's careful and rigorous proposal and validation of physical law has successfully led to great achievements and progress towards understanding the physics of the universe (Aranson and Kramer, 2002; Korteweg and De Vries, 1895; Schrödinger, 1926; FitzHugh, 1961). However, this process typically needs a rather long time. This work proposes a discovery tool that may help accelerate this discovery process, especially in the presence of complex dynamics that would otherwise take needlessly significant effort to begin uncovering and understanding (Aguirre and Sanjuán, 2002; Aguirre et al., 2009; Grebogi et al., 1982; Lai and Tél, 2011; Aihara et al., 1990). The tool automatically synthesizes the available physical features which dominate complex dynamics from data, and incorporate them into model equations which enable physicists to study them both analytically and numerically.

Data-driven discovery of physical laws allows computational methods to learn representations from data. There have been reports of many successes using data-driven methods such as learning a nonlinear mapping from time series and predict chaotic dynamics (Farmer and Sidorowich, 1987), predicting catastrophes in nonlinear systems (Wang et al., 2011), equation-free modeling multi-scale/complex systems (Kevrekidis et al., 2003), inferring models of nonlinear coupled dynamical systems (Bongard and Lipson, 2007), reconstructing networks with stochastic dynamical processes (Shen et al., 2014), extracting Koopman modes from data (Mezić, 2013; Williams et al., 2015; Tu et al., 2014; Giannakis, 2017). In addition to these studies, other methods, such as symbolic regression (Schmidt and Lipson, 2009), neural-network based tools (Raissi and Karniadakis, 2017; Raissi, 2018;

Raissi et al., 2017, 2018), make an effective attempt to data-driven discovery, though they have good prediction power, yet the learned models are hard to offer physical insights. Thus, it is desirable to develop a general framework that specifically discovers a reasonably interpretable model directly from data. In the system identification community, a standard way is to transform the model discovery problem into a regression problem and design a good optimization algorithm to solve the formulated problem (Ljung, 1986; Lindquist and Picci, 2015). More precisely, suppose the dynamics of the to-be-discovered system can be described by a partial differential equation (PDE) model of the general form:

$$u_t(x, t) = F(x, t, u, u_x, u_{xx}, u_{xxx}, \dots; \Theta), \quad (1)$$

where $u(x, t)$ is a state variable with respect to time t and space $x \in \mathbb{R}$, Θ is a parameter vector and F is an unknown function, which depends on u and its partial derivatives. To infer an accurate and interpretable representation of F using measurement data of the state variable u , a (large) set of dictionary with many potential candidate terms; for instance, $\{u, u_x, u_{xx}, u^2, uu_x, uu_{xx}, u_x u_{xx}, \dots\} \triangleq \{\Phi_1, \dots, \Phi_m\}$ (with cardinality m), will be created to represent F via $F = \sum_{i=1}^m \theta_i \Phi_i$ and the values of the parameters/weights $\{\theta_i\}_{i=1}^m$ will be estimated by solving the following regression problem

$$u_t = \Phi \theta + \nu, \quad (2)$$

where $\Phi \in \mathbb{R}^{n \times m}$ is the dictionary matrix, $u_t \in \mathbb{R}^n$ is the time derivatives of u , $\theta \in \mathbb{R}^m$ is the unknown vector to be learned and $\nu \in \mathbb{R}^n$ is the modelling error.

There are several numerical algorithms to find solutions to problems like Eq. (2). The first kind of such methods builds on the ordinary least squares regression. Examples include orthogonal reduction (Xu and Khanmohamadi, 2008; Khanmohamadi and Xu, 2009), the backward elimination scheme (Bär et al., 1999; Guo and Billings, 2006), least squares estimator (Müller and Timmer, 2004) and the alternating conditional expectation algorithm (Breiman and Friedman, 1985; Voss et al., 1999, 1998). These approaches do have an good fitting accuracy as described by (Tibshirani et al., 2015) which, however, have two major issues: prediction accuracy and interpretability in modeling a wider range of data. The second type of methods is the sparse regression approach, where a sparse priori assumption of underlying governing equations is incorporated to tackle the outstanding issue, overfitting. The work of (Brunton et al., 2016) apply a L_0 -norm penalty into the estimation problem of Eq. (2) that reflects sparse priori. The resulting sparse identification of nonlinear dynamics (SINDy) algorithm selects the key feature terms based on the value of least squares regression coefficients by a thresholding step (Mangan et al., 2016; Zhang and Schaeffer, 2018). This SINDy framework is recently extended to recover PDEs from spatiotemporal data (Rudy et al., 2017) where not only time/spatial derivatives are estimated from noisy data but also a modified algorithm, named by a Sequential Threshold Ridge regression (STRidge) algorithm, is developed to improve the robustness of the discovery. Instead of penalizing the L_0 norm of θ , (Schaeffer, 2017) used lasso algorithm, a method that combines the least-squares loss with an L_1 -constraint regularization of the coefficients, to learn active features and estimate coefficients in the underlying PDEs. In both L_0 and L_1 schemes, the identified PDEs model is parsimonious and accurate, as the penalty term explicitly encourages the solution with few nonzero components. However, it is well known that solving the

L_0 problem is NP-hard and hardly tractable with the number of features large due to the nonconvexity of the L_0 norm. The L_1 norm can provide a good approximation (convex relaxation) to the L_0 norm, but the resulting Lasso tends to over-shrink large coefficients, which leads to biased estimates (Fan and Li, 2001).

In this paper, the estimation problem in Eq. (2) with observations generated by spatiotemporal systems is solved in a Sparse Bayesian Learning (SBL) framework. Proposed in (Tipping, 2001) and further investigated in (Faul and Tipping, 2002; Rao et al., 2003; Palmer et al., 2006), the SBL framework treats the parameter θ as a random variable with sparse-inducing priori distributions determined by a set of hyperparameters. It has been shown to be more general and powerful in finding maximally sparse solution. For instance, the maximum a posteriori (MAP) estimation with a fixed weight priori is its special case (Rao et al., 2003); the Lasso problem can be interpreted as the explicit MAP with Laplace priori (Figueiredo, 2002); the L_0 -norm can be approached by a priori of the form: $p(\theta) \sim \exp(-\sum_{i=1}^m |\theta_i|^p)$ with $p \rightarrow 0$ (Wipf and Rao, 2004). Specially, (Wipf et al., 2011) establishes the equivalence of the SBL cost function and the objectives in canonical MAP-based sparse methods (e.g., the Lasso) via a dual-space analysis and shows that the SBL method maintains many desirable advantages over all possible MAP methods, such as using non-factorial coefficient priors (Wipf and Nagarajan, 2008), treating the reconstruction problem with unfavorable restricted isometry properties (RIP) (Wu and Wipf, 2012), implementing non-separable alternative via either L_2 or L_1 reweighting (Wipf and Nagarajan, 2010; Pan et al., 2012; Chang and Tomlin, 2012; Pan et al., 2016, 2018). Recently, (Zhang and Lin, 2018) explores the advantage of Bayesian inference to provide error bars and robust discovery of governing physical laws from noisy data. However, this relatively new paradigm is still hindered by lack of the theoretical analysis and an effective optimization algorithm which addresses issues brought by the SBL cost function.

Our main purpose in this paper is to present an effective SBL framework for locally recovering unknown PDEs from spatiotemporal data and conduct some theoretical analysis on the sparse discovery for spatiotemporal systems. For brevity, we refer to the proposed discovery approach as the Sparse Spatiotemporal System Discovery (S^3d) method. The key contributions of the S^3d framework include the following. First of all, we take a new perspective of the SBL algorithm. Specially, the existing methods optimizing the SBL cost function, such as the Expectation-Maximization (Wipf and Rao, 2004) and MacKay update rules (Tipping, 2001), are expressed in either the coefficient space or the latent space. Here we take the cost function expressed concurrently in the two space. The resulting difference of convex function (DC) programming is then solved by an iterative convex weighted lasso algorithm with concave-convex procedure (CCP). Secondly, many prototypical PDEs including the Navier-Stokes equation in fluid mechanics (Shu and Richards, 1992), the FitzHugh-Nagumo for nerve conduction (FitzHugh, 1961), and the Schrödinger equation in quantum mechanics (Schrödinger, 1926) are rediscovered from only spatiotemporal data, which cast off a great deal of trial and error. Finally, we apply the proposed S^3d method to real measurements from a traveling-wave convection experiment with the discovery of the normal Complex Ginzburg-Landau Equation (CGLE) model. The numerical solution to the CGLE commendably reconstructs the observed experimental phenomenon.

Theoretically, the primary concern is the convergence and consistency of the sparse estimators with measurement data generated by spatiotemporal systems. The convergence and consistency analysis are conducted under the SBL cost function expressed jointly in the coefficient space and the latent variable space. We show that under reasonable assumptions on the optimization path in joint space, with the CCP, the SBL cost function has a local minimum and can produce the maximally sparse solution. When the cost function is separable, we show that the DC programming formulation succeeds in recovering sparsity pattern and reaching small error with high probability.

The rest of this paper is organized as follows: In the next subsection we summarize the notation and conventions used in this paper. In Section 2, we introduce a bayesian framework for the problem of discovering PDEs from spatiotemporal data and estimation of coefficients. Our main results on the convergence and consistency of the parameter estimator are presented in Section 3, with the proof included in the Appendix. We then evaluate our method experimentally in Section 4 with both synthetic data and real-life data, which confirms the exact sparse recovery of the theoretical results. The comparison with the state-of-the-art algorithms are also provided in this section. We conclude this paper with a discussion in Section 5.

1.1 Notation

Throughout this paper, we denote vectors by lowercase letters, e.g., x , and matrices by uppercase letters, e.g., X . For a vector x and a matrix X , we denote its transpose by x^T and X^T , respectively. Furthermore, x_i refers to its i -th element of x . $\|x\|_2 = \sqrt{x^T x}$ and $\|x\|_1 = \sum_\ell |x_\ell|$, denotes its Euclidean norm and ℓ_1 -norm. Let $u : \Omega \rightarrow \mathbb{R}, x \in \Omega \subset \mathbb{R}^n$, $\frac{\partial u}{\partial x_i} = \lim_{h \rightarrow 0} \frac{u(x_i + h) - u(x_i)}{h}$, provide this limit exists. We usually write u_t for $\partial u / \partial t$, and similarly $u_{x_i} \triangleq \partial u / \partial x_i$, $u_{x_i x_i} \triangleq \partial^2 u / \partial x_i^2$, $u_{x_i x_j} \triangleq \partial^2 u / \partial x_i \partial x_j$, etc..

2. S³d: A Bayesian Framework to Discover PDEs from Spatiotemporal Data

This section proposes a framework of the S³d method for data-driven discovery of PDEs. The proposed S³d framework includes the following three steps: Data collection, Model specification and SBL. We will consider the dynamics of a spatiotemporal system which is governed by a PDE of the general form:

$$\frac{\partial u(x, t)}{\partial t} = F \left(x, t, u(x, t), \frac{\partial u(x, t)}{\partial x_1}, \dots, \frac{\partial u(x, t)}{\partial x_n}, \frac{\partial^2 u(x, t)}{\partial x_1^2}, \dots, \frac{\partial^2 u(x, t)}{\partial x_1 \partial x_n}, \dots, \Theta \right), \quad (3)$$

where the dynamical variable, u , is N -component with temporal and multi-dimensional spatial variables, t and $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, respectively, defining the state. Θ is a parameter vector and F is an unknown, N -dimensional function, which depends on the dynamical variable u and its derivatives.

2.1 Model Specification

In this section, we shall restrict the discussion to the case $N = 1$ for notational simplicity, yet the proposed framework can be applied to general cases when $N > 1$. We construct a

large library of candidate terms that may appear in the function F . We mainly take the model structure to be a q^{th} -order PDE:

$$u_t = \sum_{i=1}^m \theta_i f_i \left(u^{p_1} \left(\frac{\partial u}{\partial x} \right)^{p_2} \cdots \left(\frac{\partial^q u}{\partial x^q} \right)^{p_{n+1}} \right),$$

where $q \in \mathbb{N}^+$ represents the n^{th} -order PDE, $p = \sum_{l=1}^{n+1} p_l$ represents the p -th operator with $p_l \in \mathbb{Z}^+$, where $l = 1, 2, \dots, n + 1$ is the degree of nonlinearity, $f_i(\cdot)$ ($i = 1, 2, \dots, m$) is an identical mapping and the term u_t represents the output. The chosen model structure can pack many nonlinearities (e.g., polynomial u^2 , uu^2 , etc.), partial derivatives (e.g., u_x , u_{xx} , u_{xxx} , etc.), their combinations (e.g., $u^2 u_x$, $u^3 u_{xx}$, etc.). Some physics-based candidate terms (such as the reaction term $u_x - u_{xxx}$, the diffusion term $u(u - u_{xx})_x$, the convection term $u_x(u - u_{xx})$, etc.) also are included in our function library if needed. The PDE with the dictionary can be rewritten into the following regression form:

$$u_t = [1 \ u \ u_x \ u_{xx} \ u_{xxx} \ u^2 \ uu_x \ \cdots \ u^2 u_{xxx}] \cdot \theta, \quad (4)$$

where we chose the order $q = 3$ and the degree $p = 3$ for clarity. Eq. (4) then holds at all points in the domain of interest.

Now, we construct a feature matrix Φ . Each column of Φ is a compilation of all the values of a specific candidate function on the right side of Eq. (4) over all grid points, and the output vector, y , contains the values of u_t in the left side of Eq. (4) over all grid points. We state the set of linear equations representing our PDE as follows:

$$\underbrace{\begin{bmatrix} u_t(x_1, t_1) \\ u_t(x_2, t_1) \\ \vdots \\ u_t(x_{n_x}, t_{n_t}) \end{bmatrix}}_y = \underbrace{\begin{bmatrix} 1 & u(x_1, t_1) & u_x(x_1, t_1) & \cdots & u^2 u_{xxx}(x_1, t_1) \\ 1 & u(x_2, t_1) & u_x(x_2, t_1) & \cdots & u^2 u_{xxx}(x_2, t_1) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & u(x_{n_x}, t_{n_t}) & u_x(x_{n_x}, t_{n_t}) & \cdots & u^2 u_{xxx}(x_{n_x}, t_{n_t}) \end{bmatrix}}_{\text{Dictionary matrix } \Phi} \underbrace{\begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta \end{bmatrix}}_\theta. \quad (5)$$

To find the unknown coefficients θ , we need to compute the time (and spatial) derivatives of u and solve Eq. (5). Various numerical methods including the numerical differentiation (LeVeque, 2007; Li and Chen, 2008; Trefethen, 2000), the symbolic or automatic differentiation (Baydin et al., 2018) can be used for the estimation of the derivatives in the dictionary matrix Φ . In our framework, the polynomial approximation method is selected to estimate the derivatives from the noise-contained measurements. Indeed, the polynomial approximation method has been proved to be capable of effectively estimating the derivatives from the noise-contained data in (Rudy et al., 2017). In Appendix A, we explain the polynomial approximation method in detail.

A general solution to Eq. (5) can be obtained by using the least-squares method. For example, $\hat{\theta}$ is the minimizer of the following optimization problem:

$$\hat{\theta} = \arg \min_{\theta} \|y - \Phi\theta\|_2^2. \quad (6)$$

The solution, $\hat{\theta}$, is uniquely identified only if the number of grid points, $n = n_x n_t$, is larger than the number of columns in Φ and the matrix, Φ , has full column rank. However,

these two conditions, in general, are not always satisfied. In addition, if the dictionary is over-determined, the resulting optimization easily leads to overfitting.

2.2 Sparse Bayesian Learning for S³d

This section proposes a SBL algorithm based CCP to solve Eq. (5) defined by S³d. The SBL, initially introduced to solve regression and classification problems in seminal works ([Tipping, 2001](#); [Wipf and Rao, 2004](#)), performs parameter estimation via evidence maximization or type-II maximum likelihood ([Tipping, 2000](#)). From the SBL procedure, it can be seen that the SBL cost function can lead to sparse representations. [Wipf and Rao \(2004\)](#) employed the Expectation-Maximization (EM) algorithm to minimize the SBL cost function. However, EM update rule has non-ideal convergence properties. Further, [Wipf and Nagarajan \(2010\)](#) proposed a re-weighted L_1 algorithm derived from the duality theory. Here, a similar algorithm will be derived from a new perspective such that the convergence of the algorithm can be analyzed in a simpler way. We will use this algorithm to solve Eq. (5). Note that all the candidate terms in the dictionary matrix Φ and the output y are approximated using the data by numerical methods. We consider the following matrix form with the model error ν ,

$$y = \Phi\theta + \nu, \quad (7)$$

where the noise vector ν is assumed to be i.i.d. Gaussian distributed with zero mean: $\mathcal{N}(0, \sigma^2 I_n)$ with I_n an $n \times n$ identity matrix, and θ is assumed to be sparse. Then the likelihood of the output y given the coefficients θ is

$$p(y|\theta) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\|y - \Phi\theta\|^2\right). \quad (8)$$

To enforce the sparsity, we assume the following form prior distribution

$$p(\theta; \gamma) = \prod_{i=1}^m (2\pi\gamma_i)^{-\frac{1}{2}} \exp\left(-\frac{\theta_i^2}{2\gamma_i}\right), \quad (9)$$

where $\gamma := [\gamma_1, \gamma_2, \dots, \gamma_m]^T$ is a vector of m hyperparameters. Although the hyperparametric vector γ is unknown, it determines the variance of each entry of θ and ultimately produces sparsity properties. In particular, if $\gamma_i = 0$, one can get $\mathbb{P}(\theta_i = 0) = 1$ immediately. Moreover, the hyperparametric vector can be estimated by type-II maximum likelihood. To this end, we first compute the marginalized probability distribution

$$p(y; \gamma) = \int p(y|\theta)p(\theta; \gamma)d\theta = (2\pi)^{-\frac{n}{2}} |\Sigma_y|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}y^T \Sigma_y^{-1} y\right),$$

where $\Sigma_y := \sigma^2 I_n + \Phi\Gamma\Phi^T$ and $\Gamma := \text{diag}(\gamma)$. Then we estimate γ by maximizing the marginalized probability distribution function. This is equivalent to minimizing $-\log p(y; \gamma)$, giving the cost function

$$\mathcal{L}_\gamma(\gamma) = \log |\Sigma_y| + y^T \Sigma_y^{-1} y, \quad (10)$$

that is,

$$\mathcal{L}_\gamma(\gamma) = \log |\sigma^2 I_n + \Phi\Gamma\Phi^T| + y^T (\sigma^2 I_n + \Phi\Gamma\Phi^T)^{-1} y. \quad (11)$$

On the other hand, for fixed hyperparameters, one can compute the posterior distribution of θ from the Bayes formula. As a result, one has

$$p(\theta|y; \gamma) = \mathcal{N}(\mu_\theta, \Sigma_\theta), \quad (12)$$

with $\mu_\theta = \sigma^{-2}\Sigma_\theta\Phi^T y$ and $\Sigma_\theta = (\sigma^{-2}\Phi^T\Phi + \Gamma^{-1})^{-1}$. After obtaining the estimate $\hat{\gamma}$ of the hyperparameter vector, the optimal Bayes estimation of coefficient vector θ is

$$\hat{\theta} = \hat{\mu}_\theta = (\Phi^T\Phi + \sigma^2\hat{\Gamma}^{-1})^{-1}\Phi^T y = \hat{\Gamma}\Phi^T(\sigma^2I_n + \Phi\hat{\Gamma}\Phi^T)y, \quad (13)$$

where $\hat{\Gamma} = \text{diag}(\hat{\gamma})$.

