Engineering 17 (2022) 244–252

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Engineering Management—Article

Data-Driven Discovery of Stochastic Differential Equations

Yasen Wang [a,b], Huazhen Fang [c], Junyang Jin [d], Guijun Ma [a,b], Xin He [a], Xing Dai [a,d], Zuogong Yue [e], Cheng Cheng [e], Hai-Tao Zhang [b,e], Donglin Pu [d], Dongrui Wu [e], Ye Yuan [a,b,e,*], Jorge Gonçalves [e,f,g], Jürgen Kurths [h,i], Han Ding [a,b,d]

[a] School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[b] State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
[c] Department of Mechanical Engineering, University of Kansas, Lawrence, KS 66045, USA
[d] HUST-Wuxi Research Institute, Wuxi 214174, China
[e] Key Laboratory of Image Processing and Intelligent Control, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China
[f] Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK
[g] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux 4367, Luxembourg
[h] Department of Physics, Humboldt University of Berlin, Berlin 12489, Germany
[i] Department of Complexity Science, Potsdam Institute for Climate Impact Research, Potsdam 14473, Germany

ARTICLE INFO

Article history:
Received 1 October 2021
Revised 6 February 2022
Accepted 15 February 2022
Available online 23 March 2022

Keywords:
Data-driven method
System identification
Sparse Bayesian learning
Stochastic differential equations
Random phenomena

ABSTRACT

Stochastic differential equations (SDEs) are mathematical models that are widely used to describe complex processes or phenomena perturbed by random noise from different sources. The identification of SDEs governing a system is often a challenge because of the inherent strong stochasticity of data and the complexity of the system's dynamics. The practical utility of existing parametric approaches for identifying SDEs is usually limited by insufficient data resources. This study presents a novel framework for identifying SDEs by leveraging the sparse Bayesian learning (SBL) technique to search for a parsimonious, yet physically necessary representation from the space of candidate basis functions. More importantly, we use the analytical tractability of SBL to develop an efficient way to formulate the linear regression problem for the discovery of SDEs that requires considerably less time-series data. The effectiveness of the proposed framework is demonstrated using real data on stock and oil prices, bearing variation, and wind speed, as well as simulated data on well-known stochastic dynamical systems, including the generalized Wiener process and Langevin equation. This framework aims to assist specialists in extracting stochastic mathematical models from random phenomena in the natural sciences, economics, and engineering fields for analysis, prediction, and decision making.

1. Introduction

Nature, industry, and human society are teeming with phenomena, processes, and systems that evolve with the effects of random noise. Examples of such systems include the motion of Brownian particles in a fluid [1], evolution of tumors driven by tumor-immune interactions [2], price of stocks, and movement of winds. Stochastic differential equations (SDEs) are a powerful mathematical approach for modeling and analyzing systems affected by random noise. The study of SDEs originates in one of Einstein's *annus mirabilis* papers [1], which theorizes the fluctuation of Brownian particles in a thermal bath. Since then, the application of SDEs has spread across numerous scientific and engineering fields [3–7]. Traditionally, SDEs have been employed to model random phenomena based on the statistical properties of data or the experience of specialists. However, in many cases, the statistical properties of data may not resemble those of well-understood SDEs, and are thus less helpful for the discovery of underlying SDEs. Additionally, the governing SDEs of random phenomena are typically unknown or elusive. Therefore, it is crucial to develop a data-driven method for identifying the underlying SDEs of random phenomena. However, the complex behavior and strong stochasticity often present in SDE-governed systems makes the accurate identification of SDEs particularly challenging.

* Corresponding author.
E-mail address: yye@hust.edu.cn (Y. Yuan).

https://doi.org/10.1016/j.eng.2022.02.007

The key to determining an SDE is the identification of its drift and diffusion terms. The growing body of work on this subject can be roughly divided into two categories. The first lies in non-parametric identification, which seeks to build a model-free input–output mapping for drift and diffusion terms from time-series data. The Kramers–Moyal average introduces histogram-based regression (HBR) to formulate a suitable mapping for the drift and diffusion terms [8]. However, this method requires a large amount of data even for one-dimensional SDEs, and the required amount of data increases exponentially with the SDE's dimensionality. Realistically, this data requirement aggravates the burdens of sensor cost, data storage, and computational resources, which may be difficult to meet in some cases. To improve identification accuracy with a limited amount of data, kernel-based regression (KBR) [9] and polynomial-based method [10] have been proposed as more efficient mapping methods for one-dimensional SDEs. In addition, non-parametric Bayesian estimation provides another mapping approach to mitigate the high demand for time-series data [11–14]. Markov chain Monte Carlo (MCMC) methods [11,12] and Gaussian process regression (GPR) [13] have recently been proposed for the identification of SDEs. Furthermore, sparse GPRs have been developed to learn SDEs with high computational efficiency [13,14]. While the aforementioned methods can predict drift and diffusion terms adequately, they cannot model stochastic dynamical systems because they only provide black-box representations.