Consider that

$$\begin{aligned} y^T(\sigma^2I_n + \Phi\Gamma\Phi^T)^{-1}y &= \frac{1}{\sigma^2}y^Ty - \frac{1}{\sigma^4}y^T\Phi\Sigma_\theta\Phi^Ty = \frac{1}{\sigma^2}\|y - \Phi\mu_\theta\|^2 + \mu_\theta^T\Gamma^{-1}\mu_\theta \\ &= \min_{\theta} \left\{ \frac{1}{\sigma^2}\|y - \Phi\theta\|^2 + \theta^T\Gamma^{-1}\theta \right\}. \end{aligned}$$

Thereby, we define the cost function with respect to θ and γ as follows,

$$\mathcal{L}(\theta, \gamma) = \frac{1}{\sigma^2}\|y - \Phi\theta\|^2 + \theta^T\Gamma^{-1}\theta + \log|\sigma^2I_n + \Phi\Gamma\Phi^T| \stackrel{\Delta}{=} f(\theta, \gamma) + g(\theta, \gamma), \quad (14)$$

where

$$f(\theta, \gamma) = \frac{1}{\sigma^2}\|y - \Phi\theta\|^2 + \theta^T\Gamma^{-1}\theta, \quad g(\theta, \gamma) = \log|\sigma^2I_n + \Phi\Gamma\Phi^T|.$$

It is easy to verify that $f(\theta, \gamma)$ is convex for (θ, γ) and $g(\theta, \gamma)$ is concave for (θ, γ) . Define $\Omega := \mathbb{R}^m \times [0, \infty)^m$. Obviously, $\mathcal{L}(\theta, \gamma)$ is defined on the inner set $\text{int}(\Omega) := \mathbb{R}^m \times (0, \infty)^m$, and not defined on the boundary set $\partial\Omega := \Omega \setminus \text{int}(\Omega)$. In such case, it is difficult to ensure that \mathcal{L} has local minimum on the open set $\text{int}(\Omega)$. To alleviate this issue, some supplementary definitions are imposed on the boundary set $\partial\Omega$.

It is noting that from (9), if the hyperparameters are driven to zero, the associated coefficients will obey a degenerate distribution. In other words, if $\gamma_i = 0$, then $p(\theta_i; \gamma_i = 0) = \delta(\theta_i)$, which will lead to the posterior probability to satisfy

$$\mathbb{P}(\theta_i = 0 | y; \gamma_i = 0) = 1. \quad (15)$$

It is clear that the sparsity of θ and γ is closely related. More concretely, θ and γ have the same sparsity pattern, i.e., when $\theta_i = 0$, the optimal γ_i equals to zero as well, and vice versa. Note though that, such information is lost if only one variable is considered for optimization (Wipf et al., 2011). Thereout, we consider the following assumption.

Assumption 1 If θ_i and γ_i are driven to zero simultaneously, we only consider the following form of a path in (θ, γ) -space

$$\theta_i = \mathcal{O}(\gamma_i), \text{ as } \gamma_i \rightarrow 0^+. \quad (16)$$

Such an assumption is reasonable according to Eq. (13) and the KKT conditions of (14). More significantly, the limit of $\mathcal{L}(\theta, \gamma)$ at each point $(\theta, \gamma) \in \Omega$ exists under Assumption 1. In fact, $\lim_{\theta_i \rightarrow 0, \gamma_i \rightarrow 0} \theta_i^2/\gamma_i$ exists when along the path satisfying (16) in (θ, γ) -space. Based on this analysis, we supplement the definition of $\mathcal{L}(\theta, \gamma)$ to the boundary.

Definition 1 (i) For any point $(\theta, \gamma) \in \partial\Omega$, if θ_i and γ_i are zero simultaneously or $\gamma_i \neq 0$ for any index $i \in \{1, 2, \dots, m\}$, the limit of the function \mathcal{L} at point (θ, γ) exists and is unique when along the path satisfying (16), and hence we let this limit be the function value of \mathcal{L} at (θ, γ) .

(ii) If there exists an index i such that $\theta_i \neq 0$, $\gamma_i = 0$, the generalized limit of the function \mathcal{L} at point (θ, γ) is positive infinity when along the path satisfying (16), and hence we define $\mathcal{L}(\theta, \gamma) = M$, where M is a sufficiently large number.

Obviously, the boundary set can be divided into two disjoint sets, i.e., $\partial\Omega = \Omega_1 \cup \Omega_2$, where

$$\Omega_1 := \{(\theta, \gamma) | \text{there exists at least an index } i \text{ such that } \theta_i \neq 0, \gamma_i = 0\}$$

and $\Omega_2 := \partial\Omega \setminus \Omega_1$. Thereby, the SBL problem can equivalently be viewed as solving

$$(\hat{\theta}, \hat{\gamma}) \in \arg \min_{(\theta, \gamma) \in \Omega} \mathcal{L}(\theta, \gamma). \quad (17)$$

This is a DC programming problem that can be solved by CCP. The CCP is a iterative algorithm that solves the following sequence of convex programs,

$$(\theta^{(k+1)}, \gamma^{(k+1)}) \in \arg \min_{(\theta, \gamma) \in \Omega} \widehat{\mathcal{L}}(\theta, \gamma; \theta^{(k)}, \gamma^{(k)}), \quad (18)$$

or equivalently,

$$(\theta^{(k+1)}, \gamma^{(k+1)}) \in \arg \min_{(\theta, \gamma) \in \Omega} \left\{ \frac{1}{\sigma^2} \|y - \Phi\theta\|^2 + \sum_i \left(\frac{\theta_i^2}{\gamma_i} + c_i^{(k)} \gamma_i \right) \right\},$$

where

$$\begin{aligned} \widehat{\mathcal{L}}(\theta, \gamma; \theta^{(k)}, \gamma^{(k)}) &:= f(\theta, \gamma) + g(\theta^{(k)}, \gamma^{(k)}) + \langle \partial_\gamma g(\theta^{(k)}, \gamma^{(k)}), \gamma - \gamma^{(k)} \rangle, \\ c^{(k)} &:= \nabla_\gamma \log |\sigma^2 I_n + \Phi \Gamma^{(k)} \Phi^T|, \\ \Gamma^{(k)} &:= \text{diag}(\gamma_1^{(k)}, \gamma_2^{(k)}, \dots, \gamma_m^{(k)}). \end{aligned}$$

Note that, the supplementary definition for \mathcal{L} can be extended to the function $\widehat{\mathcal{L}}$. It follows that

$$\theta^{(k+1)} \in \arg \min_{\theta} \left\{ \|y - \Phi\theta\|^2 + 2\sigma^2 \sum_i \sqrt{c_i^{(k)}} |\theta_i| \right\}, \quad (19a)$$

$$\gamma_i^{(k+1)} = \frac{|\theta_i^{(k+1)}|}{\sqrt{c_i^{(k)}}}, \quad i = 1, 2, \dots, m. \quad (19b)$$

Given an initial point $(\theta^{(0)}, \gamma^{(0)}) \in \text{int}(\Omega)$, we can obtain an iterative sequence $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$. Moreover, $c^{(k)}$ can be analytically calculated by

$$c^{(k)} = \nabla_\gamma \log |\sigma^2 I_n + \Phi \Gamma^{(k)} \Phi^T| = \text{diag}[\Phi^T (\sigma^2 I_n + \Phi \Gamma^{(k)} \Phi^T)^{-1} \Phi]. \quad (20)$$

By the property of concave function, one has

$$\mathcal{L}(\theta, \gamma) \leq \widehat{\mathcal{L}}(\theta, \gamma; \theta^{(k)}, \gamma^{(k)}).$$

Combining with (18), one has

$$\begin{aligned}\mathcal{L}(\theta^{(k+1)}, \gamma^{(k+1)}) &\leq \widehat{\mathcal{L}}(\theta^{(k+1)}, \gamma^{(k+1)}; \theta^{(k)}, \gamma^{(k)}) \\ &\leq \widehat{\mathcal{L}}(\theta^{(k)}, \gamma^{(k)}; \theta^{(k)}, \gamma^{(k)}) \\ &= \mathcal{L}(\theta^{(k)}, \gamma^{(k)}).\end{aligned}\tag{21}$$

Next, easy computations give

$$\mathcal{L}(\theta^{(k)}, \gamma^{(k)}) \geq \log |\sigma^2 I_n| = 2n \log \sigma > -\infty.$$

Therefore, $\{\mathcal{L}(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ is a bounded monotonic non-increasing sequence. In fact, it is strictly monotonic before the iteration (19) reaches a local minimum (See Lemma 2). Therefore, the iterative process should be terminated when

$$\mathcal{L}(\theta^{(k)}, \gamma^{(k)}) - \mathcal{L}(\theta^{(k+1)}, \gamma^{(k+1)}) \leq \tau,\tag{22}$$

where τ is a tolerance.

In summary, the iterative process is showed in Algorithm 1.

Algorithm 1 SBL algorithm based on CCP

Require: $\Phi \in \mathbb{R}^{n \times m}$: design matrix; $y \in \mathbb{R}^n$: observation vector; τ : tolerance;

Ensure: $\theta^{(k+1)}$

- 1: initial $\Gamma^{(0)} = I_m$, $k = 0$;
 - 2: **repeat**
 - 3: $c^{(k)} = \text{diag}[\Phi^T(\sigma^2 I_n + \Phi \Gamma^{(k)} \Phi^T)^{-1} \Phi]$;
 - 4: $\theta^{(k+1)} \in \arg \min_{\theta} \left\{ \|y - \Phi \theta\|^2 + 2\sigma^2 \sum_i \sqrt{c_i^{(k)}} |\theta_i| \right\}$;
 - 5: $\gamma_j^{(k+1)} = \frac{|\theta_j^{(k+1)}|}{\sqrt{c_j^{(k)}}}$, $\Gamma^{(k+1)} = \text{diag}(\gamma^{(k+1)})$;
 - 6: $k = k + 1$
 - 7: **until** $(\mathcal{L}(\theta^{(k)}, \gamma^{(k)}) - \mathcal{L}(\theta^{(k+1)}, \gamma^{(k+1)}) \leq \tau)$
-

Remark 2 The constructed dictionary matrix from spatiotemporal data is generally high dimensional, resulting in a computationally expensive optimization problem (7). In Appendix C, we propose two strategies to reduce the computational complexity.

3. Theoretical Guarantees

In this section, we first prove the convergence of the proposed SBL algorithm based on CCP. Then, we give the properties of the SBL cost function. Finally, we prove the selection consistency and provide error bounds of the SBL algorithm.

3.1 Algorithmic Convergence

Based on the analysis in Section 2.2, the sequence $\{\mathcal{L}(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ is monotonic bounded. Thereout, $\{f(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ and $\{g(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ are bounded. To ensure the feasibility of the identification, we need the following assumption.

Assumption 2 All columns of the design matrix Φ are nonzero, i.e., $\Phi_i \neq 0$ ($i = 1, \dots, m$).

Under Assumption 2, the sequences $\{\theta^{(k)}\}_{k=0}^{\infty}$ and $\{\gamma^{(k)}\}_{k=0}^{\infty}$ must be bounded (otherwise it violates the boundedness for sequences $\{f(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$ and $\{g(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$). Let M_{γ} be an upper bound of sequence $\{\gamma^{(k)}\}_{k=0}^{\infty}$, i.e., $\gamma_i^{(k)} \leq M_{\gamma}$, for $\forall k \geq 0$, $i \in \{1, 2, \dots, m\}$. Then, by Eq. (20), one has

$$\begin{aligned}\sigma^{-2}\|\Phi_i\|^2 &\geq c_i^{(k)} = \Phi_i^T(\sigma^2 I_n + \Phi \Gamma^{(k)} \Phi^T)^{-1} \Phi_i \\ &\geq \Phi_i^T(\sigma^2 I_n + M_{\gamma} \Phi \Phi^T)^{-1} \Phi_i.\end{aligned}\quad (23)$$

Combining with (19) yields

$$0 < A \leq \frac{|\theta_i^{(k)}|}{\gamma_i^{(k)}} \leq B, \quad (24)$$

for $\gamma_i^{(k)} \neq 0$, $k \geq 0$, $i \in \{1, \dots, m\}$, where $A := \sqrt{\min_i \Phi_i^T(\sigma^2 I_n + M_{\gamma} \Phi \Phi^T)^{-1} \Phi_i}$ and $B := \sigma^{-1} \max_i \|\Phi_i\|$. This relation is consistent with Assumption 1.

By (18), define a point-to-set map:

$$\begin{aligned}\mathcal{A} : \quad \Omega &\longrightarrow \sigma(\Omega) \\ (\theta, \gamma) &\longmapsto \arg \min_{(\bar{\theta}, \bar{\gamma}) \in \Omega} \hat{\mathcal{L}}(\bar{\theta}, \bar{\gamma}; \theta, \gamma),\end{aligned}\quad (25)$$

where $\sigma(\Omega)$ stands for the σ -algebra generated by Ω . Obviously, $(\theta^{(k+1)}, \gamma^{(k+1)}) \in \mathcal{A}(\theta^{(k)}, \gamma^{(k)})$. In particular, if the dictionary matrix Φ has full column rank, $\mathcal{A}(\theta, \gamma)$ is a single point set in Ω , and hence the point-to-set map comes down to a point-point map. A fixed point of the map \mathcal{A} is a point (θ, γ) that satisfies $\{(\theta, \gamma)\} = \mathcal{A}(\theta, \gamma)$, whereas a generalized fixed point of map \mathcal{A} is a point (θ, γ) that satisfies $(\theta, \gamma) \in \mathcal{A}(\theta, \gamma)$. Let S be the generalized fixed point set of \mathcal{A} . Then, S is a stationary point set of (14), as shown in Lemma 1.

Lemma 1 If $(\theta^*, \gamma^*) \in S$, then (θ^*, γ^*) is a stationary point of the program (14).

Proof. If $(\theta^*, \gamma^*) \in S$, by the iteration (19) generated from map \mathcal{A} , there exists $z^* \in \partial \|\theta^*\|_1$ such that

$$\left\{ \begin{array}{l} \Phi_i^T(\Phi \theta^* - y) + \sigma^2 \sqrt{c_i^*} z_i^* = 0, \end{array} \right. \quad (26a)$$

$$\left\{ \begin{array}{l} \frac{|\theta_i^*|}{\sqrt{c_i^*}} = \gamma_i^*, i = 1, \dots, m, \end{array} \right. \quad (26b)$$

where $c^* := \nabla_{\gamma} \log |\sigma^2 I_n + \Phi \Gamma \Phi^T| \Big|_{\gamma=\gamma^*}$. Moreover, it is easy to show that (26) is the KKT conditions of (14). Hence, (θ^*, γ^*) is a stationary point of (14). \blacksquare

To show the convergence of the iterative sequence $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$, we introduce the following lemmas.

Lemma 2 Let $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$ be an iterative sequence generated by the point-to-set map \mathcal{A} with an initial point $(\theta^{(0)}, \gamma^{(0)}) \in \text{int}(\Omega)$. Then, one has

- (1) All points $(\theta^{(k)}, \gamma^{(k)})$ are in a compact set $F \subseteq \Omega$.
- (2) \mathcal{A} is monotonically decreasing with respect to \mathcal{L} , i.e., $\mathcal{L}(\bar{\theta}, \bar{\gamma}) \leq \mathcal{L}(\theta, \gamma)$ for any $(\bar{\theta}, \bar{\gamma}) \in \mathcal{A}(\theta, \gamma)$, and
 - (i) if $(\theta, \gamma) \notin S$, $\mathcal{L}(\bar{\theta}, \bar{\gamma}) < \mathcal{L}(\theta, \gamma)$, $\forall (\bar{\theta}, \bar{\gamma}) \in \mathcal{A}(\theta, \gamma)$,
 - (ii) if $(\theta, \gamma) \in S$, then either the algorithm terminates or $\mathcal{L}(\bar{\theta}, \bar{\gamma}) \leq \mathcal{L}(\theta, \gamma)$, $\forall (\bar{\theta}, \bar{\gamma}) \in \mathcal{A}(\theta, \gamma)$.

Proof. We first verify claim (1). According to the former analysis, we see that the sequence $\{\theta^{(k)}\}_{k=0}^{\infty}$ is bounded. Let m_{θ} and M_{θ} be the lower and upper bounds of sequence $\{\theta^{(k)}\}_{k=0}^{\infty}$ respectively. Then, there exists a compact set $F = [m_{\theta}, M_{\theta}]^m \times [0, M_{\gamma}]^m \subseteq \Omega$ such that $(\theta^{(k)}, \gamma^{(k)}) \in F$ for $k = 0, 1, \dots$

Next, we verify claim (2). By the definition of \mathcal{A} , one has

$$\mathcal{L}(\bar{\theta}, \bar{\gamma}) \leq \widehat{\mathcal{L}}(\bar{\theta}, \bar{\gamma}; \theta, \gamma) \leq \widehat{\mathcal{L}}(\theta, \gamma; \theta, \gamma) = \mathcal{L}(\theta, \gamma), \quad (27)$$

for any $(\bar{\theta}, \bar{\gamma}) \in \mathcal{A}(\theta, \gamma)$, i.e., \mathcal{A} is monotonically decreasing with respect to \mathcal{L} . Moreover, if $(\theta, \gamma) \notin S$, then $\mathcal{L}(\bar{\theta}, \bar{\gamma}) < \mathcal{L}(\theta, \gamma)$. In fact, if $\mathcal{L}(\bar{\theta}, \bar{\gamma}) = \mathcal{L}(\theta, \gamma)$, by (27), $\widehat{\mathcal{L}}(\bar{\theta}, \bar{\gamma}; \theta, \gamma) = \widehat{\mathcal{L}}(\theta, \gamma; \theta, \gamma)$, and therefore $(\theta, \gamma) \in S$. \blacksquare

Lemma 3 Let $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$ and $\{(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)})\}_{k=0}^{\infty}$ be two sequences in $\text{int}(\Omega) \cup \Omega_2$. If the following conditions:

- (1) $(\theta^{(k)}, \gamma^{(k)}) \rightarrow (\theta^*, \gamma^*) \in \Omega$,
- (2) $(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}) \rightarrow (\theta^{**}, \gamma^{**}) \in \text{int}(\Omega) \cup \Omega_2$,
- (3) $|\bar{\theta}_i^{(k)}|/\bar{\gamma}_i^{(k)}$ is bounded on $\{\bar{\gamma}_i^{(k)} \neq 0\}$,
- (4) $\bar{\theta}_i^{(k)}$ and $\bar{\gamma}_i^{(k)}$ are both zero and not zero at the same time,
- (5) $(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}) \in \mathcal{A}(\theta^{(k)}, \gamma^{(k)})$

hold. Then, one has

$$(\theta^{**}, \gamma^{**}) \in \mathcal{A}(\theta^*, \gamma^*). \quad (28)$$

Proof. See Appendix. \blacksquare

Theorem 3 Let $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$ be an iterative sequence generated by the point-to-set map \mathcal{A} with an initial point $(\theta^{(0)}, \gamma^{(0)}) \in \text{int}(\Omega)$. Then, the sequence $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$ converges to a stationary point of (14). Moreover,

$$\lim_{k \rightarrow \infty} \mathcal{L}(\theta^{(k)}, \gamma^{(k)}) = \mathcal{L}(\theta^*, \gamma^*), \quad (29)$$

where (θ^*, γ^*) is a stationary point of (14).

Proof. By Lemma 2, one has that all points $(\theta^{(k)}, \gamma^{(k)})$ are in a compact set. Then, there exists a convergent subsequence

$$(\theta^{(k_s)}, \gamma^{(k_s)}) \longrightarrow (\theta^*, \gamma^*), \text{ as } s \rightarrow \infty. \quad (30)$$

Combining with the monotonicity of \mathcal{A} with respect to \mathcal{L} , one has

$$\mathcal{L}(\theta^{(k_{s+1})}, \gamma^{(k_{s+1})}) \leq \mathcal{L}(\theta^{(k_s)}, \gamma^{(k_s)}) < +\infty. \quad (31)$$

By (24), one has $(\theta^*, \gamma^*) \in \text{int}(\Omega) \cup \Omega_2$. In fact, if $(\theta^*, \gamma^*) \in \Omega_1$, then by (30), one has

$$\lim_{s \rightarrow \infty} \mathcal{L}(\theta^{(k_s)}, \gamma^{(k_s)}) = +\infty.$$

This is a contradiction with (31). Next, if $(\theta^*, \gamma^*) \in \text{int}(\Omega)$, then by the continuity of \mathcal{L} on $\text{int}(\Omega)$, one has

$$\lim_{s \rightarrow \infty} \mathcal{L}(\theta^{(k_s)}, \gamma^{(k_s)}) = \mathcal{L}(\theta^*, \gamma^*); \quad (32)$$

If $(\theta^*, \gamma^*) \in \Omega_2$, then by (24) and the supplementary definition of \mathcal{L} on Ω_2 , Eq. (32) holds.

On the other hand, for the sequence $\{(\theta^{(k_{s+1})}, \gamma^{(k_{s+1})})\}_{s=1}^\infty$, there exists a convergent subsequence

$$(\theta^{(k_{s_j}+1)}, \gamma^{(k_{s_j}+1)}) \longrightarrow (\theta^{**}, \gamma^{**}), \text{ as } j \rightarrow \infty. \quad (33)$$

Similar with the proof of Eq. (32), we obtain that

$$\lim_{j \rightarrow \infty} \mathcal{L}(\theta^{(k_{s_j}+1)}, \gamma^{(k_{s_j}+1)}) = \mathcal{L}(\theta^{**}, \gamma^{**}). \quad (34)$$

By Lemma 2, we can conclude that the sequence $\{\mathcal{L}(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ is monotonic bounded, and therefore is convergent. Combining with Eqs. (32) and (34), one has

$$\mathcal{L}(\theta^*, \gamma^*) = \mathcal{L}(\theta^{**}, \gamma^{**}). \quad (35)$$

Furthermore, by (30) and Eq. (32), one has

$$(\theta^{(k_{s_j})}, \gamma^{(k_{s_j})}) \longrightarrow (\theta^*, \gamma^*), \text{ as } j \rightarrow \infty \quad (36)$$

and

$$\lim_{j \rightarrow \infty} \mathcal{L}(\theta^{(k_{s_j})}, \gamma^{(k_{s_j})}) = \mathcal{L}(\theta^*, \gamma^*). \quad (37)$$

Consider that $(\theta^{(k_{s_j}+1)}, \gamma^{(k_{s_j}+1)}) \in \mathcal{A}(\theta^{(k_{s_j})}, \gamma^{(k_{s_j})})$. Combining with (19) and (24), the sequences $\{(\theta^{(k_{s_j})}, \gamma^{(k_{s_j})})\}$ and $\{(\bar{\theta}^{(k_{s_j}+1)}, \bar{\gamma}^{(k_{s_j}+1)})\}$ satisfy the conditions of Lemma 3. Applying Lemma 3 yields

$$(\theta^{**}, \gamma^{**}) \in \mathcal{A}(\theta^*, \gamma^*). \quad (38)$$

Then, by (35), (38), and Lemma 2, one has

$$(\theta^*, \gamma^*) \in S.$$

Hence, by Lemma 1, (θ^*, γ^*) is a stationary point of (14). Moreover, the Eq. (29) follows from the convergence of $\{\mathcal{L}(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ and Eq. (32). \blacksquare

3.2 Properties of the Cost Function in (θ, γ) -space

This section mainly discusses the properties of cost function $\mathcal{L}(\theta, \gamma)$ whose minimization corresponds with maximally sparse solutions. In particular, we will focus on establishing the relationship between the cost function in (θ, γ) -space and the one in θ -space. This relationship facilitates us to understand how the underlying cost function promotes sparsity.

Let (θ^*, γ^*) be a local minimum point of \mathcal{L} . It is noting that a local minimum point of \mathcal{L} must be a stationary point of \mathcal{L} . Then, (θ^*, γ^*) satisfies the KKT conditions (26). Thereout, θ^* and γ^* have the same sparsity pattern. Consider that the log-det penalty term in $\mathcal{L}(\theta, \gamma)$ is concave for γ , which is frequently used for promoting sparse solution. Consequently, $\mathcal{L}(\theta, \gamma)$ can reach a local minimum at a sparse solution. To illustrate this result, we present a theorem that connects local minimum of \mathcal{L} with local minimum of the cost function in θ -space.

Theorem 4 Define the θ -space cost function

$$\mathcal{L}_\theta(\theta) := \|y - \Phi\theta\|^2 + \sigma^2 h(\theta), \quad (39)$$

with penalty

$$h(\theta) := \min_{\gamma \succeq 0} \left\{ \sum_i \frac{\theta_i^2}{\gamma_i} + \log |\sigma^2 I_n + \Phi \Gamma \Phi^T| \right\}. \quad (40)$$

Then, under Assumption 1 and Assumption 2, (θ^*, γ^*) is a global (local) minimum of (14) iff θ^* is a global (local) minimum of (39) with $\gamma_i^* = \theta_i^* / \sqrt{c_i^*}$, where c_i^* is the same as in (26).

Proof. Obviously, under Assumption 1, the term θ_i^2 / γ_i is definable. If (θ^*, γ^*) is a global minimum of (14), then it satisfies the KKT conditions (26), and hence θ^* is a global minimum of (39) with $\gamma_i^* = \theta_i^* / \sqrt{c_i^*}$. In turn, let θ^* be a global minimum of (39). With θ fixed, global minimum γ of the optimization problem from (40) satisfies the equation: $\gamma_i = \theta_i / \sqrt{c_i}$, where $c := \nabla_\gamma \log |\sigma^2 I_n + \Phi \Gamma \Phi^T|$. Then, there must exist some γ^* that minimizes $\mathcal{L}(\theta^*, \gamma)$ such that $\gamma_i^* = \theta_i^* / \sqrt{c_i^*}$. It leads to that (θ^*, γ^*) satisfies the KKT condition of (14). Note that $\mathcal{L}_\theta(\theta)$ is a strict upper bound on $\mathcal{L}(\theta, \gamma)$ with $\mathcal{L}_\theta(\theta) = \min_{\gamma \succeq 0} \mathcal{L}(\theta, \gamma)$. Hence, (θ^*, γ^*) is a global minimum of (39).

The relationship between global solutions to (14) and (39) can be extended to local solutions as well. In fact, since the optimization problem from (40) is convex with respect to the reparameterization of γ given by $v_i := \log \gamma_i$, all of its minima (if multiple exist) are connected. In other words, the minimization problem

$$h(\theta) := \min_v \left\{ \sum_i e^{-v_i} \theta_i^2 + \log |\sigma^2 I_n + \Phi e^\Upsilon \Phi^T| \right\},$$

where $\Upsilon := \text{diag}(v)$, is convex for v , and therefore all local minima are connected. Hence, the above result for global minima is also hold for local minima. \blacksquare

Theorem 4 shows that global (local) minimum of (14) can be obtained by minimizing a least squares with concave penalty. Note that concave, non-decreasing regularization

functions are well-known to enforce sparsity. Since $h(\theta)$ is such a function, it can induce sparsity to some extent. Furthermore, $h(\theta)$ provides a tighter approximation of ℓ_0 norm than ℓ_1 norm while generating many fewer local minimum than using ℓ_0 norm. However, usually $h(\theta)$ is non-factorable, i.e., $h(\theta) \neq \sum_j h_i(\theta_i)$, unlike traditional penalty, for example, ℓ_1 -norm penalty and ℓ_2 -norm penalty. It leads to the difficulty of studying the behaviors of the optimal solution of (39).

3.3 Selection Consistency and Error Bounds

In this subsection, we will compare the sparsity pattern of the local minimum of (14) with that of real coefficient vector in (5). For convenience, we denote the real coefficient vector θ in (7) by θ^{real} . Theorem 3 showed that the iterative sequence $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ generated from Algorithm 1 converges to a stationary point (θ^*, γ^*) of (14). And, each local minimum of (14) must be stationary point. Therefore, it is necessary to explore the connections between θ^* and θ^{real} . For the iterative sequence $\{(\theta^{(k)}, \gamma^{(k)})\}_{k=0}^\infty$ generated from Algorithm 1, one has

$$\begin{cases} \Phi_i^T (\Phi \theta^{(k+1)} - y) + \sigma^2 \sqrt{c_i^{(k)}} z_i^{(k+1)} = 0, \end{cases} \quad (41a)$$

$$\begin{cases} \frac{|\theta_i^{(k+1)}|}{\sqrt{c_i^{(k)}}} = \gamma_i^{(k+1)}, i = 1, \dots, m, \end{cases} \quad (41b)$$

where $c^{(k)} := \text{diag}[\Phi^T(\sigma^2 I_n + \Phi \Gamma^{(k)} \Phi^T)^{-1} \Phi]$ and $z^{(k+1)} := [z_1^{(k+1)}, \dots, z_n^{(k+1)}]^T \in \partial \|\theta^{k+1}\|_1$.

To simplify the Algorithm 1, we consider a special case (i.e., an orthonormal design $\Phi^T \Phi = I_m$). In such case, one has

$$\begin{aligned} & (\sigma^2 I_n + \Phi \Gamma^{(k+1)} \Phi^T)^{-1} \\ &= \sigma^{-2} I_n - \sigma^{-4} \Phi \Gamma^{(k+1)} \left[I_m + \sigma^{-2} \Gamma^{(k+1)} \Phi^T \Phi \right]^{-1} \Phi^T \\ &= \sigma^{-2} I_n - \sigma^{-2} \Phi \Gamma^{(k+1)} \left[\sigma^2 I_m + \Gamma^{(k+1)} \right]^{-1} \Phi^T, \end{aligned}$$

and hence,

$$\begin{aligned} c_i^{(k+1)} &= \sigma^{-2} \Phi_i^T \Phi_i - \sigma^{-2} \Phi_i^T \Phi \Gamma^{(k+1)} \left[\sigma^2 I_m + \Gamma^{(k+1)} \right]^{-1} \Phi^T \Phi_i \\ &= \sigma^{-2} - \sigma^{-2} e_i^T \Gamma^{(k+1)} \left[\sigma^2 I_m + \Gamma^{(k+1)} \right]^{-1} e_i \\ &= \sigma^{-2} - \sigma^{-2} \frac{\gamma_i^{(k+1)}}{\sigma^2 + \gamma_i^{(k+1)}} \\ &= \frac{1}{\sigma^2 + \gamma_i^{(k+1)}} = \frac{\sqrt{c_i^{(k)}}}{\sigma^2 \sqrt{c_i^{(k)}} + |\theta_i^{(k+1)}|}, \end{aligned} \quad (42)$$

where the first equality follows from the Woodbury equation. By Theorem 1, we see that $\{(\theta^{(k)}, \gamma^{(k)})\}$ converges to the stationary point (θ^*, γ^*) of (14). And, all stationary points

of (14) satisfy the KKT conditions (26). Hence, $c^{(k)}$ converges to c^* as k tends to infinity. Next, take the limit of Eq. (42), we obtain an implicit equation:

$$\sigma^2 c_i^* + |\theta_i^*| \sqrt{c_i^*} - 1 = 0, \quad (43)$$

which can be viewed as a quadratic equation with respect to $\sqrt{c_i^*}$. It can be easily calculated

$$\sqrt{c_i^*} = \frac{-|\theta_i^*| + \sqrt{|\theta_i^*|^2 + 4\sigma^2}}{2\sigma^2} = \frac{2}{|\theta_i^*| + \sqrt{|\theta_i^*|^2 + 4\sigma^2}}. \quad (44)$$

At this time, the weighted update rule is similar with the one in (Candes et al., 2008). Similarly, one has from (41)

$$\Phi_i^T (\Phi \theta^* - y) + \sigma^2 \sqrt{c_i^*} z_i^* = 0, \quad (45)$$

where $z^* := [z_1^*, z_2^*, \dots, z_m^*]^T$. Substituting Eqs. (7) and (44) into Eq. (45) yields

$$\theta_i^* - \theta_i^{\text{real}} + \frac{-|\theta_i^*| + \sqrt{|\theta_i^*|^2 + 4\sigma^2}}{2} z_i^* = \Phi_i^T v. \quad (46)$$

Obviously, $\theta_i^* = \theta_i^{\text{real}}$ if $\sigma^2 = 0$ and $v = 0$. To compare the true coefficient vector θ^{real} and the estimated one θ^* , we introduce the support as a performance measure. Here, the support of a given vector θ is defined as

$$\mathbb{S}(\theta) := \{i : |\theta_i| > \xi\},$$

where $\xi > 0$ is a fixed threshold that allows us to neglect some very small nonzero coefficients compared to those of coefficients. Let θ_{\min} be the minimum absolute value of all nonzero entries of θ^{real} . Then, when $\xi < \theta_{\min}$, one has $\mathbb{S}(\theta^{\text{real}}) = \{i : \theta_i^{\text{real}} \neq 0\}$, and hence $\theta_{\min} := \min_{i \in \mathbb{S}(\theta^{\text{real}})} |\theta_i^{\text{real}}|$. Moreover, we tackle the related issues with noise, demonstrating that the estimator θ^* is selection consistent in recovering sparsity pattern.

Definition 5 A estimator θ^* is selection consistent if $\mathbb{P}(\mathbb{S}(\theta^*) = \mathbb{S}(\theta^{\text{real}})) \rightarrow 1$, as $\sigma^2 \rightarrow 0$.

Theorem 6 Assume that $\Phi^T \Phi = I_m$. If $2\sigma\sqrt{\log 2m} \leq \xi < \frac{1}{2}\theta_{\min} - \sigma$, then $\mathbb{S}(\theta^{\text{real}}) = \mathbb{S}(\theta^*)$ with probability at least $1 - e^{-\frac{\xi^2}{4\sigma^2}}$, i.e., θ^* is selection consistent. Moreover,

$$|\theta_i^{\text{real}} - \theta_i^*| \leq \xi + \frac{\sigma^2}{\xi}, \quad \forall i \in \mathbb{S}(\theta^{\text{real}}).$$

Proof. For any $i \in \mathbb{S}(\theta^{\text{real}})$, one has $i \in \mathbb{S}(\theta^*)$ if $|\Phi_i^T v| < \frac{1}{2}\theta_{\min}$. In fact, if $i \notin \mathbb{S}(\theta^*)$, then by Eq. (46), one has

$$\theta_i^* + \frac{-|\theta_i^*| + \sqrt{|\theta_i^*|^2 + 4\sigma^2}}{2} z_i^* = \theta_i^{\text{real}} + \Phi_i^T v. \quad (47)$$

When $|\Phi_i^T v| < \frac{1}{2}\theta_{\min}$, one has

$$|\theta_i^{\text{real}} + \Phi_i^T v| \geq |\theta_i^{\text{real}}| - |\Phi_i^T v| \geq \frac{1}{2}\theta_{\min},$$

and

$$\left| \theta_i^* + \frac{-|\theta_i^*| + \sqrt{|\theta_i^*|^2 + 4\sigma^2}}{2} z_i^* \right| \leq \xi + \sigma.$$

By assumption, it is contradictory to Eq. (47). For any $i \notin \mathbb{S}(\theta^{\text{real}})$, one has $i \notin \mathbb{S}(\theta^*)$ if $|\Phi_i^T v| < \xi$. In fact, if $i \in \mathbb{S}(\theta^*)$, then from Eq. (46) and assumption

$$|\Phi_i^T v| = \frac{|\theta_i^* + \sqrt{|\theta_i^*|^2 + 4\sigma^2} z_i^*|}{2} > \xi, \quad (48)$$

which is a contradiction. Hence, $\mathbb{S}(\theta^{\text{real}}) = \mathbb{S}(\theta^*)$ holds on $\{\max_i |\Phi_i^T v| < \xi\}$. Since $\Phi^T \Phi = I_m$, then $\Phi_i^T v$, $i = 1, \dots, m$, are independent Gaussian random variables with zero mean and variance σ^2 . By the Hoeffding inequality, one has

$$\mathbb{P} [|\Phi_i^T v| \geq \xi] \leq 2e^{-\frac{\xi^2}{2\sigma^2}}.$$

Thereby,

$$\begin{aligned} \mathbb{P} \left[\max_i |\Phi_i^T v| < \xi \right] &= \prod_i \mathbb{P} [|\Phi_i^T v| < \xi] \\ &\geq \left(1 - 2e^{-\frac{\xi^2}{2\sigma^2}} \right)^m \\ &\geq 1 - e^{-\frac{\xi^2}{4\sigma^2}}, \end{aligned}$$

where the last inequality follows from the fact that $(1-x)^m \geq 1 - mx$, $\forall x \in [0, 1]$. Next, we establish the bounds of error $\theta_i^{\text{real}} - \theta_i^*$. For any $i \in \mathbb{S}(\theta^{\text{real}})$, by Taylor formula, there exists an real number $\eta \in (0, 4\sigma^2)$ such that

$$\sqrt{|\theta_i^*|^2 + 4\sigma^2} = |\theta_i^*| + \frac{2\sigma^2}{\sqrt{|\theta_i^*|^2 + \eta}}.$$

Substituting it into Eq. (46) yields

$$|\theta_i^{\text{real}} - \theta_i^*| = \left| \Phi_i^T v - \frac{\sigma^2}{\sqrt{|\theta_i^*|^2 + \eta}} \right| \leq \xi + \frac{\sigma^2}{\xi}.$$

This completes the proof. ■

Remark 7 Under the scaling $\sigma^2 \rightarrow 0$, the condition of Theorem 6 holds, and moreover that the probability of success converges to one. Additionally, from the proof process, we see that the estimator θ^* is sign-consistent, i.e., $\mathbb{P}(\text{sign}(\theta^*) = \text{sign}(\theta^{\text{real}})) \rightarrow 1$, as $\sigma^2 \rightarrow 0$, with the convention: $\text{sign}(\theta_i) = 0$, $|\theta_i| \leq \xi$.

The above analysis implies that the log-det term in (14) is factorable in the case of orthogonal design and can be expressed in the form independently of Φ . More generally, if $\Phi^T \Phi$ is diagonal, the log-det term is still factorable and has similar results.

4. Experiments

This section performs several experiments to demonstrate and validate our presented method. All codes to reproduce the results are publicly available at <https://github.com/HAIRLAB/S3d>. First, in §4.1, we illustrate the discovery process with S³d on eight prototypical PDEs used in various scientific fields and had several different mathematical forms. We seek to show that the proposed S³d method handles model diversity and typical errors or difficulties that arise in modeling. Secondly, in §4.2, we demonstrate the applicability of our S³d method to real data collected from an experiment on convection in an ethanol-water mixture in a long, narrow, annular container heated from below (Voss et al., 1999). We reconstruct the underlying PDE known as the Complex Ginzburg-Landau Equation (CGLE) without any *a priori* information. Finally in §4.3, we compare the performance of S³d, STRidge and DR in a identification task for the original dataset generated in (Rudy et al., 2017). It has been observed that S³d achieves better performance in terms of smaller parametric error. In addition, the number of samples that is required by S³d for training is smaller than the STRidge method.