In contrast, parametric approaches focus on identifying the model structures of the drift and diffusion terms of an SDE, and are more advantageous in revealing the underlying physical laws of stochastic dynamical systems. For scalar homogeneous SDEs with known model structures, the KBR and least-squares methods are combined to estimate parameters [15]. For non-linear dynamical systems, a framework called sparse identification of non-linear dynamics (SINDy) has been applied to determine governing equations under the assumption that the model structure is sparse in the space of possible basis functions [16]. Based on HBR and SINDy, sparse learning approaches have been proposed to identify drift and diffusion terms [17,18]. However, these approaches exhibit the same drawbacks as HBR because they initially rely on HBR to estimate the drift and diffusion terms.

An emerging technique for identifying model structures in the fields of system identification and signal processing is sparse Bayesian learning (SBL) [19–25], which aims to find a parsimonious representation from basis functions based on input–output data by striking a balance between model complexity and accuracy. The present study leverages this technique to discover the underlying SDEs of stochastic dynamical systems using relatively limited time-series data. The implementation of the proposed algorithm can be summarized in terms of the following two stages. First, theoretical expressions for the drift and diffusion terms are derived by discretizing SDEs with the Euler–Maruyama method. Then, the discovery problem of SDEs is cast into an input–output regression problem for the drift and diffusion terms based on the central limit theorem. Although the binning operation can be used to estimate the values of the drift and diffusion terms at the selected point to formulate the input–output regression problem, this operation is data-hungry and suffers from the curse of dimensionality. Owing to the analytical tractability of SBL, a more efficient method that can be implemented using relatively limited quantities of time-series data in practice is proposed herein. The enhanced capabilities and robustness of the proposed algorithm, named Bayesian identification of SDEs (BISDEs), were demonstrated on well-known SDEs against those of the state-of-the-art method. Additionally, the proposed BISDE algorithm was validated over a wide range of simulated and real-world systems.

2. Methods

2.1. Mathematical expressions for the drift and diffusion terms

In this study, we consider the n-dimensional SDEs in a general form as follows:

$$\mathrm{d}\boldsymbol{x}(t) = \boldsymbol{f}(\boldsymbol{x}(t))\mathrm{d}t + \boldsymbol{G}(\boldsymbol{x}(t))^{\frac{1}{2}}\mathrm{d}\boldsymbol{W}(t) \tag{1}$$

where $\boldsymbol{x}(t)\in\mathbb{R}^n$ is the state vector at time t, $\boldsymbol{f}(\boldsymbol{x}(t))\in\mathbb{R}^n$ is the state-dependent drift vector, $\boldsymbol{G}(\boldsymbol{x}(t))\in\mathbb{R}^{n\times n}$ is the positive-definite diffusion matrix, and $\boldsymbol{W}(t)$ is an n-dimensional standard Brownian motion or Wiener process. It is assumed that each entry for $\boldsymbol{f}(\boldsymbol{x}(t))$ and $\boldsymbol{G}(\boldsymbol{x}(t))$ is a continuous function. However, the exact model structures for $\boldsymbol{f}(\boldsymbol{x}(t))$ and $\boldsymbol{G}(\boldsymbol{x}(t))$ are unknown. The objective of this study is to identify them from the relatively limited time-series data of $\boldsymbol{x}(t)$.

Because analytical solutions for SDEs are generally unavailable, we resort to the Euler–Maruyama discretization. To ensure the existence and uniqueness of the solution of the SDE in Eq. (1) and the feasibility of the Euler–Maruyama method, we assume that the drift and diffusion terms satisfy the local Lipschitz and Khasminskii-type conditions (Section S1 in Appendix A). Furthermore, under these conditions, the numerical solution based on the Euler–Maruyama method converges to the exact solution in probability [26,27].

Applying the Euler–Maruyama method to the SDE in Eq. (1), we have

$$\widehat{\boldsymbol{x}}((k+1)\Delta t) - \widehat{\boldsymbol{x}}(k\Delta t) = \boldsymbol{f}\big(\widehat{\boldsymbol{x}}(k\Delta t)\big)\Delta t + \boldsymbol{G}\big(\widehat{\boldsymbol{x}}(k\Delta t)\big)^{\frac{1}{2}}\sqrt{\Delta t}\boldsymbol{\epsilon}_k \tag{2}$$

where k is the discretization time index, Δt is the discretization time step, $\widehat{\boldsymbol{x}}(k\Delta t)$ is the numerical solution of $\boldsymbol{x}(k\Delta t)$, and $\boldsymbol{\epsilon}_k \sim \mathcal{N}(0, \boldsymbol{I}_n)$, a normal distribution with mean 0 and covariance matrix $\boldsymbol{I}_n$ with $\boldsymbol{I}_n$ as an $n\times n$ identity matrix. It is more convenient to perform the analysis by considering continuous-time approximations. Hence, for any $t\in(k\Delta t,(k+1)\Delta t)$, we set

$$\widehat{\boldsymbol{x}}(t) = \widehat{\boldsymbol{x}}(k\Delta t) \tag{3}$$

Given $\widehat{\boldsymbol{x}}(k\Delta t)$, it is evident that $\widehat{\boldsymbol{x}}((k+1)\Delta t)$ in Eq. (2) satisfies the Gaussian distribution. This implies that, for any $t\geq 0$,

$$\widehat{\boldsymbol{x}}(t+\Delta t)\mid\widehat{\boldsymbol{x}}(t) \sim \mathcal{N}\big(\widehat{\boldsymbol{x}}(t) + \boldsymbol{f}\big(\widehat{\boldsymbol{x}}(t)\big)\Delta t, \boldsymbol{G}\big(\widehat{\boldsymbol{x}}(t)\big)\Delta t\big) \tag{4}$$