4.1 Discovery of Prototypical PDEs

In this section, we collect our synthetic data by generating datasets from numerical solutions to canonical PDEs. The data collected on each PDE is composed into snapshot matrix $U \in \mathbb{R}^{n_x \times n_t}$. We denote the i th row and j th column of U by $u(x_i, t_j)$. To illustrate the robustness of our method in more realistic settings, we add 1% Gaussian noise into the synthetic data. We write our synthetic measurements with added Gaussian noise as

$$y(x_j, t_i) = u(x_j, t_i) + \eta_{j,i},$$

where $\eta_{j,i}$ takes a normal distribution $\mathcal{N}(\mu, \sigma_u^2)$ with mean $\mu = 0$ and standard deviation σ_u . The error from the noise-contained data is quantified by means of the root mean square error (RMSE),

$$Err(y) = \sqrt{\frac{1}{n_x n_t} \sum_{j=1}^{n_x} \sum_{i=1}^{n_t} (u(x_j, t_i) - y(x_j, t_i))^2}.$$

We also use different error norms for assessing our results. We compute the RMSE for evaluating the impact of noise and the mean-square error (MSE) and Standard Deviation (STD) for assessing our discovery results.

4.1.1 SINE-GORDON EQUATION

We begin our examples with the S³d method by considering the propagation of a slip in an infinite chain of elastically bound atoms lying over a fixed lower chain of similar atoms (Barone et al., 1971), whose dynamics are governed by the sine-Gordon equation (Ablowitz et al., 1973),

$$u_{tt} = u_{xx} - \sin(u), \quad (49)$$

where u_{xx} represents the elastic interaction energy between neighboring atoms, u_{tt} represents neighboring atoms' kinetic energy, and $\sin(u)$ represents the potential energy due to

the fixed lower chain. Eq. (49) admits various types of special solutions, such as small-amplitude solutions, traveling-wave solutions, and envelope wave solutions (Barone et al., 1971). Our objective is to extract the dynamical regime from a breather solution to Eq. (49) in the following form:

$$u(x, t) = 4 \arctan \left(\frac{\sin(t/\sqrt{2})}{\cosh(x/\sqrt{2})} \right).$$

For our analysis using S³d, we use a dataset consisting of 256 measurements of the breather solution. The measurements are equally sampled in time with an interval of $\Delta t = 0.0784$. We uniformly discretize the spatial domain $[-12.4, 12.4]$ of interest into 511 equal spaces (so that $m = 512$) using a grid step of $\Delta x = 0.0485$. We show a visualization of the discrete breather solution in Fig. 1. It is seen that the solution is not only symmetric but also periodic (a solitary wave propagates with periodicity $2\sqrt{2}\pi$). We organize the data into snapshot matrix, $U \in \mathbb{R}^{512 \times 256}$, and focus on discovering the governing equation with the nonlinear term $\sin(u)$ from our small periodic dataset.

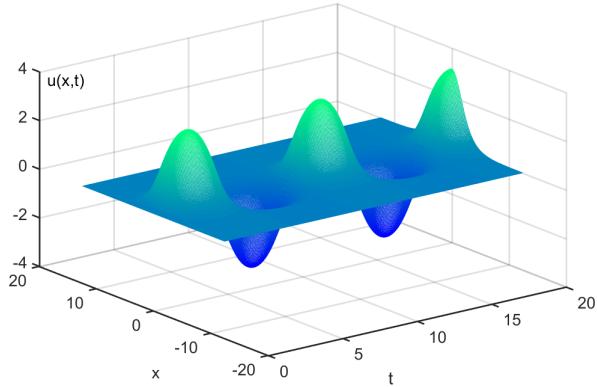


Figure 1: Numerical solution for Sine-Gordon equation (49).

First, we test the S³d method on a small dataset selected randomly from the synthetic data. We construct a dictionary consisting of $m = 25$ basis functions. The basis functions contain any combination of the state, $u(x, t)$, and the corresponding derivatives up to the third-order. In addition, we include periodic trigonometric functions such as, $\sin(u)$, $\cos(u)$, $\sin(u)\sin(u)$, $\sin(u)\cos(u)$, $\cos(u)\cos(u)$ in our dictionary. From the full dataset, the derivatives we include among the candidate terms can be approximated using the second-order finite difference method. Then, we randomly sub-sample 10 data points from the domain $[-2.6935, 2.1112] \times [0.3922, 4.2353]$, in which a wave package forms. Finally, when applying S³d, we correctly discover the Sine-Gordon equation with an accuracy of MSE=0.0706%, STD = 0.0895%, which is shown in Table 1. To check the criticality of the subsample domain for our discovery results, we conduct the same experiments on other domains where the wave package does not form. Our results show that, indeed, the subsampled domain is essential for Eq. (49).

Next, we perform a sensitivity analysis of S³d on noisy data. We choose to add 1% Gaussian noise into our synthetic data and compare this noisy data with the original one, and find a RMSE of $Err = 0.0084$. It is observed from Table 1 that more data points are required to correctly identify the underlying dynamics. In contrast with the noise-free case, we consider a domain in which the second wave package forms and randomly sub-sample 50 data points out of the 131072 ($512 \times 256 = 131072$) data points. Our results show that disturbance from noise results in a small loss of precision.

This example with the Sine-Gordon equation highlights the S³d method's ability to discover PDEs with only a small sample of the system's states. This example also shows that small samples including essential information regarding the system dynamics play key roles in detecting the underlying trigonometric function ($\sin(u)$).

Table 1: S³d Method for Sine-Gordon equation.

	points	$u_{tt} = u_{xx} - \sin(u)$	mean(err)±std(err)
Identified PDE (no noise)	10	$u_{tt} = 0.9999u_{xx} - 0.9987 \sin(u)$	0.0706%±0.0895%
Identified PDE (with noise)	50	$u_{tt} = 0.9920u_{xx} - 0.9982 \sin(u)$	0.4874%±0.4407%

4.1.2 FISHER'S EQUATION

For our second example, we consider the Fisher's equation:

$$u_t = \alpha u - \beta u^2 + du_{xx}, \quad (50)$$

which governs bacterial population dynamics ([Özış and Köröḡlu, 2008](#)) as an one-dimensional reaction diffusion model. The positive constant, α , denotes growth rate, β is the competition parameter and d is the diffusion coefficient. Fisher's equation has a long-standing history in mathematical modeling of propagation phenomena in distributed dissipative systems.

We consider the initial condition with a flat roof in the middle ([Mittal and Arora, 2010](#))

$$u_0(x) = \begin{cases} e^{10(x+1)}, & x < -1, \\ 1, & -1 \leq x \leq 1, \\ e^{-10(x-1)}, & x > 1, \end{cases}$$

and set coefficients $\alpha = \beta = 1$ and $d = 0.1$ in Eq. (50). The numerical solution, u , is computed using a difference scheme with spatial step length, $\Delta x = 0.06$, and time step length, $\Delta t = 0.01$. In Fig. 2, we plot the evolution of the solution, u . We observe that, initially, both the reaction term, $u(1-u)$, and the diffusion term, u_{xx} , have contours that round around the edges. As time goes on, the whole contour reduces and at the final stage, the limiting contour appears. The collected data is appropriate for identification since the limiting wave fronts and the limiting speed strictly depend on the system. We store the data into snapshot matrix $U \in \mathbb{R}^{201 \times 1000}$.

With results shown in Table 2, we chose $m = 16$ basis functions as our model class, specifically and attempted to discover the few active terms in the underlying dynamics from

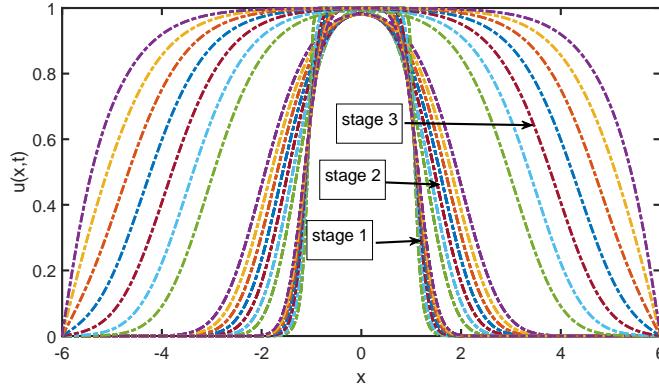


Figure 2: Numerical solution to Eq. (50). Shown is the contour plots of Stage 1 from $t = 0$ to $t = 0.4$ with $\Delta t = 0.06$; the contour plots of Stage 2 from $t = 1$ to $t = 3$ with $\Delta t = 0.3$; the contour plots of Stage 3 from $t = 5$ to $t = 10$ with $\Delta t = 0.8$.

Table 2: S³d Method for Fisher's equation.

	points	$u_t = u - u^2 + 0.1u_{xx}$	mean(err) \pm std(err)
Identified PDE (no noise)	10000	$u_t = 0.9998u - 1.0002u^2 + 0.0997u_{xx}$	0.0960% \pm 0.1342%
Identified PDE (with 1 % noise)	10000	$u_t = 1.0053u - 1.0079u^2 + 0.0971u_{xx}$	1.4104% \pm 1.3007%

our chosen basis functions, $1, u, u^2, u^3, u_x, u_{xx}, u_{xxx}, uu_x, uu_{xx}, uu_{xxx}, u^2u_x, u^2u_{xx}, u^2u_{xxx}, u^3u_x, u^3u_{xx}, u^3u_{xxx}$. We compute the derivatives using the explicit difference scheme. Furthermore, instead of specifying the re-sampling region, we randomly take 10000 out of total 201000 data points. Our method exactly identifies the main features u, u^2, u_{xx} from the random samples. The computed MSE and STD between the identified parameter and the true parameters is very small, as shown in Table 2.

We continue to test our method with noisy data. In this case, we must handle the error due to the numerical solution, in addition to the measurement error due to noise. We compute the RMSE, $Err = 0.0042$ from the noisy data. The RMSE is enhanced when computing the derivatives using the second-order difference method. For the added noise level, we measure the RMSE for the derivatives u_t, u_{xx} as $Err = 0.3023$ for term u_t , $Err = 2.96$ for term u_{xx} . To reduce the RMSE, we choose to, instead, approximate the derivatives using the polynomial interpolation method. Compared to the FD method, the measured RMSE for polynomial interpolation method is much smaller, with $Err = 0.0523$ for term u_t , $Err = 0.0888$ for term u_{xx} . We choose to randomly sub-sample 10000 data points from the noisy data, just as we did in the noiseless case. As shown in Table 2, with these design choices, we discover Eq. (50) with relatively good accuracy.

This example with Eq. (50) shows that both having an initial condition that stimulates the important features of a dynamical equation and having an accurate method for estimat-

ing the derivatives play important roles in the discovery process. In these two experiments, our method permits the use of random samples from the full dataset, similar to the Sine-Gordon equation. However, for Eq. (50), we randomly sub-sample the data from the entire domain.

4.1.3 KORTEWEG-DE VRIES EQUATION

Our third example is with the KdV equation with the form,

$$u_t = -uu_x - \epsilon u_{xxx}, \quad (51)$$

where u represents the height of wave at position x and time t , the nonlinear term uu_x represents steepening of the wave, the linear dispersive term u_{xxx} represents the spreading of the wave, and ϵ is a non-zero real constant. The KdV equation describes the unidirectional propagation of shallow water waves over a flat bottom, which theoretically explains the stability of the solution in the experiments of Scott Russell in 1834 ([Korteweg and De Vries, 1895](#)). In this example, we show the S^3d method on the two interacting soliton case. The interaction between different solitary waves yields particle-like behavior, which is an interesting phenomenon in the shallow water wave problem. We set $\epsilon = 0.000484$ and try to validate S^3d 's ability to discover the KdV equation with small ϵ .

We use the high-order compact difference method in ([Li and Chen, 2008](#)) to simulate the collision of a double soliton. The corresponding initial values are specified as follows:

$$u_0(x) = 3c_1 \operatorname{sech}^2(a_1(x - x_1)) + 3c_2 \operatorname{sech}^2(a_2(x - x_2)),$$

with $c_1 = 0.3, c_2 = 0.1, x_1 = 0.4, x_2 = 0.8, a_1 = 0.5\sqrt{c_1/\epsilon}, a_2 = 0.5\sqrt{c_2/\epsilon}$. Here, the physical domain is $x \in [0, 2]$, which is uniformly discretized into 255 equal spaces with a spatial grid, $\Delta x = 0.0078$. At the outset ($t = 0$), we place two solitary waves of different amplitudes along the x-axis, and both of them move in the left direction. We run the simulation until $t = 2$ with time steps of $\Delta t = 0.0015$. Fig. 3 shows the time evolution of $u(x, t)$ over a range of time values $[0, 2]$. We observe that the soliton with higher amplitude travels at a faster speed before interacting with the other soliton. After the two solitons separate, the two solitons still preserve their original amplitude. We organize the data into the snapshot matrix, $U \in \mathbb{R}^{256 \times 1301}$.

Using snapshot matrix, U , we first approximate the time derivatives and the spatial derivatives in a dictionary matrix consisting of $m = 12$ basis functions of the form, $1, u, u^2, u_x, u_{xx}, u_{xxx}, uu_x, uu_{xx}, uu_{xxx}, u^2u_x, u^2u_{xx}, u^2u_{xxx}$. We approximate the time derivatives using the second order central difference scheme. To test S^3d on a small dataset, we subsample 10000 data points out of the full 333056 data points. We show the interaction between these two solitons in Fig. 3. To guarantee simultaneous, equal representation of the two solitons in the subsampled data, we subsample the 10000 data points on the domain, $[0.3922, 0.7765] \times [0, 0.3062]$. Table 3 lists the discovery results. We see that the S^3d method, only using a small portion of data, can effectively identify the KdV equation with $MSE = 0.0855\%$ and $STD = 0.0145\%$.

In Table 3, we also show our results in the noisy case. After adding 1% Gaussian noise into the original dataset, we compute a $RMSE = 0.002$ between the synthetic noisy data and the original one. This error makes it difficult to approximate the derivatives. To address

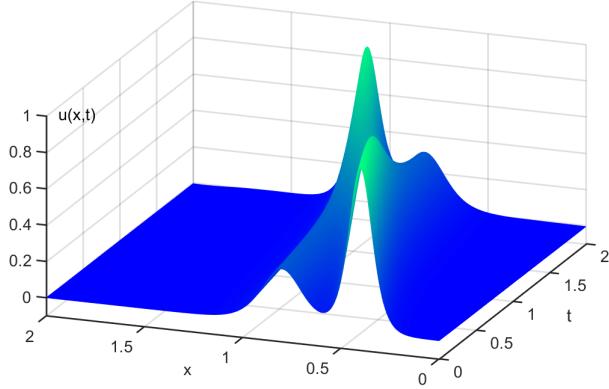


Figure 3: Numerical solution for Korteweg-de Vries equation (51).

this error, we use the polynomial approximation with good noise immunity to estimate our derivatives. Then, by increasing the number of data points (18750 of full data points), we can correctly discover the KdV equation and are able to detect parameters that highly coincide with the true values, yielding $MSE = 0.9982\%$ and $STD = 0.9995\%$.

Table 3: S³d Method for Korteweg-de Vries equation.

	points	$u_t = -0.000484u_{xxx} - uu_x$	mean(err)±std(err)
Identified PDE (no noise)	10000	$u_t = -0.000484u_{xxx} - 0.999247uu_x$	0.0855%± 0.0145%
Identified PDE (with noise)	18750	$u_t = -0.000483u_{xxx} - 0.982950uu_x$	0.9982%± 0.9995%

4.1.4 FITZHUGH-NAGUMO EQUATION

For our fourth example, we discover the FitzHugh-Nagumo model with two components in the following form:

$$\begin{aligned} u_t &= d \cdot u_{xx} + f(u, \alpha) - w, \\ w_t &= \beta u - \gamma w, \end{aligned} \tag{52}$$

where, $u(x, t)$ denotes the transmembrane electrical potential at position x and time t , and w denotes the likelihood that a particular class of ion channel is open (Hoffman et al., 1997). The diffusion coefficient, d , represents the electrical conductivity. The function, $f(u, \alpha) \triangleq u(u - \alpha)(1 - u)$, represents the reactive properties of the medium with reactive coefficient, $\alpha = 0.2$. We are particularly interested in the minimally reactive coefficients, $\gamma = 0.001 \ll 1$, $\beta = 0.002 \ll 1$, and diffusion coefficient, $d = 1$. Eq. (52) is a simplified version to the Hodgkin-Huxley model for nerve conduction, which generally can be used to model activation and deactivation dynamics of a spiking neuron (FitzHugh, 1961; Nagumo et al., 1962).

First, we collect the data from a numerical solution to Eq. (52) with the following initial condition: $u_0 = \exp(-x^2)$, $w_0 = 0.2 \exp(-(x+2)^2)$ and zero boundary conditions. We compute the numerical solution using the finite element method. We create a uniform spatial grid from the interval of interest, $x \in [-15, 15]$, by dividing the interval into 512 equal subintervals with a fine mesh size of $\Delta x = 0.0587$. Then, we discretize the spatial derivative terms in Eq. (52) (Skeel and Berzins, 1990). We integrate the resulting ODE to obtain a finite element solution to Eq. (52). Fig. 4 shows the evolution of state variables u and w . The sampled data is organized in two data matrices, $U, W \in \mathbb{R}^{512 \times 401}$, with the time step size, $\Delta t = 0.05$.

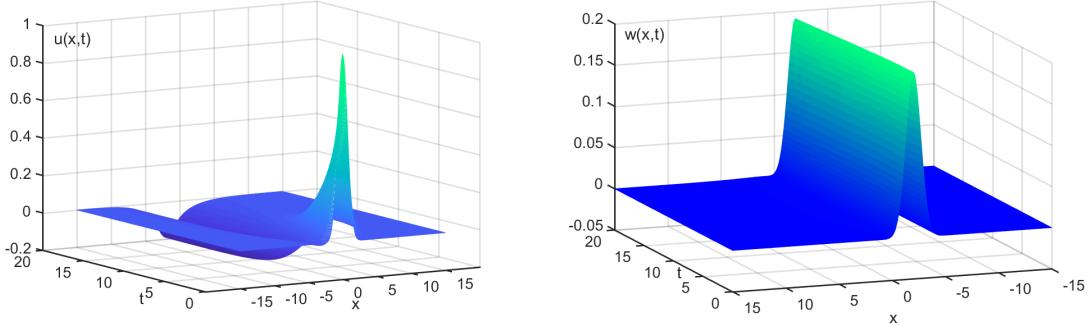


Figure 4: The numerical solution to the FitzHugh-Nagumo equation (52) with $\gamma = 0.001$, $\beta = 0.002$.

Next, we test the ability of the S³d method to identify Eq. (52) from small datasets randomly selected from the full synthetic dataset. For this, we sub-sample the collected data from time $t_1 = 0.25$, to time $t_2 = 5.2$, with 100 spatial points. We use a model class consisting of $m = 25$ basis functions for our candidate terms, including the derivative of u up to the third order. We approximate the derivative terms in the candidate dictionary using an explicit scheme. By applying the S³d algorithm on the subsampled 10000 data points, we correctly identify Eq. (52) with a relatively good accuracy. The results is summarized in Table 4. Note that the dynamics in the selected region incorporate the interaction between the reactive term and the diffusion term in Eq. (52), enabling our method to discover the Eq. (52) from the subsampled data points.

To illustrate the robustness of the S³d method, we add 1% Gaussian noise to the simulated data. The synthetic noisy data is compared with the original simulation data at each of the grid points, and the measured RMSE is $Err(y) = 0.000735$ for u and $Err(y) = 0.000382$ for w . With the noisy data, we perform the same experiment as we did for the data without noise. In the experiment, we sub-sample the noise-contained data from time $t_1 = 0$ to time $t_2 = 8.95$ with 100 spatial points. In addition, we find that the finite difference scheme is sensitive to the noise. Instead, we propose using polynomial interpolation to approximate the derivatives. For different data points and the same dictionary, the proposed S³d method correctly identifies Eq. (52) (see Table 4).

Table 4: S³d discovers the FitzHugh-Nagumo equation.

	points	$u_t = d \cdot \Delta u + u(u - \alpha)(1 - u) - \omega$ $\omega_t = \beta u - \gamma \omega$	mean(err)±std(err)
Identified PDE (no noise)	10000 (u) 10000 (w)	$u_t = -0.1991u + 1.2004u^2$ $-1.0031u^3 + 1.0023u_{xx} - 1.0005\omega$ $\omega_t = -0.000993\omega + 0.001999u$	0.2649%±0.2462%
Identified PDE (with 1% noise)	18000 (u) 27000 (w)	$u_t = -0.1981u + 1.1762u^2$ $-0.9499u^3 + 1.0025u_{xx} - 0.9920\omega$ $\omega_t = -0.001004\omega + 0.001952u$	1.6795%±1.6690%

Finally, for an even more intensive study of our method, we further test the proposed S³d method on very small parameters. We take $\gamma = 0.0001$, $\beta = 0.0002$ and plot the evolution of the computed solution u and w in Fig. 5. We observe that compared with the

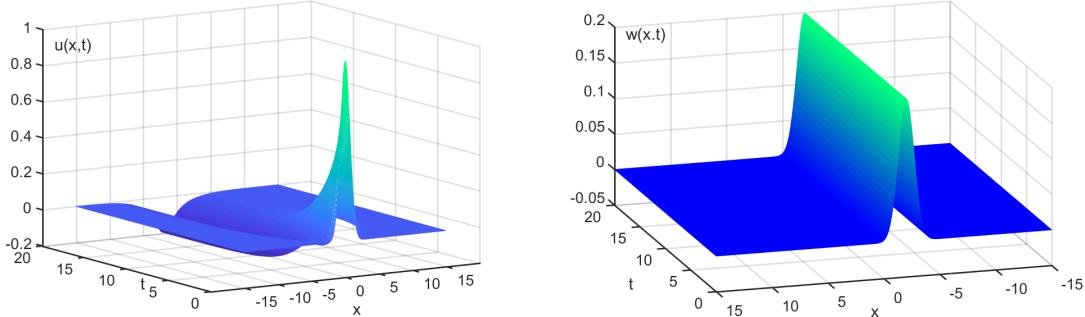


Figure 5: The numerical solution to the FitzHugh-Nagumo equation (52) with $\gamma = 0.0001$, $\beta = 0.0002$.

dynamics in Fig. 4, the smaller coefficients produce a slightly different behavior. Despite the different behavior, our method is able to effectively infer the small parameters β and γ as $\beta = 0.0002$ and $\gamma = 0.000099$.

4.1.5 NONLINEAR SCHRÖDINGER'S EQUATION

Now, we showcase the discovery process with S³d on collected data that is in complex domain. To deal with the complex data, we extend the S³d method to the system with complex variables, as introduced in Appendix D. We consider the following nonlinear Schrödinger's equation (also called the Gross-Pitaevskii equation) for condensed neutral atoms in an harmonic trap (Edwards and Burnett, 1995):

$$iu_t = -\frac{1}{\epsilon}(|u|^2 - 1)u - \epsilon u_{xx}, \quad (53)$$

where $u(x, t)$ is the Bose-Einstein condensate (BEC) wave function. In mean field theory, Eq. (53) is a classical approximation that describes a quantum mechanical nonrelativistic

many body system with a two body δ -function interaction, which is difficult to observe. However, we show that the S^3d method can extract the dynamical regimes from the data collected from the wave function.

First, we obtain the discrete version of the wave function, $u(x, t)$, using the Fourier spectral method. We consider the initial condition, $u(x, 0) = 4x^2 \exp(-2x^2) \exp(1i)$, and choose parameter $\epsilon = 0.3$. We perform the simulation on the domain, $[-\pi, \pi] \times [0, 8]$, with time step, $\Delta t = 0.016$, and spatial step, $\Delta x = 0.0123$, such that $n_t = 501$ is the number of measurements form the snapshot matrices, $U \in \mathbb{R}^{512 \times 501}$. Fig. 6 shows the evolution of the (condensate) wave function. Note, the solution has sufficiently variability in both space and time. The identified results from the synthetic data and synthetic noise data is shown

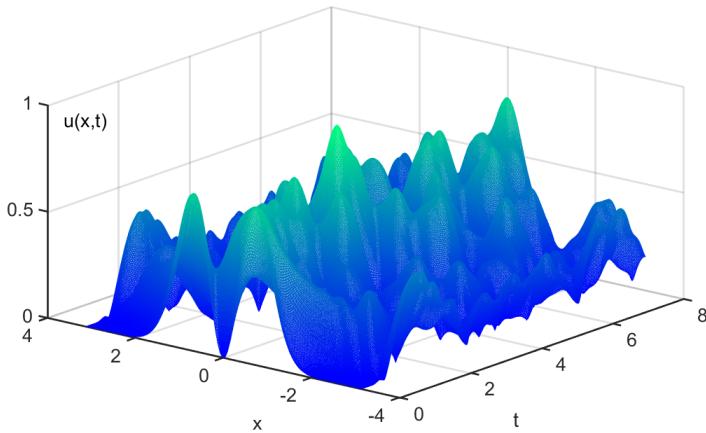


Figure 6: Numerical solution for nonlinear Schrödinger's equation (53).

in Table 5. We use $m = 40$ basis functions for our candidate terms; these terms include $u, |u|, u_x, u_{xx}, u_{xxx}$ and their combinatorial terms. Since the wave is variable, we choose to use a high precision method (i.e. the Pade scheme (Li and Chen, 2008) for noise free data and the polynomial interplant for the noisy data case) for estimating the derivatives in the example. Additionally, our experiment shows that the method to estimate the derivatives is critical for the success of discovery process.

Table 5: S^3d Method for Schrödinger's equation.

	points	$u_t = i\frac{3}{10}u_{xx} + i\frac{10}{3} u ^2u - i\frac{10}{3}u$	mean(err) \pm std(err)
Identified PDE (no noise)	10000	$u_t = i0.3u_{xx} + i3.3333u u ^2 - i3.3333u$	0.0011% \pm 0.0007%
Identified PDE (with noise)	10000	$u_t = i0.3013u_{xx} + i3.2947u u ^2 - i3.3187u$	0.6775% \pm 0.4162%

4.1.6 KLEIN-GORDON EQUATION

Consider the Klein-Gordon equation with a cubic nonlinearity

$$u_{tt} = u_{xx} - u - u^3, \quad (54)$$

which occurs as a relativistic wave equation. The state variable, $u(x, t)$, represents the wave displacement at position, x , and time, t . The Klein-Gordon equation is related to the Schrödinger equation and has many applications such as spin waves or nonlinear optics (Dodd et al., 1982). We rewrite Eq. (54) as a first-order system,

$$\mathbf{u}_t = A\mathbf{u} + F(\mathbf{u}),$$

with $\mathbf{u} = [u \ u_t]^T$, $F(\mathbf{u}) = [0 \ -u - u^3]^T$, and $A = \begin{bmatrix} 0 & 1 \\ \partial_{xx} & 0 \end{bmatrix}$. This first-order system is similar to the Fisher's model in Eq. (50), but with a different nonlinearity. To identify the Klein-Gordon equation from measurement data, we regard the second derivative of u in time as the output, Y .

To generate the data, we use the finite difference method to solve Eq. (54) with the initial condition, $u_0 = x^2$, and with zero boundary condition. We divide the domain, $[0, 1] \times [0.003, 3.003]$, into a 101×1001 mesh with the step size, $\Delta x = 0.01$, and time step, $\Delta t = 0.003$. In Fig. 7, we depict the evolution of the numerical solution, u . The discrete

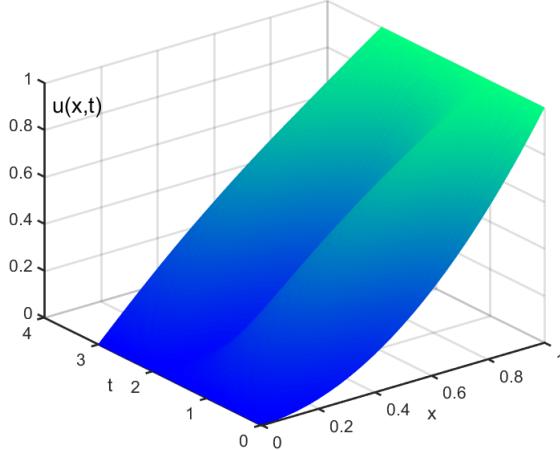


Figure 7: Numerical solution for Klein-Gordon equation (54).

version of the numerical solution forms the snapshot matrix, $U \in \mathbb{R}^{101 \times 1001}$.

Table 6 lists our identified results for the noise-free and the noisy cases, respectively. In the noise-free case, we randomly sample 10000 out of a total of 101×1001 data points in the original dataset. The candidate terms consist of $m = 20$ basis functions, including the derivatives of u up to the third order. We use the explicit difference scheme to compute the derivative terms in the dictionary matrix, Φ . The MSE and STD in Table 6 confirm the accuracy of the identified results.

Table 6: S³d Method results for Klein-Gordon equation.

	points	$u_{tt} = u_{xx} - u - u^3$	mean(err)±std(err)
Identified PDE (no noise)	10000	$u_{tt} = 0.9994u_{xx} - 0.9995u - 0.9998u^3$	0.0426%± 0.0220%
Identified PDE (with 1% noise)	15500	$u_{tt} = 0.9987u_{xx} - 0.9833u - 1.0303u^3$	1.61%± 1.4513%

In the noisy case, we add 1% Gaussian noise into the original dataset and compare the synthetic noisy data with the original one. The computed RMSE is $Err(y) = 0.003$. The noisy experiment demonstrates that Eq. (54) is sensitive to noise.

We use the Proper Orthogonal Decomposition (POD) method to denoise our data. In Appendix B, we introduce the POD method in detail. By taking the first 4 POD modes to construct the POD basis, i.e., $\Psi \in \mathbb{R}^{101 \times 4}$, we obtain the de-noised dataset, $\tilde{U} = \Psi A \in \mathbb{R}^{101 \times 1000}$, with the POD coefficients, A , for each POD mode, $A = \Psi^T U$. With our de-noised dataset, \tilde{U} , we estimate the derivative in the candidate terms using the polynomial approximation, which is proven to be more robust than the explicit difference method. We pay particular attention to the region from time $t_1 = 0.603$ to $t_2 = 2.1$ with a compact spatial interval, $[0.3, 0.6]$. The results in Table 6 confirms the effectiveness of the use de-noising and discovery methods.

4.1.7 KURAMOTO-SIVASHINSKY EQUATION

Consider the Kuramoto-Sivashinsky (KS) equation subject to periodic boundary condition, $u(x + L, t) = u(x, t)$:

$$u_t + uu_x + u_{xxxx} + u_{xx} = 0, \quad (55)$$

which models a small perturbation, $u(x, t)$, of a metastable planar front or interface. Here, L is the size of a typical pattern scale and all other dimensional parameters are eliminated by rescaling. The KS equation can exhibit chaotic behavior (Hyman and Nicolaenko, 1986) and is known to have an inertial manifold (Foias et al., 1988; Constantin et al., 2012). This complexity is one of the KS equation's interesting features. In the KS equation, the fluctuations generated by the instability is dissipated by the (stabilizing) fourth-order derivative, u_{xxxx} . However, estimating the term from the data is challenging and results in large deviation. In this KS equation example, we exhibit the S³d method's ability to handle PDEs with a fourth derivative term.

To obtain the data, we simulate Eq. (55) with the Fourier spectral method in (Trefethen, 2000). We start with the initial condition on interval, $[0, 32\pi]$:

$$u_0(x) = \cos(x/L)(1 + \sin(x/L)),$$

with $L = 16$. Note, the sampling period is crucial for estimating the derivatives of the candidate terms. To estimate the derivatives, we use a time step of $\Delta t = 0.1$ and a spatial step of $\Delta x = 0.0491$ with $n_x = 2048$ spatial points and $n_t = 1001$ time points. We carry out the pseudo-code in (Trefethen, 2000) to solve the KS equation. We plot the time evolution

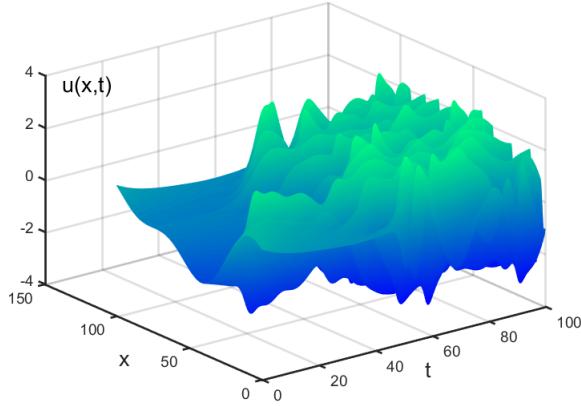


Figure 8: Numerical solution for Kuramoto-Sivashinsky equation (55).

of Eq. (55) in Fig. 8. The data are stored into the snapshot matrix, $U \in \mathbb{R}^{2048 \times 1001}$.

We choose a pool of $m = 36$ basis functions, including the derivatives of solution u up to the fifth order. We use the fourth-order compact Pade scheme to approximate the derivative terms. As expected, the resulting derivatives are more accurate than that are obtained using other methods. We further subsample the generated data from time $t_1 = 40$ to time $t_2 = 40.6$ using almost all of the spatial points and then we use the sub-sampled 14203 data points for the sparse identification step. By adjusting the regularization parameter λ in our proposed algorithm, our method correctly identifies Eq. (55). We present the discovery results in Table 7.

Table 7: S³d Method for Kuramoto-Sivashinsky equation.

	points	$u_t = -uu_x - u_{xx} - u_{xxxx}$	mean(err) \pm std(err)
Identified PDE (no noise)	14203	$u_t = -1.0000uu_x - 1.0000u_{xx}$ $-1.0000u_{xxxx}$	0.0022% \pm 0.0001%
Identified PDE (with 1% noise)	59210	$u_t = -0.9095uu_x - 0.9214u_{xx}$ $-0.9238u_{xxxx}$	8.1781% \pm 0.7678%

We add 1% Gaussian noise to the original dataset and the measured RMSE is $Err = 0.0107$. In this example, we apply the POD method to de-noise our data. By projecting the data onto the 33 POD modes obtained with the a threshold of 99.99%, we generate a new snapshot, $\tilde{U} \in \mathbb{R}^{2048 \times 1001}$, from the original snapshot: $\tilde{U} = \Psi A$ with Ψ consisting of 33 POD modes and the corresponding POD coefficients, A . The measured RMSE of $Err = 0.0028$ is much smaller than the noisy snapshot. Using the new snapshot, \tilde{U} , we use the polynomial interpolation method to estimate the derivatives. We are interested in re-sampling the time domain, [4, 96]. We use 59210 out of the total 2050048 data points for the sparse identification step. However, with the large dataset, our S³d method suffers from the curse of dimensionality.

To reduce computational time, we apply the SVD technique to the dictionary matrix, $\Phi \in \mathbb{R}^{59210 \times 36}$, giving us a 72 dimensional subspace denoted as Ψ . We project the dictionary matrix onto the 72 dimensional subspace such that we obtain the reduced output vector, $\bar{Y} = \Psi^T Y \in \mathbb{R}^{72 \times 1}$, and the reduced dictionary matrix, $\bar{\Phi} = \Psi^T \Phi \in \mathbb{R}^{72 \times 36}$. We show in Table 7 that applying the S³d method to data with de-noising from the POD method and dimensionality reduction works well.

4.1.8 NAVIER-STOKES EQUATION

Consider the incompressible Navier-Stokes (NS) equations for the unsteady two-dimensional flows on torus with vorticity/stream function formulation ([Shu and Richards, 1992](#)):

$$\begin{aligned}\omega_t + u\omega_x + v\omega_y &= (\omega_{xx} + \omega_{yy})/Re, \\ \psi_{xx} + \psi_{yy} &= -\omega,\end{aligned}\tag{56}$$

where $u = \psi_y$ represents the horizontal velocity component, $v = -\psi_x$ represents the vertical velocity component, $\omega = u_y - v_x$ represents the vorticity and Re is the Reynolds number (based on the radius of the cylinder and the free-stream velocity, V_∞). The vorticity formulation is attractive for the accurate solution of high Reynolds number planar or axisymmetric NS equations ([Strang, 2007](#)). In this example with the NS equations, the flow field is described by four field quantities, each of which needs to be measured for data.

We use a combination of the Fourier spectral method and the Crank-Nicolson method ([Strang, 2007](#)) to solve Eq. (56) with initial condition in $x \times y \in [0, 2\pi] \times [0, 2\pi]$:

$$\begin{aligned}\omega(x, y) &= \exp\left(-\frac{1}{5}[x^2 + (y + \frac{\pi}{4})^2]\right) + \exp\left(-\frac{1}{5}[x^2 + (y - \frac{\pi}{4})^2]\right) \\ &\quad - \frac{1}{2} \exp\left(-\frac{2}{5}[(x - \frac{\pi}{4})^2 + (y - \frac{\pi}{4})^2]\right).\end{aligned}$$

We perform the spatial discretization on a uniform grid with $\Delta x = 0.0628$ and $\Delta y = 0.0628$. By substituting the Fourier approximation of the solution, u, v, ω , into Eq.(56), we further integrate the resultant ODEs using a time step of $\Delta t = 0.1$ using the Crank-Nicolson scheme. In Fig. 9, we plot the 2D NS pseudo-spectral solver on the torus with Reynolds

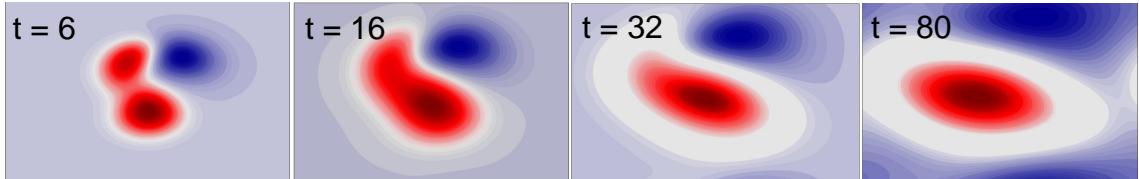


Figure 9: Flow field for 2D NS equation (56) on the torus depicted in time $t = 6, t = 16, t = 32$ and $t = 80$.

number, $Re = 100$. The discrete version of the numerical solution forms the snapshot matrices, $U, V, W \in \mathbb{R}^{100 \times 100 \times 1001}$.