From Eq. (4), we can derive expressions for the drift and diffusion terms. First, for any point $\xi\in\mathbb{R}^n$, the drift $\boldsymbol{f}(\xi)$ can be obtained by the conditional expectation as follows:

$$\boldsymbol{f}(\xi) = \frac{1}{\Delta t}E\big[\widehat{\boldsymbol{x}}(t+\Delta t) - \widehat{\boldsymbol{x}}(t)\mid\widehat{\boldsymbol{x}}(t)=\xi\big] \tag{5}$$

where E denotes the expectation operator.

Similar to the case of drift, the diffusion $\boldsymbol{G}(\xi)$ is given by the conditional covariance as follows:

$$\boldsymbol{G}(\xi) = \frac{1}{\Delta t}E\Big[\big(\widehat{\boldsymbol{x}}(t+\Delta t) - \widehat{\boldsymbol{x}}(t) - \boldsymbol{f}(\widehat{\boldsymbol{x}}(t))\Delta t\big)\big(\widehat{\boldsymbol{x}}(t+\Delta t) - \widehat{\boldsymbol{x}}(t) - \boldsymbol{f}(\widehat{\boldsymbol{x}}(t))\Delta t\big)^{\mathrm{T}} \mid \widehat{\boldsymbol{x}}(\mathrm{t}) = \xi\Big] \tag{6}$$

where T denotes the transpose operator.

While the drift and diffusion terms are estimated based on the numerical solution $\widehat{\boldsymbol{x}}(t)$, we can prove that when $\Delta t\to 0$, all entries of $\boldsymbol{f}(\widehat{\boldsymbol{x}}(t))$ and $\boldsymbol{G}(\widehat{\boldsymbol{x}}(t))$ converge to the corresponding entries of $\boldsymbol{f}(\boldsymbol{x}(t))$ and $\boldsymbol{G}(\boldsymbol{x}(t))$ in probability, respectively. The following proposition summarizes the result, in which the subscript for the entry index is omitted for notational simplicity.

Proposition 1. Suppose that the SDE in Eq. (1) satisfies the local Lipschitz and Khasminskii-type conditions, and each entry of the drift and diffusion terms is a continuous function. Then, for any $\tau>0$,

$$\lim_{\Delta t\to 0}\left(\sup_{0\le t\le\tau}\left|\boldsymbol{f}(\widehat{\boldsymbol{x}}(t))-\boldsymbol{f}(\boldsymbol{x}(t))\right|\right)=0 \tag{7}$$

in probability, and

$$\lim_{\Delta t\to 0}\left(\sup_{0\le t\le\tau}\left|\boldsymbol{G}(\widehat{\boldsymbol{x}}(t))-\boldsymbol{G}(\boldsymbol{x}(t))\right|\right)=0 \tag{8}$$

in probability.

Proof: refer to Section S2 in Appendix A.

According to the central limit theorem, $\boldsymbol{f}(\xi)$ can be computed approximately from the collected time series $\{\widehat{\boldsymbol{x}}(t_i)\}_{i=1}^m$ based on Eq. (5). However, computing $\boldsymbol{G}(\xi)$ is more complicated. From Eq. (6), we notice that computing $\boldsymbol{G}(\xi)$ requires not only the data but also $\boldsymbol{f}(\xi)$. One viable method is to use the identified drift to estimate the value of $\boldsymbol{f}(\xi)$.

2.2. Inferring the drift term

After collecting the time series $\{\widehat{\boldsymbol{x}}(t_i)\}_{i=1}^m$, $\boldsymbol{f}(\xi)$ can be estimated according to the central limit theorem as follows:

$$\frac{\sum_{s=1}^{K}\left[\widehat{\boldsymbol{x}}(t_{j_s+1})-\widehat{\boldsymbol{x}}(t_{j_s})\ \middle|\ \widehat{\boldsymbol{x}}(t_{j_s})=\xi\right]}{K\Delta t}\approx\boldsymbol{f}(\xi)+\boldsymbol{\varepsilon}_1 \tag{9}$$