Table 8: S³d Method for NS equation.

	points	$\omega_t(x, t) = 0.01\omega_{xx} + 0.01\omega_{yy}$ $-u\omega_x - v\omega_y$	mean(err)±std(err)
Identified PDE (no noise)	10000	$\omega_t(x, t) = 0.01\omega_{xx} + 0.01\omega_{yy}$ $-1.0068u\omega_x - 0.9987v\omega_y$	0.37%±0.2305%
Identified PDE (with 1% noise)	20000	$\omega_t(x, t) = 0.0101\omega_{xx} + 0.0097\omega_{yy}$ $-1.0051u\omega_x - 1.0013v\omega_y$	0.9447%±1.3498%

We report the discovery results in Table 8. The dictionary matrix, Φ , consists of $m = 60$ basis functions. The underlying dynamics is described with the three state variables, u, v, w . Thus, our dictionary of basis functions include the linear combinations of the variables, $u, v, \omega, u_x, v_x, u_{xx}, v_{xx}$. With snapshots, U, V, W , we use the explicit difference scheme to approximate the first or second order derivatives in the noise-free case. We see from the discovery results that our method has strong performance with the noise-free dataset, a small MSE, and a small STD. We enumerate the re-sampling region in Table 8.

In the noisy case, we use the 4th-order compact Pade scheme with the S³d method. We add 1% Gaussian noise to the original snapshot and the resulting RMSE is $Err = 0.4823$ for u , $Err = 0.4338$ for v , and $Err = 1.7111$ for w . We apply the POD method to de-noise our data. We reshape the snapshot U into a 10000×1001 matrix, \bar{U} ; we do the same shaping for W and V . By projecting the data onto 33 POD modes obtained using a threshold of 99.99 fixed, we generate a new snapshot, $\tilde{U} \in \mathbb{R}^{10000 \times 1001}$, from the matrix, \bar{U} : $\tilde{U} = \Psi A$ with Ψ consisting of 33 POD modes and the corresponding POD coefficients, A . Then, we reverse the process such that snapshots, $\bar{U}, \bar{V}, \bar{W} \in \mathbb{R}^{100 \times 100 \times 1001}$. The computed RMSE is 0.0056 for u , 0.0051 for v , and 0.0412 for w . We pay special attention to the Pade scheme. With the de-noised dataset, we approximate the first-order derivatives in each grid point and then we compute the second-order derivatives using the Pade scheme. After we compute the dictionary matrix, we correctly discover the NS equation as shown in Table 8.

4.2 Discovery of Complex Ginzburg-Landau Equations from Binary-Fluid Convection Experiment

This section specifies the discovery of the complex Ginzburg-Landau equation (CGLE) from experimental measurements. Traveling-wave (TW) convection in binary fluids is a known, mainstream ansatz technique for studying the physical mechanism of non-equilibrium pattern-forming systems ([Cross and Hohenberg, 1993](#)). Substantial experiments on TW convection unfolds many different dynamical states of TW. A major challenge in the scientific study of TW is quantitatively understanding and predicting the underlying dynamics. First-principle models based on the CGLE and its variants ([Cross, 1988; Newell, 1974](#)) have long been used to quantitatively explain experimental observations. We apply the proposed S³d method to see whether we can discover the CGLE from experimental data alone.

4.2.1 DATA DESCRIPTION

The experimental data comes from an experiment conducted in an annular cell ([Voss et al., 1999; Kolodner, 1992](#)), detailed in Table 9:

Table 9: Overview of the considered real-world TW data. More experimental details for each dataset can be found in (Voss et al., 1999; Kolodner, 1992).

Experiment Label	bifurcation parameter	Samples	Sampling interval	Space interval	Truncation
cal06172/car06172	$\gamma = 9.32 \times 10^{-3}$	988×180	1.5625	0.458167	228
cal06182/car06182	$\gamma = 4.22 \times 10^{-3}$	800×180	1.5625	0.458167	40
cal06192/car06192	$\gamma = 1.77 \times 10^{-3}$	1000×180	1.5625	0.458167	240
cal06212/car06212	$\gamma = 6.38 \times 10^{-3}$	1000×180	1.5625	0.458167	240
cal06222/car06222	$\gamma = 12.07 \times 10^{-3}$	1031×180	1.5625	0.458167	271
cal06242/car06242	$\gamma = 14.03 \times 10^{-3}$	1000×180	1.5625	0.458167	240
cal06252/car06252	$\gamma = 16.28 \times 10^{-3}$	1000×180	1.5625	0.458167	240

- The special letter “cal” and “car” respectively represent the left-going waves and right-going ones. The files record the experimental data;
- Seven bifurcation parameters, ε , scaled by the characteristic time, τ_0 (i.e., $\gamma = \varepsilon\tau_0^{-1}$), represent the same experiments but conducted with different bifurcation parameters;
- All samples used for the analysis of S^3d are truncated from the 21th sampling point to the 780th sampling point, such that a total of 180×760 data points are used for the analysis. Then, we form the snapshot matrices, $U \in \mathbb{C}^{180 \times 760}$.

We define the measured left-going TW state (e.g. “cal06172”, “cal06182”) as the left-going complex amplitudes, A_L , in Eq. (57). We do the same for the data for the right-going complex amplitudes A_R , such as “car06172”, “cal06182”, etc.. Note that the data is being collected periodically. Then, we use the circular Pade scheme to approximate the derivatives of the amplitude, A_R (or A_L), from the full dataset (i.e., 180×760).

In Fig.10, we show the dynamics of the weakly nonlinear wave packets for $\varepsilon\tau_0 = 1.77 \times 10^{-3}$, which propagate around the system in opposite directions.

4.2.2 DISCOVERY OF CGLE

We summarize the best-fit basis functions and parameters for each database in Table 10. Here, we illustrate the performance of the S^3d method in two steps: first, it determines the key nonlinearities and then recovers the coefficients from experimental data. We introduce 18 basis functions to construct the dictionary: $A_{L,R}$, $|A_{L,R}|$, $A_{L,R}^2$, $|A_{L,R}|^2 A_{L,R}^3$, $|A_{L,R}|^3$, $A_{L,R}^2 |A_{L,R}|$, $A_{L,R} |A_{L,R}|^2$, $\partial_x A_{1,2}$, $\partial_{xx} A_{1,2}$, $A_{L,R} \partial_x A_{L,R}$, $A_{L,R} \partial_{xx} A_{L,R}$, $A_{L,R}^2 \partial_x A_{L,R}$, $A_{L,R}^3 \partial_x A_{L,R}$, $A_{L,R}^3 \partial_{xx} A_{L,R}$, $|A_{R,L}|^2 A_{L,R}$. The above amounts to 3rd-order Volterra expansions of $A_{L,R}$ and $|A_{L,R}|$. We investigate the performance of the S^3d algorithm on the last three sets of data with larger signal-to-noise ratio (e.g. $\varepsilon\tau_0^{-1} \geq 12.07 \times 10^{-3}$) as discussed in (Voss et al., 1999) in Table 9. We take the maximum iteration number to be $k_{max} = 15$ and empirically adjust the regularization parameter, λ .

In our experiments, we observe that the S^3d algorithm yields a stable result (shown in Fig. 11): for each data set, there exists a critical number of the basis functions before which

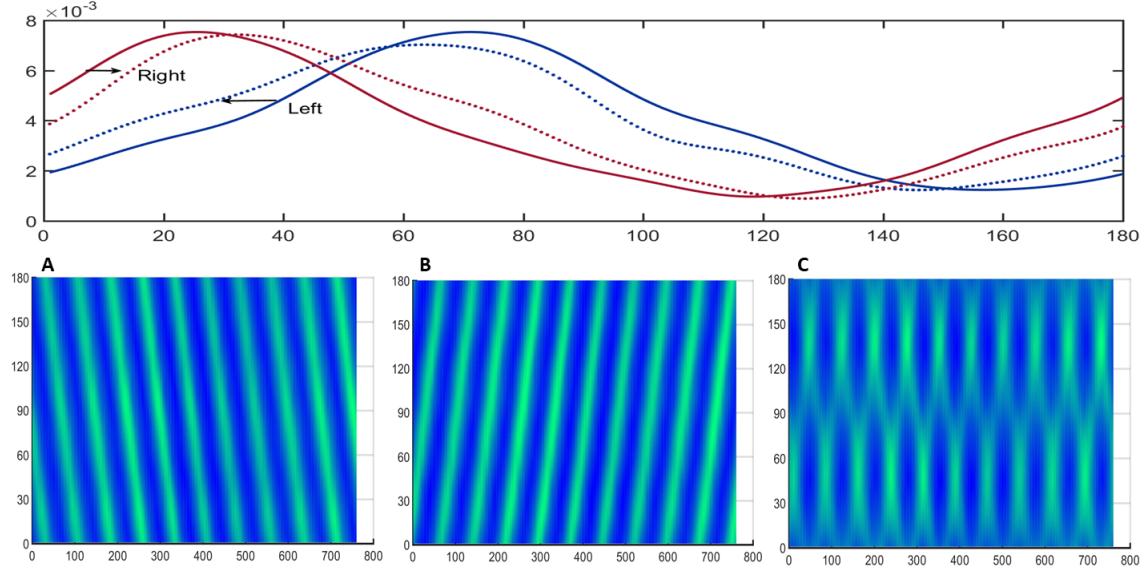


Figure 10: **The dynamics of the observed TW state for $\varepsilon\tau_0 = 1.77 \times 10^{-3}$.** The top line shows the physical change of the amplitudes of the left-going (blue full line and dashed line) and right-going (red full curve and dashed line) waves components at a particular time and at the next time's data. A: The behavior of left-going amplitudes ; B: The behavior of right-going amplitudes; C: The behavior of the sum of left- and right-going amplitudes.

the fitting-error slowly increases and after which the error begins to grow rapidly. The table in Fig. 11 enumerates the critical number for the bifurcation parameter, $\gamma = 14.03 \times 10^{-3}$; specifically, we discover 5 key nonlinearities. In addition, the discovered key nonlinearities are consistent for two other datasets with bifurcation parameter $\gamma = 12.07 \times 10^{-3}$ and bifurcation parameter $\gamma = 16.28 \times 10^{-3}$.

The following discovered equation has the following form:

$$\begin{aligned} \tau_0(\partial_t + s\partial_x)A_R &= \varepsilon(1 + ic_0)A_R + \omega_0^2(1 + ic_1)\partial_x^2A_R + g(1 + ic_2)|A_R|^2A_R \\ &\quad + h(1 + ic_3)|A_L|^2A_R, \\ \tau_0(\partial_t - s\partial_x)A_L &= \varepsilon(1 + ic_0)A_L + \omega_0^2(1 + ic_1)\partial_x^2A_L + g(1 + ic_2)|A_L|^2A_L \\ &\quad + h(1 + ic_3)|A_R|^2A_L. \end{aligned} \tag{57}$$

In this model (CGLE), A_R (or A_L) is the complex amplitude of a right-going (or left-going) wave with group velocity s , the parameter ω_0 is a characteristic length scale and τ_0 is a characteristic time which is determined experimentally by measuring the growth rate, $\gamma = \varepsilon\tau_0^{-1}$, at several values of ε and fitting the slope, $c_0 - c_3$ are dispersion coefficients, g is a nonlinear saturation parameter, and h is a nonlinear coupling coefficient which reflects the stabilizing interaction between oppositely propagating traveling-wave components (Voss et al., 1999). CGLE is frequently used to model non-equilibrium patterning-forming systems.

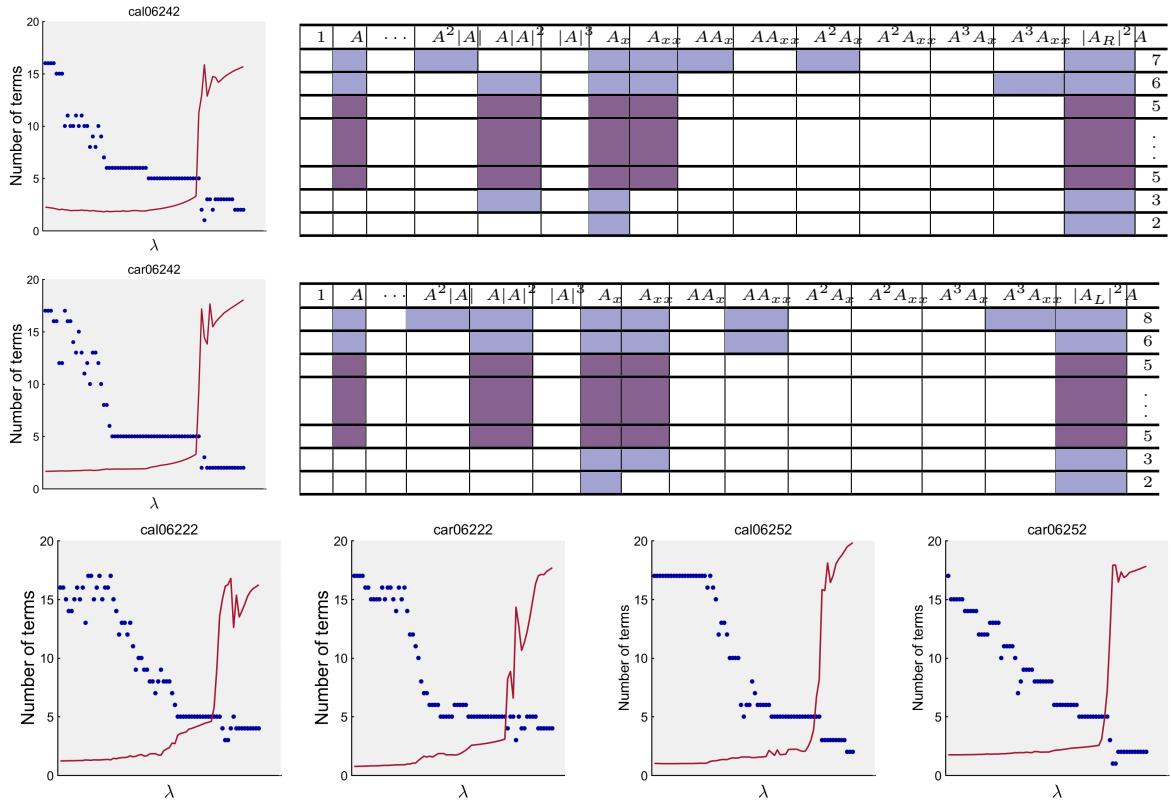


Figure 11: **The accelerations of fitting error due to change in the number of feature terms.** We empirically adjust the regulation parameter, λ , while observing the change of the fitting error. Three sets of data were chosen to test S³d: car/cal06242, car/cal06222, and car/cal06252 from top to bottom. For each dataset, the fitting error slowly increased and tended to a relatively stable value (5 features in Eq. (57) emerge), then that error began to rapidly grow. In the table above, we show the selected basis functions for different λ .

Thus, the structural analysis of the CGLE model further allows us to quantitatively explain the weakly nonlinear dynamics of TW convection. To further fine-tune the coefficients of CGLE, we apply the S³d algorithm to each data set in Table 10 with the discovered nonlinearities. To show our identified results in an accessible form, we introduce the concept of the leave-one-out in machine learning to discuss the best method to choose the regularization parameters λ and obtain the best-fit coefficients. We accept the parameters that lead to the smallest fitting error on the test sets: $err = \frac{\|Y - \Phi\omega\|^2}{\|Y\|^2}$, by which, the identified coefficients are reported in Table 10. We can see that the identified results resemble the theoretical values and experimental ones (Voss et al., 1999; Kolodner, 1992; Kolodner et al., 1995) and reference therein, except for some obscure parameters in theory and experiment.

Previous works (Voss et al., 1999) developed nonlinear regression analysis using prior knowledge in physics to infer the CGLE model for the description of binary-fluid convection

Table 10: **Summary of the identified parameters of seven sets of real data:** The data are labeled according to “right”(car) and “left”(cal). Each row represents the identified coefficients by S^3d .

Coefficients in the discovered complex Ginzburg-Landau equation									
Experiment Label	s	$\varepsilon\tau_0^{-1}$	$\xi_0^2\tau_0^{-1}$	$g\tau_0^{-1}$	$h\tau_0^{-1}$	$\varepsilon\tau_0^{-1}c_0$	$\xi_0^2c_1\tau_0^{-1}$	$gc_2\tau_0^{-1}$	$hc_3\tau_0^{-1}$
cal06172cb	0.5895	0.0065	0.3392	-17.1619	-6.4595	0.0443	-0.2185	-89.0851	-139.7500
car06172cb	0.5982	0.0079	0.9210	17.7800	-25.2965	0.0640	-0.1243	-206.3442	-150.6041
cal06182cb	0.6268	0.0057	0.0000	-78.0836	-0.8076	-0.0094	-0.3774	0.0000	-154.4844
car06182cb	0.6308	0.0081	0.0935	-90.2362	-35.0632	-0.0059	-0.4397	-55.8157	-157.3323
cal06192cb	0.5930	0.0054	0.0148	-139.4494	-2.5230	-0.0066	-0.6457	-37.6346	-152.7833
car06192cb	0.5916	0.0087	0.0158	-198.6799	-34.9115	-0.0066	-0.9189	-68.1182	-141.3105
cal06212cb	0.6063	0.0057	0.0639	-45.0095	-7.9602	0.0033	-0.2732	-35.3580	-154.5495
car06212cb	0.6131	0.0042	0.0202	-26.1553	-33.0555	0.0019	-0.1489	-17.7879	-168.7832
cal06222cb	0.5741	0.0173	1.7492	40.1657	-15.1882	0.0281	-0.0424	-379.5853	-136.5903
car06222cb	0.5771	0.0127	1.3000	29.9928	-16.6945	0.0114	-0.0894	-272.3181	-161.5083
cal06242cb	0.5509	0.0138	1.0701	16.2302	-17.3846	0.0555	-0.1505	-234.8376	-144.0718
car06242cb	0.5597	0.0150	1.1368	13.1359	-4.5687	0.0567	-0.2108	-237.9784	-149.1559
cal06252cb	0.5634	0.0199	1.1690	2.1061	-5.2084	0.0867	-0.2690	-267.3739	-169.0035
car06252cb	0.5481	0.0131	1.1075	22.3429	-3.9619	0.0794	-0.1245	-234.6182	-179.6777

near onset. We employ the proposed S^3d to discover CGLE from experimental data and lead to comparable parameters of the oscillatory behavior in the experimental system. The proposed method advances the theory without any prior knowledge in physics and can consistently discover the underlying mechanistic PDE from multiple datasets with different bifurcation parameters.

4.2.3 VALIDATION

To further validate our method, we suggest simulating the identified CGLE to reconstruct the experimental chaos observed in the one-dimensional TW convection. We employ the

Fourier spectral method (Trefethen, 2000) to simulate the identified CGLE model

$$\begin{aligned}\partial_t A_R + s\partial_x A_R &= \varepsilon\tau_0^{-1}(1+ic_0)A_R + \xi_0^2\tau_0^{-1}(1+ic_1)\partial_x^2 A_R \\ &\quad + g\tau_0^{-1}(1+ic_2)|A_R|^2 A_R + h\tau_0^{-1}(1+ic_3)|A_L|^2 A_R, \quad 0 < x < L, t > 0, \\ \partial_t A_L - s\partial_x A_L &= \varepsilon\tau_0^{-1}(1+ic_0)A_L + \xi_0^2\tau_0^{-1}(1+ic_1)\partial_x^2 A_L \\ &\quad + g\tau_0^{-1}(1+ic_2)|A_L|^2 A_L + h\tau_0^{-1}(1+ic_3)|A_R|^2 A_L, \quad 0 < x < L, t > 0,\end{aligned}\tag{58}$$

where parameter values are given in Table 10. We consider system length of $L = 82.0119$ and computational time of $T = 625$ ($t_0 = 0, t_f = 625$), which are determined based on a sampling interval with spatial step, $\Delta x = 0.458167$, and time step, $\Delta t = 1.5625$. We discretize the left-going complex wave amplitude, A_L , and the right-going complex wave amplitude, A_R , in space on a uniform 180 mesh with spectral approximations of the spatial derivatives, and integrated in time using an explicit Runge-Kutta formula with 401 time points.

Note, we conducted the experiments in an annular container where the TW system comprises periodic boundary conditions (Voss et al., 1999). This corresponds to the boundary condition:

$$A_L(x, t) = A_L(x + L, t), A_R(x, t) = A_R(x + L, t).$$

Here, we take any column of the snapshot matrix, U , as initial conditions. Unequivocally, we take the 622th column of the snapshot matrix, U , as initial values for data labelled “cal/car06172”, the 600th column for data labelled “cal/car06192”, the 600th column for data labelled “cal/car06212”, and the 1st column for data labelled “cal/car06182”. Given the initial and boundary conditions, we obtain the discretized version of the solution, A_L, A_R , to Eq. (58) and plot their time evolution.

Fig. 12 presents the comparison between the simulation and the experiments on TW convection. The simulations in the first three row of Fig. 12 have a resolution of 180×400 samples and the forth row has a resolution of 180×200 . Both resolutions correspond to reconstruction times $T = 625$ and $T = 312.5$, respectively. It is observed that the simulations agree well with the experiments. In addition, we also find that the reconstruction results are different under different initial values. In particular, we take the 658th column taken from data labelled “cal/car06192” as the initial value. The regular bursts of left- and right-going TW are seen with the simulation time up to $T = 1187$, and the amplitudes of wave approach the real values, as shown in Fig. 13. Thus, the reconstruction from our identified CGLE model bears a striking resemblance to the experimentally observed nonlinear states of TW convection on binary-fluid convection.

4.3 Comparison of S³d, STRidge and Douglas-Rachford

We first compare the S³d method with the STRidge regression algorithm or PDE functional identification of nonlinear dynamics (PDE-FIND) proposed in (Rudy et al., 2017). We consider the original dataset in (Rudy et al., 2017). These dataset are generated by simulating the KdV equation with single- and double- soliton solution, the Burger’s equation, the Quantum Harmonic Oscillator and the nonlinear Schrödinger equation. Here we neglect the simulation details and refer readers to (Rudy et al., 2017). Table 11 shows the identified results of applying the S³d method to the dataset. We use a much smaller number

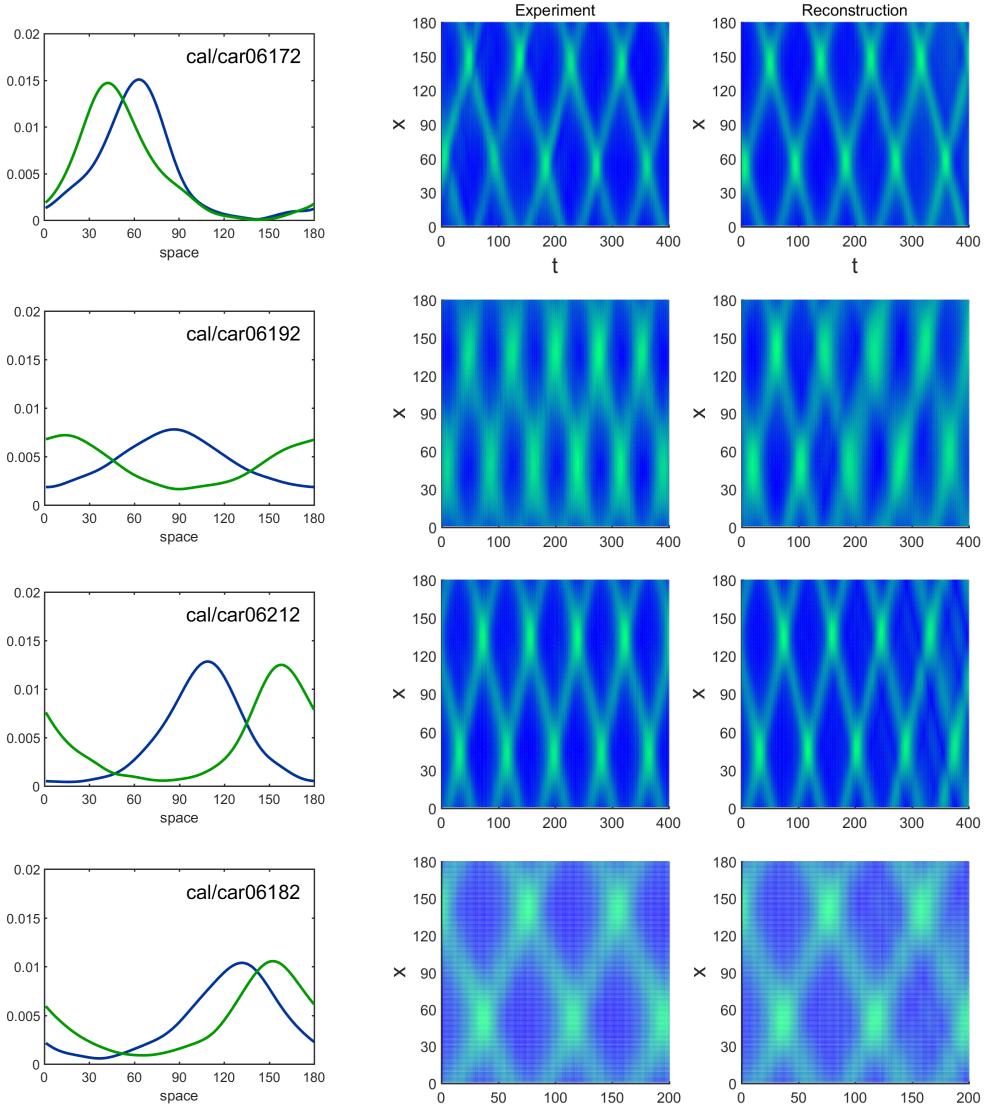


Figure 12: **Comparison of spatiotemporal evolution of the superposition of the left- and right-TW amplitude between the simulation and the experiment.** The different initial values: the 622^{th} column of the snapshot matrix U for data labelled “cal/car06172”, the 600^{th} column for data labelled “cal/car06192”, the 600^{th} column for data labelled “cal/car06212” and the 1^{st} column for data labelled “cal/car06182”, are selected for simulation (the first column). The first 400 (or 200) snapshots of the real data and the simulation data are shown with a resolution of 180×400 (column 2 and column 3).

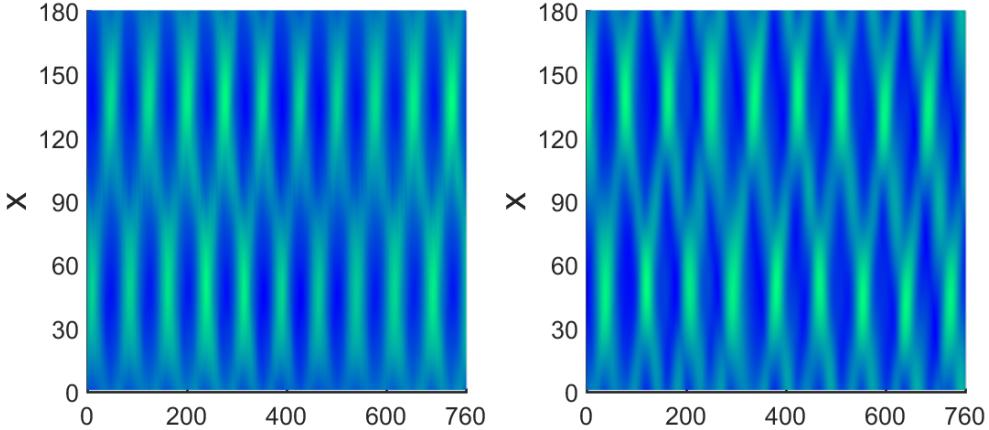


Figure 13: **Comparison of spatiotemporal evolution of the superposition of the left- and right- TW amplitude between the simulation and the experiment.** Using the extended time $T = 1187$ and initial value (i.e., 658th column of snapshot matrix), a regular busts of TW is seen .

of samples for a better discovery of these PDEs in terms of smaller parametric error. We further investigate the impact of noise on both methods, S³d and PDE-FIND, using data generated from the Quantum Harmonic Oscillator and the nonlinear Schrödinger equation ([Rudy et al., 2017](#)). In particular, we consider data with 1% Gaussian noise, as shown in Table 12. S³d is able to discover both equations from such datasets.

We then apply the PDE-FIND algorithm to Data 1(QH) in Table 12. For PDE-FIND, we perform a two-stage grid search for the regularization parameter λ and tolerance value d_{tol} . In the first stage, we search through a coarse grid from $\lambda = 10^{-8}$ to 10^{-1} and $d_{tol} = 10^{-5}$ to 8.5×10^3 , such that a total of 1152 sets of parameters are tested for the PDE identification in order to make a fair comparison. However, the PDE-FIND is not able to discover the QH equation for any combinations: either it discovers the true terms u_{xx} and $\frac{x^2}{2}u$ together with a few redundant terms or fails to discover the true terms. To further demonstrate the results, in the second stage, we perform a more fine-grained search in the vicinity of that tuple (λ, d_{tol}) that gives the most closest results. More specifically, the regularization parameter λ varies now from 10^{-4} to 0.0864 with the step 10^{-4} , such that a total of 124,352 sets of parameters are picked. By this, the PDE-FIND method still fail in identifying the QH equation.

For nonlinear Schrödinger equation, the same two-stage grid search are used for the identification of the PDE using PDE-FIND. A total of 14,256 sets of parameters were exhaustively tested for the discovery. However, the PDE-FIND fails in discovering the nonlinear Schrödinger equation (the code used for comparison was in the Github repository).

Now, we compare our method with the work in ([Schaeffer, 2017](#)), which applies the Douglas-Rachford algorithm ([Lions and Mercier, 1979; Combettes and Pesquet, 2011](#)) to learn PDEs from the data generated by the viscous Burgers' equation, inviscid Burgers' equation, the Swift-Hohenberg equation, the Cahn-Hilliard equation, et al. However, the

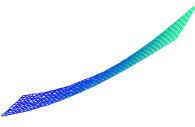
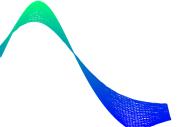
method makes an unrealistic assumption that only time derivatives are corrupted by additive Gaussian noise (see the section ‘‘Simulation and numerical experiments’’ in (Schaeffer, 2017)). Because the measured variables are always associated with noise, it will not only lead to variations in the time-derivatives but also to higher-order derivatives. Such variations will further propagate through the later identification process, which significantly increases the difficulty in identifying PDEs. The proposed S³d method weakens their assumption by allowing adding Gaussian noise to the measured variables. As a result, both time and spatial derivatives are corrupted by noise. Indeed, this is a major challenge in the discovery process.

We compared the identification performance to the datasets generated by the NLS equation and the QH equation. All these equations were discovered by our proposed method. Note that there exist four tunable parameters in the Douglas-Rachford algorithm (ours has one tuning parameter): γ , μ , the maximum number of iterations (MaxIt), balancing parameter λ . Starting with $\gamma = 0.1$, $\mu = 0.1$, $\lambda = 10^{-3}$ and MaxIt = 5000 (which are picked according to examples in their papers), a grid search method was used to find its best set of parameters at reasonable ranges of these parameters. In total, we checked a total of 26244 sets of parameters for the PDE identification, which took more than 24 hours to complete. The algorithm failed to identify the correct features of the NLS equation, and gave many redundant terms across all searches. Similarly, we observed a similar failure of PDE discovery for the QH equation. In summary, the proposed Douglas-Rachford algorithm failed to discover those equations with the same limited and noisy data that was used (successfully) in S³d.

Table 11: Comparison of identified results using PDE-FIND method and S³d method with the original database in (Rudy et al., 2017).

Example		S ³ d method		PDE-FIND method	
Terms	Identified results	No. of samples	MSE(err)-STD(err)	No. of samples	MSE(err)-STD(err)
Korteweg-de Vries (Single-soliton)					
	$u_t = -4.9892u_x (c = 5)$	2560	0.2159 %±0.0%	12800	0.3745 %±0.0%
	$u_t = -1.0004u_x (c = 1)$		0.0375%±0.0%		0.0820%±0.0%
	$u_t = -6.0552uu_x - 1.0312u_{xxx}$	5120	2.0221%±1.558%	25600	2.5931%±1.3601%
Korteweg-de Vries (Double-soliton)					
	$u_t = -5.9853uu_x - 0.9978u_{xxx}$	9000	0.2339%±0.0151%	102912	0.9572%±0.2322%
	$u_t = -5.9250uu_x - 0.9867u_{xxx}$	16600	1.2894%±0.0548%	102912	7.4727%±4.9306%
Burgers’ equation					
	$u_t = -0.9999uu_x + 0.1000u_{xx}$	2000	0.0051%±0.0054%	25856	0.1595%±0.0608%
	$u_t = -1.0010uu_x + 0.1001u_{xx}$	5000	0.0760%±0.0366%	25856	1.9655%±1.0000%
Quantum Harmonic Oscillator					
	$u_t = 0.5001iu_{xx} - 0.9999i\frac{x^2}{2}u$	2000	0.0117%±0.0055%	205312	0.2486%±0.0128%
	$u_t = 0.4996iu_{xx} - 1.0005i\frac{x^2}{2}u$	2000	0.0685%±0.0273%	205312	9.6889%±6.9705%
Nonlinear Schrödinger equation					
	$u_t = 0.4999iu_{xx} + 0.9999i u ^2u$	7000	0.0199%±0.0090%	256512	0.0473%±0.0147%
	$u_t = 0.4997iu_{xx} + 0.9961i u ^2u$	7200	0.2250%±0.2383%	256512	3.0546%±1.2193%

Table 12: Discovering the Schrödinger equation and the nonlinear Schrödinger equation from local dynamics using S³d.

Data 1(QH)	 500	$u_t(x, t) = 0.4732iu_{xx}(x, t) - 0.9451i\frac{x^2}{2}u(x, t)$
Data 2(NLS)	 1800	$u_t(x, t) = 0.5005iu_{xx}(x, t) + 0.9958iu u ^2(x, t)$

5. Conclusion and Discussion

In summary, our proposed S³d algorithm discovers the dynamics underlying a state of traveling-wave convection and many canonical PDEs from measured spatiotemporal data. The merit of the proposed method is its ability to freely construct a model class with candidate functions and automatically select the key ones that reproduce the observed spatiotemporal patterns. Benefiting from sparsity, the inferred PDEs are parsimonious and accurate, enabling interpretability. One of the most important details of applying the S³d method is the approach for estimating the derivatives. In our work, we use the finite difference method, the polynomial interpolation method and the spectral method. In the various examples, we observe that we are additionally able to robustly handle noise in the measurement data with polynomial interpolation and the Pade scheme.

Our proposed S³d method exhibits the ability to extract the governing amplitude equations solely from high-quality experimental data. Although the dynamics we refer to in this work are Eulerian dynamics described by PDEs, S³d, as a general method, is also able to infer ODEs and static functional relations using datasets in (Schmidt and Lipson, 2009; Rudy et al., 2017). S³d unifies results for the discovery of natural laws (Schmidt and Lipson, 2009; Bongard and Lipson, 2007). Future work will focus on two important aspects of extending the method for a wider range of practical applications. First, the current proposed method relies on estimating temporal and spatial derivatives of the measured state variables from data; there are cases that the system variables are not necessarily observable. Second, the current method can reconstruct equations that are linear with respect to the parameters. We are extending this proposed method to cases that are nonlinear with the parameters.

We expect the S³d method to be useful for the modeling of spatiotemporal dynamics from experimental data. This framework, as demonstrated through numerous examples, could potentially accelerate the discovery of new laws and furthermore stimulate physical explanations for the discovered equations, which lead to the discovery of the underlying mechanisms.

Acknowledgments

We wish to thank Dr. Paul Kolodner for useful discussion and allowing us to use the experimental data. All synthetic data and codes used in this manuscript are publicly available on GitHub <https://github.com/HAIRLAB/S3d>.

Appendix A. Polynomial approximation

For noise-contaminated data, polynomial approximation is a better choice to alleviate effects due to noise. With sampled data, $u(x_j, t_i)$, at time, t_i , with $j = 1, \dots, n_x$, we construct an approximation of the q -order derivative $\frac{\partial^q u(x,t)}{\partial x^q}$ by selecting the following sequence of polynomials of degree $p \in \mathbb{N}^+$ with $q < p$:

$$L_p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_p x^p,$$

subject to $L_p(x_j) = u(x_j, t_i)$. For example,

$$\begin{aligned} a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_p x_1^p &= u(x_1, t_i), \\ a_0 + a_1 x_2 + a_2 x_2^2 + \dots + a_p x_2^p &= u(x_2, t_i), \\ \dots \dots & \\ a_0 + a_1 x_{n_x} + a_2 x_{n_x}^2 + \dots + a_p x_{n_x}^p &= u(x_{n_x}, t_i). \end{aligned} \tag{59}$$

Further, we write the above Eq. (59) into matrix form and solve for parameters using a QR factorization. Then, with $u(x, t_i) = L_p(x)$, we compute the q th-order derivative, $\frac{\partial^q u(x,t)}{\partial x^q}$. We demonstrate in our experiment that the error values caused by noise can be removed using polynomial approximation, leading a closer-to-real estimation of the derivative.

Appendix B. Proper Orthogonal Decomposition

Data preprocessing is a key sub-step in S³d that removes noise in our measured datasets. There are many established approaches for rejecting noise in a set of data. For example, one can employ damping or thresholding (Donoho, 1995; Elad and Aharon, 2006; Buades et al., 2005). The Proper Orthogonal Decomposition (POD) (Berkooz et al., 1993; Sirovich, 1987), originally a compression method, is a data processing algorithm that extracts coherent structures with a single temporal frequency from a numerical or experimental data-sequence. In this work, we instead choose to use POD for de-noising our datasets since the extracted structures, called the POD modes, do not contain directions with small variance including small information signals or signals related to noise.

We determine the POD modes from the snapshot matrix, U , by minimizing the Frobenius norm of the difference between U and ΨA ,

$$\min_{\Psi, A} \frac{1}{2} \|U - \Psi A\|_F^2 \quad s.t. \quad \langle \psi_i, \psi_j \rangle = \delta_{ij}$$

where each column of Ψ is a POD mode and A is the coefficient matrix. We solve this minimization by using the singular value decomposition (SVD) on the empirical correlation matrix $C = UU^T$, yielding:

- the $n_x \times p$ matrix comprised of the POD modes: Ψ ;
- the $p \times p$ diagonal matrix with p eigenvalues, $\{\lambda_i\}_{i=1}^p$: Σ ; and
- the $p \times n_t$ matrix whose rows represent the POD coefficients for each POD mode: A ;

Then, we capture the optimal POD modes with a sufficiently high threshold. In our work, we fix our threshold at 0.9999, allowing us to choose integer, $r \leq p$, with

$$\sum_{i=1}^r \lambda_i^2 / \sum_{i=1}^p \lambda_i^2 \geq 0.9999.$$

where $\lambda_{i+1} \leq \lambda_i$. We obtain the new, filtered snapshot matrix, $\tilde{\mathbf{U}}$, with $\tilde{\mathbf{U}} = \Psi A$, with Ψ taking on the first r columns and A taking on the first r rows. In the previously stated minimization problem, we compare the error between the new snapshot matrix, $\tilde{\mathbf{U}}$, and the original snapshot matrix, \mathbf{U} , and choose to use $\tilde{\mathbf{U}} \in \mathbb{R}^{n_x \times n_t}$ instead of \mathbf{U} in the remaining steps.