where K is the number of $\widehat{\boldsymbol{x}}(t_{j_s})$ equal to ξ with $\widehat{\boldsymbol{x}}(t_{j_s})\in\{\widehat{\boldsymbol{x}}(t_i)\}_{i=1}^m$, and $\boldsymbol{\varepsilon}_1$ is a vector in which each entry is Gaussian distributed with mean zero and variance proportional to $1/K$. This equation bridges the empirical estimation with the SBL, as the noise can be modeled as Gaussian noise. Thus, the drift identification problem can be transformed into an input–output regression problem. This equation also implies that drift entry can be identified independently. Without loss of generality, suppose that $\xi(i)\in\mathbb{R}^n$, $\boldsymbol{f}_r(\xi(i))\in\mathbb{R}$, $i=1,2,\cdots,N$ denote the input–output data of the rth drift entry, and $\boldsymbol{f}_r(\xi(i))$ is estimated using K_i time-series data points from Eq. (9). Let

$$\begin{aligned}\boldsymbol{X}&=\begin{bmatrix}|&|&|&|\\ \xi(1)&\xi(2)&\cdots&\xi(N)\\ |&|&|&|\end{bmatrix}^{\mathrm{T}},\\ \boldsymbol{Y}&=[\boldsymbol{f}_r(\xi(1))\ \boldsymbol{f}_r(\xi(2))\ \cdots\ \boldsymbol{f}_r(\xi(N))]^{\mathrm{T}}\end{aligned} \tag{10}$$

We assume that the rth drift entry is a linear combination of some basis functions that belong to a library of candidate functions. It is usually desirable to make the library sufficiently large to allow for a thorough search and determination of the underlying model structure. For many practical systems, the drift is sparse in the space spanned by the basis functions because it comprises only a few terms. In addition, any available prior knowledge of the considered stochastic dynamical system can guide us in selecting the basis functions more efficiently. An example library can consist of constant, linear, and polynomial terms as follows:

$$\boldsymbol{\Phi}=\left[1\ \boldsymbol{X}\ \boldsymbol{X}^2\ \boldsymbol{X}^3\cdots\right] \tag{11}$$

where $\boldsymbol{X}^2,\boldsymbol{X}^3$, and so on, denote higher polynomials. For example, each column of $\boldsymbol{X}^2$ can be specified as the element-wise product of $\xi(i)$ and $\xi(j)$ (where i can be equal to j). The remaining problem lies in estimating the basis functions' weights. Precise identification of the sparse weight vector is crucial for identifying the model structure of the rth drift entry.

To estimate the weight vector, we can approximately solve the following regression equation derived from Eq. (9):

$$\boldsymbol{Y}=\boldsymbol{\Phi}\boldsymbol{\theta}+\boldsymbol{\varepsilon} \tag{12}$$

where $\boldsymbol{\Phi}\in\mathbb{R}^{N\times M}$ is the constructed library matrix, M is the number of basis functions, and $\boldsymbol{\theta}=[\theta_1\ \theta_2\ \cdots\ \theta_M]^{\mathrm{T}}$ is the weight vector. The noise vector $\boldsymbol{\varepsilon}$ is assumed to follow a Gaussian distribution $\mathcal{N}(0,\boldsymbol{\Psi})$, where $\boldsymbol{\Psi}$ is a diagonal matrix with the ith element being σ^2/K_i, and σ^2 is a scale parameter of variance. First, given that the model structure of the drift is sparse in the span of the selected basis functions, we impose a sparsity-promoting Gaussian prior with mean zero and variance γ_i on weight θ_i. Hence, $\boldsymbol{\theta}$ is denoted as a random vector with the initial probability distribution $p(\boldsymbol{\theta};\boldsymbol{\gamma})$, where $\boldsymbol{\gamma}=[\gamma_1\ \gamma_2\ \cdots\ \gamma_M]^{\mathrm{T}}$. Based on the maximum a posteriori (MAP) principle, the mean of the posterior distribution of $\boldsymbol{\theta}$,

$$p(\boldsymbol{\theta}\mid\boldsymbol{Y};\boldsymbol{\Psi},\boldsymbol{\gamma})\propto p(\boldsymbol{Y}\mid\boldsymbol{\theta};\boldsymbol{\Psi})p(\boldsymbol{\theta};\boldsymbol{\gamma}) \tag{13}$$

is selected as the estimate of $\boldsymbol{\theta}$. Here, $p(\boldsymbol{Y}\mid\boldsymbol{\theta};\boldsymbol{\Psi})$ is the likelihood function arising from Eq. (12). However, the estimation of $\boldsymbol{\theta}$ involves $\boldsymbol{\gamma}$. In order to set a reasonable value of $\boldsymbol{\gamma}$, we can maximize the type-II likelihood function $p(\boldsymbol{Y};\boldsymbol{\Psi},\boldsymbol{\gamma})=\int p(\boldsymbol{Y}\mid\boldsymbol{\theta};\boldsymbol{\Psi})p(\boldsymbol{\theta};\boldsymbol{\gamma})\mathrm{d}\boldsymbol{\theta}$. Consequently, after obtaining its optimal value, denoted as $\boldsymbol{\gamma}^*$, we have the following:

$$\boldsymbol{\theta}=\left(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{\Phi}+\boldsymbol{\Gamma}^{*-1}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{Y} \tag{14}$$

where $\boldsymbol{\Gamma}^*=\mathrm{diag}(\boldsymbol{\gamma}^*)$.

For a system of interest, it may be difficult to obtain many measurements $\widehat{\boldsymbol{x}}(t_i)$ being equal to ξ. A useful trick to approximate the conditional expectation $\boldsymbol{f}(\xi)$ is to treat any data point $\widehat{\boldsymbol{x}}(t_{j_s})$ that falls into a small neighborhood of ξ as ξ; or, specifically,

$$\frac{\sum_{s=1}^{K}\left[\widehat{\boldsymbol{x}}(t_{j_s+1})-\widehat{\boldsymbol{x}}(t_{j_s})\ \middle|\ \widehat{\boldsymbol{x}}(t_{j_s})\in(\xi-\boldsymbol{\delta},\xi+\boldsymbol{\delta})\right]}{K\Delta t}\approx\boldsymbol{f}(\xi)+\boldsymbol{\varepsilon}_1 \tag{15}$$

where $\boldsymbol{\delta}=[\delta_1\ \delta_2\ \cdots\ \delta_n]^{\mathrm{T}}$ is the hyperparameter vector that is used to control the neighborhood size. This technique, known as the binning operation, has proven to be effective in Refs. [8,17,18]. In this manner, the data points are divided into $\prod_{j=1}^{n}(\{\max[\widehat{\boldsymbol{x}}_j(t_i)]-\min[\widehat{\boldsymbol{x}}_j(t_i)]\}/2\delta_j)$ bins for an n-dimensional SDE, where $\widehat{\boldsymbol{x}}_j(t_i)$ is the jth entry of $\widehat{\boldsymbol{x}}(t_i)$. Consequently, this approach suffers from the curse of dimensionality as an increase in an SDE's dimensionality leads to an exponentially increasing number of bins and data, if one intends to preserve overall approximation accuracy. It is also difficult to balance the number of bins and the accuracy of approximation. To address the above issues, we developed a more efficient strategy for formulating the regression equation to identify the model structure of drift. First, we employed an equivalent realization of the regression equation in Eq. (12) to derive an identical weight vector.

Theorem 1. The weight vector identified from

$$\widetilde{\boldsymbol{Y}}=\widetilde{\boldsymbol{\Phi}}\,\widetilde{\boldsymbol{\theta}}+\widetilde{\boldsymbol{\varepsilon}} \tag{16}$$

that is, $\widetilde{\boldsymbol{\theta}}$ is identical to that identified in Eq. (12), that is, $\boldsymbol{\theta}$, where

$$\widetilde{\boldsymbol{Y}}=\left[\widetilde{\boldsymbol{f}}_r(\xi(1))\ \widetilde{\boldsymbol{f}}_r(\xi(2))\ \cdots\ \widetilde{\boldsymbol{f}}_r(\xi(N))\right]^{\mathrm{T}} \tag{17}$$

$$\begin{aligned}\widetilde{\boldsymbol{f}}_r(\xi(j))=\Bigg[&\frac{1}{\Delta t}\left(\widehat{\boldsymbol{x}}_r(t_{j_1+1})-\widehat{\boldsymbol{x}}_r(t_{j_1})\right)\mid\widehat{\boldsymbol{x}}(t_{j_1})=\xi(j)\ ,\cdots,\\ &\frac{1}{\Delta t}\left(\widehat{\boldsymbol{x}}_r\left(t_{j_{K_j}+1}\right)-\widehat{\boldsymbol{x}}_r\left(t_{j_{K_j}}\right)\right)\ \middle|\ \widehat{\boldsymbol{x}}\left(t_{j_{K_j}}\right)=\xi(j)\Bigg]\end{aligned} \tag{18}$$

$$\widetilde{\boldsymbol{\Phi}}=\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}}\otimes\boldsymbol{\Phi}_1\\ \vdots\\ \boldsymbol{a}_N^{\mathrm{T}}\otimes\boldsymbol{\Phi}_N\end{bmatrix},\ \boldsymbol{a}_i=\underbrace{[1\ \cdots\ 1]}_{K_i} \tag{19}$$

where $\otimes$ is the Kronecker product, $\widehat{\boldsymbol{x}}_r$ is the rth entry of $\widehat{\boldsymbol{x}}$, $\boldsymbol{\Phi}_i$ is the ith row of $\boldsymbol{\Phi}$, and $\widetilde{\boldsymbol{\varepsilon}}$ follows the Gaussian distribution $\mathcal{N}(0,\sigma^2\boldsymbol{I})$ with $\boldsymbol{I}$ as a $\sum_{i=1}^{N}K_i\times\sum_{i=1}^{N}K_i$ identity matrix.

Proof: refer to Section S2 in Appendix A.

Remark 1. For any permutation matrix $\boldsymbol{P}$, we have $\boldsymbol{P}\widetilde{\boldsymbol{Y}} = \boldsymbol{P}\widetilde{\boldsymbol{\Phi}}\,\widetilde{\boldsymbol{\theta}} + \boldsymbol{P}\widetilde{\boldsymbol{\varepsilon}}$. Hence, the elements of $\widetilde{\boldsymbol{Y}}$ can be rearranged in the order of time sequence.