Appendix C. Reduce Computational Burden

Our first strategy is to identify the dynamics from subsampled data instead of the full data. Let \mathbf{U}_s be a small portion of the total snapshot matrix \mathbf{U} . For example, we can choose the value of solution $u(x, t)$ at all space points and in time interval $[t_k, t_{k+p}]$ such that $\mathbf{U}_s \in \mathbb{R}^{n_x \times p}$. This subsampling is similar to constructing a measurement matrix $C \in \mathbb{R}^{n_x p \times n_x n_t}$ for y in Eq. (5), which results in linear equations similar to Eq. (5)

$$Cy = C\Phi\theta \rightarrow y = \Phi\theta \quad (60)$$

where $y \in \mathbb{R}^{n_x p \times 1}$, $\Phi \in \mathbb{R}^{n_x p \times m}$, and $\theta \in \mathbb{R}^{m \times 1}$ with m representing the total number of candidate terms. Then, we solve the linear equations with the proposed algorithm in the case $mp \ll mn$.

Our next strategy is to reduce the dimension of our problem. Whether $\Phi \in \mathbb{R}^{n_x n_t \times m}$ in Eq. (5) or $\Phi \in \mathbb{R}^{n_x p \times m}$ in Eq. (60), we can use SVD to obtain a set of transform basis. Specifically, using the economy-size SVD of Φ in Eq. (60), we get

$$\Phi = \underbrace{\begin{bmatrix} \psi_1 & \psi_2 & \cdots & \psi_r \end{bmatrix}}_{\text{Transform basis}} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r \end{bmatrix}}_{\text{Truncated singular value}} \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{bmatrix}.$$

Then, we write the optimal low-dimensional representation as

$$\bar{y} = \bar{\Psi}\theta,$$

where $\bar{y} = \Psi^T y \in \mathbb{R}^{r \times 1}$, $\bar{\Psi} = \Psi^T \Phi \in \mathbb{R}^{r \times m}$ with $r \ll n_x p$. The proposed dimensionality reduction method is especially useful in applications where there is a large number of measurements for the sparse identification algorithm.

Depending on the size of the data, either or both the above methods are selected for reducing computational burden from high dimensionality.

Appendix D. Systems with Complex Variables

In certain applications, we consider a standard matrix form, $\tilde{y} = \tilde{\Phi}\tilde{\theta}$, where $\tilde{y}, \tilde{\theta}$ are complex vectors and $\tilde{\Phi}$ is a complex matrix. Given a complex vector \tilde{y} , we consider the following one-to-one mapping to a real vector y :

$$y = \begin{bmatrix} \operatorname{Re} \tilde{y} \\ \operatorname{Im} \tilde{y} \end{bmatrix}.$$

Given complex matrix, $\tilde{\Phi}$, we consider the following construction of real matrix Φ :

$$\Phi = \begin{bmatrix} \operatorname{Re}(\tilde{\Phi}) & -\operatorname{Im}(\tilde{\Phi}) \\ \operatorname{Im}(\tilde{\Phi}) & \operatorname{Re}(\tilde{\Phi}) \end{bmatrix}.$$

With the above mapping and matrix construction, we are able to write that

$$\tilde{y} = \tilde{\Phi}\tilde{\theta} \Leftrightarrow y = \Phi\theta.$$

Once θ is determined, we can obtain $\tilde{\theta}$. Thus, we can relate the discovery of systems with complex variables with the discovery of systems with real variables and are able to use the solution presented for the discovery of systems with real variables.

Appendix E. Proofs of Lemmas

Proof. [Proof of Lemma 3] By the definition of $\widehat{\mathcal{L}}$, and the items (1), (2), (3), (4), $\lim_{k \rightarrow \infty} \widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^{(k)}, \gamma^{(k)})$ exists, and therefore the sequence $\{\widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^{(k)}, \gamma^{(k)})\}_{k=0}^{\infty}$ is bounded.

We first show that

$$\lim_{k \rightarrow \infty} \widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^{(k)}, \gamma^{(k)}) = \widehat{\mathcal{L}}(\theta^{**}, \gamma^{**}; \theta^*, \gamma^*). \quad (61)$$

To this end, it suffices to show that

$$\lim_{k \rightarrow \infty} \widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^*, \gamma^*) = \widehat{\mathcal{L}}(\theta^{**}, \gamma^{**}; \theta^*, \gamma^*). \quad (62)$$

If $(\theta^{**}, \gamma^{**}) \in \operatorname{int}(\Omega)$, then one has from the continuity of $\widehat{\mathcal{L}}(\theta, \gamma; \theta^*, \gamma^*)$ on $\operatorname{int}(\Omega)$ that (62) holds; If $(\theta^{**}, \gamma^{**}) \in \Omega_2$, then by the supplementary definition of $\widehat{\mathcal{L}}(\theta, \gamma; \theta^*, \gamma^*)$ on Ω_2 , (62) holds. Thus, (61) holds.

Next, we show (28). To complete the proof let us assume that $(\theta^{**}, \gamma^{**}) \notin \mathcal{A}(\theta^*, \gamma^*)$ and establish a contradiction. Since $\widehat{\mathcal{L}}(\theta, \gamma; \theta^*, \gamma^*)$ is convex, $\mathcal{A}(\theta^*, \gamma^*)$ is nonempty. Let $(\tilde{\theta}, \tilde{\gamma})$ be any point in $\mathcal{A}(\theta^*, \gamma^*)$. Then, one has

$$\widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta^*, \gamma^*) < \widehat{\mathcal{L}}(\theta^{**}, \gamma^{**}; \theta^*, \gamma^*).$$

Select a sufficiently small positive number ε such that

$$2\varepsilon < \widehat{\mathcal{L}}(\theta^{**}, \gamma^{**}; \theta^*, \gamma^*) - \widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta^*, \gamma^*).$$

By (61), there exists a positive number k_1 such that

$$\widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^{(k)}, \gamma^{(k)}) > \widehat{\mathcal{L}}(\theta^{**}, \gamma^{**}; \theta^*, \gamma^*) - \varepsilon,$$

for any $k > k_1$. By the definition of $\widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta, \gamma)$, we know that $\widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta, \gamma)$ with respect to (θ, γ) is continuous on Ω . Then, there exists a positive number k_2 such that

$$\widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta^{(k)}, \gamma^{(k)}) < \widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta^*, \gamma^*) + \varepsilon,$$

for any $k > k_2$. Putting together these pieces above yields

$$\widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^{(k)}, \gamma^{(k)}) > \widehat{\mathcal{L}}(\tilde{\theta}, \tilde{\gamma}; \theta^{(k)}, \gamma^{(k)}),$$

for any $k > \max\{k_1, k_2\}$. However, we have from the item (5) that

$$\widehat{\mathcal{L}}(\bar{\theta}^{(k)}, \bar{\gamma}^{(k)}; \theta^{(k)}, \gamma^{(k)}) \leq \widehat{\mathcal{L}}(\theta, \gamma; \theta^{(k)}, \gamma^{(k)}),$$

for any $(\theta, \gamma) \in \Omega$. It is a contradiction, and hence, (28) holds. \blacksquare

References

- Mark J. Ablowitz, David J. Kaup, Alan C. Newell, and Harvey Segur. Method for solving the sine-gordon equation. *Physical Review Letters*, 30(25):1262, 1973.
- Jacobo Aguirre and Miguel A.F. Sanjuán. Unpredictable behavior in the duffing oscillator: Wada basins. *Physica D: Nonlinear Phenomena*, 171(1-2):41–51, 2002.
- Jacobo Aguirre, Ricardo L. Viana, and Miguel A.F. Sanjuán. Fractal structures in nonlinear dynamics. *Reviews of Modern Physics*, 81(1):333, 2009.
- Kazuyuki Aihara, T. Takabe, and Masashi Toyoda. Chaotic neural networks. *Physics Letters A*, 144(6–7):333–340, 1990.
- Igor S. Aranson and Lorenz Kramer. The world of the complex ginzburg-landau equation. *Reviews of Modern Physics*, 74(1):99, 2002.
- Markus Bär, Rainer Hegger, and Holger Kantz. Fitting partial differential equations to space-time dynamics. *Physical Review E*, 59(1):337, 1999.
- A. Barone, F. Esposito, C.J. Magee, and A.C. Scott. Theory and applications of the sine-gordon equation. *La Rivista del Nuovo Cimento*, 1(2):227–267, 1971.
- Atilim Gunes Baydin, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research*, 18:1–43, 2018.
- Gal Berkooz, Philip Holmes, and John L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25(1):539–575, 1993.

- Josh Bongard and Hod Lipson. Automated reverse engineering of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 104(24):9943–9948, 2007.
- Leo Breiman and Jerome H Friedman. Estimating optimal transformations for multiple regression and correlation. *Journal of the American statistical Association*, 80(391):580–598, 1985.
- Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, page 201517384, 2016.
- Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.
- Emmanuel J Candes, Michael B Wakin, and Stephen P Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier analysis and applications*, 14(5-6):877–905, 2008.
- Young Hwan Chang and Claire Tomlin. Data-driven graph reconstruction using compressive sensing. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 1035–1040. IEEE, 2012.
- Patrick L. Combettes and Jean-Christophe Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.
- Peter Constantin, Ciprian Foias, Basil Nicolaenko, and Roger Temam. *Integral manifolds and inertial manifolds for dissipative partial differential equations*, volume 70. Springer Science & Business Media, 2012.
- Mark C. Cross. Structure of nonlinear traveling-wave states in finite geometries. *Physical Review A*, 38(7):3593, 1988.
- Mark C. Cross and Pierre C. Hohenberg. Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3):851, 1993.
- Roger K. Dodd, Hedley C. Morris, J.C. Eilbeck, and J.D. Gibbon. Soliton and nonlinear wave equations. *London and New York, Academic Press, 1982, 640 p.*, 1982.
- David L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995.
- Mark Edwards and K. Burnett. Numerical solution of the nonlinear schrödinger equation for small samples of trapped neutral atoms. *Physical Review A*, 51(2):1382, 1995.
- Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image processing*, 15(12):3736–3745, 2006.

- Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360, 2001.
- J. Doyne Farmer and John J. Sidorowich. Predicting chaotic time series. *Physical Review Letters*, 59(8):845, 1987.
- Anita C Faul and Michael E Tipping. Analysis of sparse bayesian learning. In *Advances in neural information processing systems*, pages 383–389, 2002.
- Mário Figueiredo. Adaptive sparseness using jeffreys prior. In *Advances in neural information processing systems*, pages 697–704, 2002.
- Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, 1961.
- Ciprian Foias, George R. Sell, and Roger Temam. Inertial manifolds for nonlinear evolutionary equations. *Journal of Differential Equations*, 73(2):309–353, 1988.
- Dimitrios Giannakis. Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Applied and Computational Harmonic Analysis*, 2017.
- Celso Grebogi, Edward Ott, and James A. Yorke. Chaotic attractors in crisis. *Physical Review Letters*, 48(22):1507, 1982.
- Lingzhong Guo and Stephen A Billings. Identification of partial differential equation models for continuous spatio-temporal dynamical systems. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 53(8):657–661, 2006.
- Dax A. Hoffman, Jeffrey C. Magee, Costa M. Colbert, and Daniel Johnston. K+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. *Nature*, 387(6636):869, 1997.
- James M. Hyman and Basil Nicolaenko. The kuramoto-sivashinsky equation: a bridge between pde’s and dynamical systems. *Physica D: Nonlinear Phenomena*, 18(1-3):113–126, 1986.
- Ioannis G. Kevrekidis, C. William Gear, James M. Hyman, Panagiotis G. Kevrekidid, Olof Runborg, Constantinos Theodoropoulos, et al. Equation-free, coarse-grained multiscale computation: Enabling mocroscopic simulators to perform system-level analysis. *Communications in Mathematical Sciences*, 1(4):715–762, 2003.
- Omid Khanmohamadi and Daolin Xu. Spatiotemporal system identification on nonperiodic domains using chebyshev spectral operators and system reduction algorithms. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(3):033117, 2009.
- Paul Kolodner. Counterpropagating quasilinear wave packets in binary-fluid convection. *Physical Review Letters*, 69(17):2519, 1992.

Paul Kolodner, Said Slimani, Nadine Aubry, and Ricardo Lima. Characterization of dispersive chaos and related states of binary-fluid convection. *Physica D: Nonlinear Phenomena*, 85(1-2):165–224, 1995.

Diederik Johannes Korteweg and Gustav De Vries. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 5(39):422–443, 1895.

Ying-Cheng Lai and Tamás Tél. *Transient chaos: complex dynamics on finite time scales*, volume 173. Springer Science & Business Media, 2011.

Randall J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*, volume 98. SIAM, 2007.

Jichun Li and Yi-Tung Chen. *Computational partial differential equations using MATLAB*. Chapman and Hall/CRC, 2008.

Anders Lindquist and Giorgio Picci. *Linear stochastic systems: A geometric approach to modeling, estimation and identification*, volume 1. Springer, 2015.

Pierre-Louis Lions and Bertrand Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.

Lennart Ljung. *System Identification: Theory for the User*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1986. ISBN 0-138-81640-9.

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):52–63, 2016.

Igor Mezić. Analysis of fluid flows via spectral properties of the koopman operator. *Annual Review of Fluid Mechanics*, 45:357–378, 2013.

R.C. Mittal and Geeta Arora. Efficient numerical solution of fisher’s equation by using b-spline method. *International Journal of Computer Mathematics*, 87(13):3039–3051, 2010.

TG Müller and Jens Timmer. Parameter identification techniques for partial differential equations. *International Journal of Bifurcation and Chaos*, 14(06):2053–2060, 2004.

Jinichi Nagumo, Suguru Arimoto, and Shuji Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.

Alan C. Newell. Nonlinear wave motion. *Providence, RI, American Mathematical Society (Lectures in Applied Mathematics)*, 15:237, 1974.

Turgut Özış and Canan Köroğlu. A novel approach for solving the fisher equation using exp-function method. *Physics Letters A*, 372(21):3836–3840, 2008.

Jason Palmer, Kenneth Kreutz-Delgado, Bhaskar D Rao, and David P Wipf. Variational em algorithms for non-gaussian latent variable models. In *Advances in neural information processing systems*, pages 1059–1066, 2006.

Wei Pan, Ye Yuan, Jorge Gonçalves, and Guy-Bart Stan. Reconstruction of arbitrary biochemical reaction networks: A compressive sensing approach. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 2334–2339. IEEE, 2012.

Wei Pan, Ye Yuan, Jorge Gonçalves, and Guy-Bart Stan. A sparse bayesian approach to the identification of nonlinear state-space systems. *IEEE Transactions on Automatic Control*, 61(1):182–187, 2016.

Wei Pan, Ye Yuan, Lennart Ljung, Jorge Gonçalves, and Guy-Bart Stan. Identification of nonlinear state-space systems from heterogeneous datasets. *IEEE Transactions on Control of Network Systems*, 5(2):737–747, 2018.

Maziar Raissi. Deep hidden physics models: Deep learning of nonlinear partial differential equations. *Journal of Machine Learning Research*, 473(19):1–24, 2018.

Maziar Raissi and George Em Karniadakis. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 357:125–141, 2017.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (part ii): data-driven discovery of nonlinear partial differential equations. *arXiv preprint arXiv:1711.10566*, 2017.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Multistep neural networks for data-driven discovery of nonlinear dynamical systems. *arXiv preprint arXiv:1801.01236*, 2018.

Bhaskar D Rao, Kjersti Engan, Shane F Cotter, Jason Palmer, and Kenneth Kreutz-Delgado. Subset selection in noise based on diversity measure minimization. *IEEE transactions on Signal processing*, 51(3):760–770, 2003.

Samuel H. Rudy, Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Data-driven discovery of partial differential equations. *Science Advances*, 3(4):e1602614, 2017.

Hayden Schaeffer. Learning partial differential equations via data discovery and sparse optimization. *Proceedings of the Royal Society A*, 473(2197):20160446, 2017.

Michael Schmidt and Hod Lipson. Distilling free-form natural laws from experimental data. *Science*, 324(5923):81–85, 2009.

Erwin Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical Review*, 28(6):1049, 1926.

Zhesi Shen, Wen-Xu Wang, Ying Fan, Zengru Di, and Ying-Cheng Lai. Reconstructing propagation networks with natural diversity and identifying hidden sources. *Nature Communications*, 5:4323, 2014.

- Chang Shu and Bryan E. Richards. Application of generalized differential quadrature to solve two-dimensional incompressible navier-stokes equations. *International Journal for Numerical Methods in Fluids*, 15(7):791–798, 1992.
- Lawrence Sirovich. Turbulence and the dynamics of coherent structures. i. coherent structures. *Quarterly of Applied Mathematics*, 45(3):561–571, 1987.
- Robert D. Skeel and Martin Berzins. A method for the spatial discretization of parabolic equations in one space variable. *SIAM Journal on Scientific and Statistical Computing*, 11(1):1–32, 1990.
- Gilbert Strang. *Computational science and engineering*, volume 791. Wellesley: Wellesley-Cambridge Press, 2007.
- Robert Tibshirani, Martin Wainwright, and Trevor Hastie. *Statistical learning with sparsity: the lasso and generalizations*. Chapman and Hall/CRC, 2015.
- Michael E Tipping. The relevance vector machine. In *Advances in neural information processing systems*, pages 652–658, 2000.
- Michael E Tipping. Sparse bayesian learning and the relevance vector machine. *Journal of machine learning research*, 1(Jun):211–244, 2001.
- Lloyd N. Trefethen. *Spectral methods in MATLAB*, volume 10. SIAM, 2000.
- Jonathan H. Tu, Clarence W. Rowley, Dirk M. Luchtenburg, Steven L. Brunton, and J. Nathan Kutz. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(2):391–421, 2014.
- H Voss, MJ Bünnner, and Markus Abel. Identification of continuous, spatiotemporal systems. *Physical Review E*, 57(3):2820, 1998.
- Henning U. Voss, Paul Kolodner, Markus Abel, and Jürgen Kurths. Amplitude equations from spatiotemporal binary-fluid convection data. *Physical Review Letters*, 83(17):3422, 1999.
- Wen-Xu Wang, Rui Yang, Ying-Cheng Lai, Vassilios Kovanis, and Celso Grebogi. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Physical Review Letters*, 106(15):154101, 2011.
- Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data–driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.
- David Wipf and Srikantan Nagarajan. Iterative reweighted ℓ_1 and ℓ_2 methods for finding sparse solutions. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):317–329, 2010.
- David P Wipf and Srikantan S Nagarajan. A new view of automatic relevance determination. In *Advances in neural information processing systems*, pages 1625–1632, 2008.

- David P Wipf and Bhaskar D Rao. Sparse bayesian learning for basis selection. *IEEE Transactions on Signal processing*, 52(8):2153–2164, 2004.
- David P Wipf, Bhaskar D Rao, and Srikantan Nagarajan. Latent variable bayesian models for promoting sparsity. *IEEE Transactions on Information Theory*, 57(9):6236–6255, 2011.
- Yi Wu and David P Wipf. Dual-space analysis of the sparse linear model. In *Advances in Neural Information Processing Systems*, pages 1745–1753, 2012.
- Daolin Xu and Omid Khanmohamadi. Spatiotemporal system reconstruction using fourier spectral operators and structure selection techniques. *Chaos*, 18(4):043122, 2008.
- Linan Zhang and Hayden Schaeffer. On the convergence of the sindy algorithm. *arXiv preprint arXiv:1805.06445*, 2018.
- Sheng Zhang and Guang Lin. Robust data-driven discovery of governing physical laws with error bars. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2217):20180305, 2018.