Theorem 1 shows a constructive method for identifying $\boldsymbol{\theta}$ in Eq. (12) by considering the alternate regression equation in Eq. (16). Its advantages are significant. First, Eq. (16) implies that the "derivative" can be used as the output, obviating the need for a binning operation to reduce the amount of required data and avoid the curse of dimensionality in practice. Second, unlike the binning operation, the value of $\widehat{\boldsymbol{x}}(t_{j_s}) \in (\boldsymbol{\xi} - \boldsymbol{\delta}, \boldsymbol{\xi} + \boldsymbol{\delta})$ is kept instead of being replaced with $\boldsymbol{\xi}$ when we formulate the library matrix and output vector, which avoids the consequent approximation error. Finally, our experiments consistently suggest that the amount of required data is reduced and robust against the SDE's dimensionality.

2.3. Inferring the diffusion term

Once the drift is successfully identified, a regression equation can be formulated to identify the diffusion term. Based on Eq. (6) for diffusion, we consider

$$\frac{\sum_{s=1}^{K}\left\{\left[\widehat{\boldsymbol{x}}(t_{j_s+1})-\widehat{\boldsymbol{x}}(t_{j_s})-\boldsymbol{f}\left(\widehat{\boldsymbol{x}}(t_{j_s})\right)\Delta t\right]\left[\widehat{\boldsymbol{x}}(t_{j_s+1})-\widehat{\boldsymbol{x}}(t_{j_s})-\boldsymbol{f}\left(\widehat{\boldsymbol{x}}(t_{j_s})\right)\Delta t\right]^{\mathrm{T}} \mid \widehat{\boldsymbol{x}}(t_{j_s}) = \xi\right\}}{K\Delta t} \approx \boldsymbol{G}(\boldsymbol{\xi}) + \boldsymbol{\varepsilon}_2 \tag{20}$$

where $\boldsymbol{\varepsilon}_2$ is a matrix with each entry being Gaussian distributed with mean zero and variance proportional to $1/K$. The identified drift was used to estimate the corresponding exact drift value. Following similar lines as in the identification of the drift, we can formulate a regression equation to identify the diffusion term and leverage the SBL approach to solve it. Note that the selection of basis functions here can be different from those used for drift identification.

2.4. Model validation

Next, we evaluated the proposed BISDE algorithm through simulations and experiments. For simulations, because the exact model structures are known, we can compare the mean squared error (MSE) between the data generated by the true drift (diffusion) and the fitted drift (diffusion) at the measurements to assess the identified models' performance. For real experimental data, owing to the absence of ground truth, we designed three criteria to assess the identified models' performance.

Criterion 1: If a widely accepted model exists, we can leverage it to assess the identified model. If a good consistency is found between them, the effectiveness of such a model is cross verified. Otherwise, obvious inconsistency may imply the potential to determine a more efficient model whose performance can be assessed by Criteria 2 or 3.

Criterion 2: If the real dynamic process is approximately stationary, we can compare the analytic probability density function (PDF) of the identified model derived by solving the stationary Fokker–Planck equation with the empirical PDF of the time-series. A good match between them suggests the soundness of identification. Non-stationary processes can be transformed into stationary processes by constructing algebraic or logarithmic increments, or by other methods [28].

Criterion 3: When the PDF cannot be solved analytically or estimated numerically for stationary processes, we can run the identified model with initial conditions identical to those of real data, and then assess the identified model's performance by comparing the empirical PDF of the simulated data to that of the real data.

3. Results

3.1. Discovering the Langevin equation using simulated data

We first applied the proposed BISDE algorithm to the Langevin equation (Section S3.1 in Appendix A), which plays an important role in physics [29,30], as shown in Fig. 1. Despite its simple mathematical form, it took physicists nearly a hundred years to discover this equation. Specifically, the Langevin equation describes the dynamics of Brownian particles over all time scales, which overcomes the shortcomings of Einstein's theory in Ref. [1]. Consider a Brownian particle immersed in a fluid (Fig. 1(a)). Its random movements are driven by collision with liquid molecules from all directions due to thermal motion, and its velocity follows the Langevin equation.

The data were obtained by uniformly discretizing the time interval [0, 1000] with a time step $\Delta t = 0.04$. The basis functions consist of the constant term, polynomials in $\boldsymbol{x}$ up to the order of 15, and exponential functions with exponents from $\boldsymbol{x}$ to $10\boldsymbol{x}$. In this example, the same basis functions were employed for both drift and diffusion identification. If prior information relating to drift or diffusion, such as symmetry and periodicity, is available, the basis functions should be specified for each term. Figs. 1(b) and (c) show that BISDE successfully identifies the Langevin equation with very high accuracy. This example with the Langevin equation highlights the ability of BISDE to assist physicists in identifying the underlying SDEs of random phenomena from relatively limited quantities of time-series data.

3.2. Discovering the dynamics of bearing vibration from experimental data

Next, we showcase the discovery of the dynamics of the rolling bearing vibration from the Case Western Reserve University (CWRU) bearing dataset. With the rapid development of modern industries, rotating machines are being widely used in manufacturing systems and household appliances. Although the rolling bearing has found wide and indispensable use in these machines, it is also ranked as the top component related to machinery defects [31–33]. A bearing fault reduces machine life and performance, lowers the quality of workpieces, and causes safety risks and even casualties in extreme cases. Consequently, bearing fault diagnoses have become a popular topic in the engineering community. In general, vibration signals are considered the most informative data for evaluating bearing defects, as any fault in the bearings can affect the vibration dynamics [34]. Therefore, determining the dynamics of bearing vibration signals at the fault-free and faulty stages can provide knowledge about potential bearing defects.

The CWRU dataset is an open-source dataset that is used to explore the dynamics of normal and faulty bearings. The original test stand and its cross-sectional view are shown in Figs. 2(a) and (b), respectively. The stand consists of an electric motor, an encoder or torque transducer, and a dynamometer. Data were collected at 48 kHz under three different states: ① normal bearings (NBs), ② inner race faults (IRFs), and ③ ball faults (BFs). Single-point faults were introduced to the drive-end and fan-end bearings with fault diameters of 7, 14, and 21 mils (1 mil = 0.0254 mm). Every faulted bearing was reinstalled into the motor and tested for motor loads of 1–3 horsepower (hp; 1 hp = 0.7457 kW).

We focus on the results of normal and faulted fan-end bearings of 1 hp herein; a more comprehensive comparison can be found in Section S3.2 in Appendix A. Because there is a sufficient amount of data, the binning operation can provide reasonable estimates for

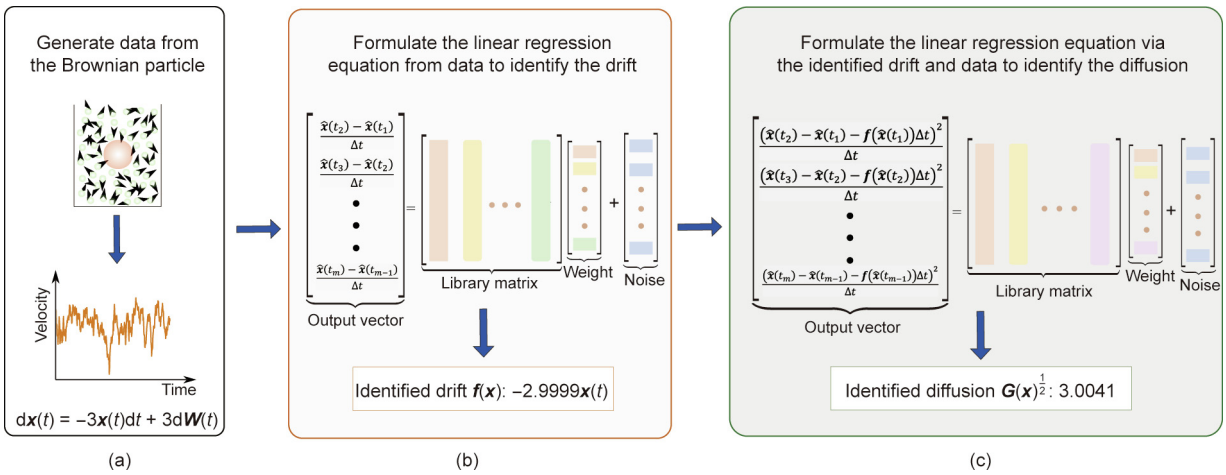

Fig. 1. Pipeline of the BISDE algorithm on the Langevin equation. (a) A Brownian particle (red dot) immersed in a fluid with its velocity satisfying the Langevin equation; (b, c) formulate the linear regression equation to identify the drift and diffusion from the space of candidate basis functions, respectively.

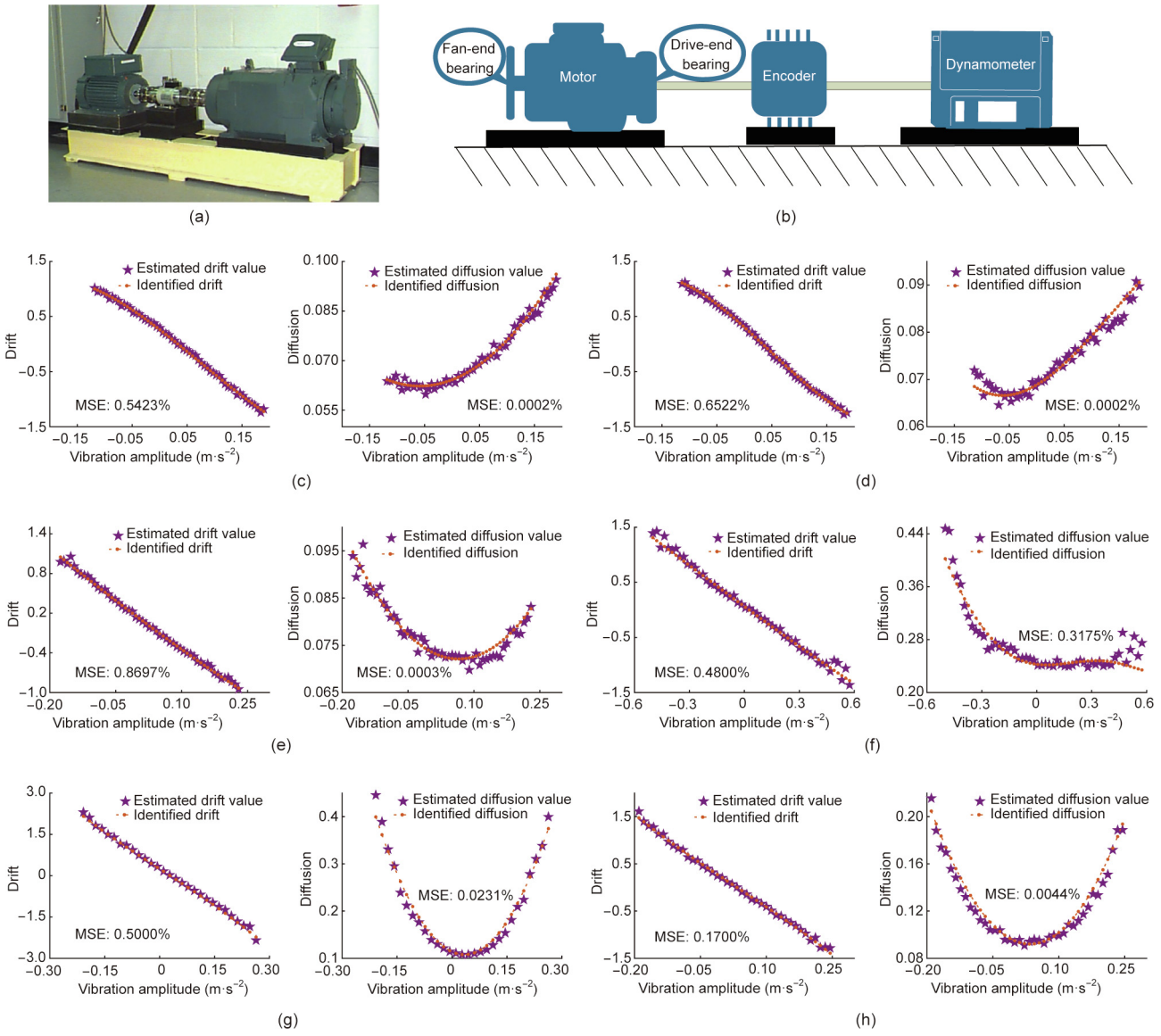

Fig. 2. Discovering the dynamics of normal and faulted fan-end bearings. (a) Bearing experimental platform of the CWRU dataset; (b) a cross-sectional view of the original test stand, consisting of a motor (left), a torque transducer/encoder (center), and a dynamometer (right); (c, d) identification results for NBs of 1 and 2 hp, respectively; (e, f) identification results for BF and IRF with the fault diameter of 7 mils, respectively (motor loads: 1 hp); (g, h) identification results for BF and IRF with the fault diameter of 14 mils, respectively (motor loads: 1 hp).

the drift and diffusion terms. Moreover, the estimates can be used as benchmarks to verify the identification results of BISDE. To illustrate that our method can reduce the required amount of data, BISDE only used about one-tenth of the dataset to discover the underlying dynamics. We present the comparison between the drift and diffusion estimates of the binning operation and the corresponding drift and diffusion identified by BISDE under different operating conditions in Figs. 2(c)–(h); the identified model's performance was assessed using the MSE. This demonstrates that the identified drift and diffusion can accurately capture the dynamics of the bearing vibration signals. The identified models of NBs under different motor loads can aid operators or practitioners in conducting early diagnoses of faults to prevent disastrous consequences and reduce maintenance expenses.

3.3. Discovering the canonical stochastic models and applications

Finally, we applied the proposed BISDE algorithm to determine several canonical and real-world SDEs. The simulation models are based on common physical systems and stochastic processes (Sections S3.3–S3.5 in Appendix A). A two-dimensional simulation model is used to validate the power of the BISDE algorithm in identifying multidimensional SDEs from a limited amount of data. Real-world systems include stock price fluctuations, wind speeds, and oil prices (Sections S3.6–S3.8 in Appendix A). Enabling an identification framework for such stochastic dynamical systems can help practitioners improve system design and develop more efficient system management strategies for different scenarios. A more detailed illustration can be found below and in Section S3. Additionally, data and code implementations are available at https://github.com/HAIRLAB/BISDE.

Fig. 3 summarizes the simulated and real-world systems to be identified. Each category of three examples is marked with a specific background color. The first and fourth rows illustrate the simulated and real-world systems, respectively. The second row shows the simulated sample paths with colors denoting their probability density values, whereas the fifth row only shows real sample paths due to the lack of ground-truth information. The third and sixth rows assess the performance of the identified models. For simulated systems, we can compare the MSE between the data points generated by the real and identified drift/diffusion at measurements to assess its performance. We adopted Criteria 1, 2, and 3 to assess the identified models' performance of real-world systems.

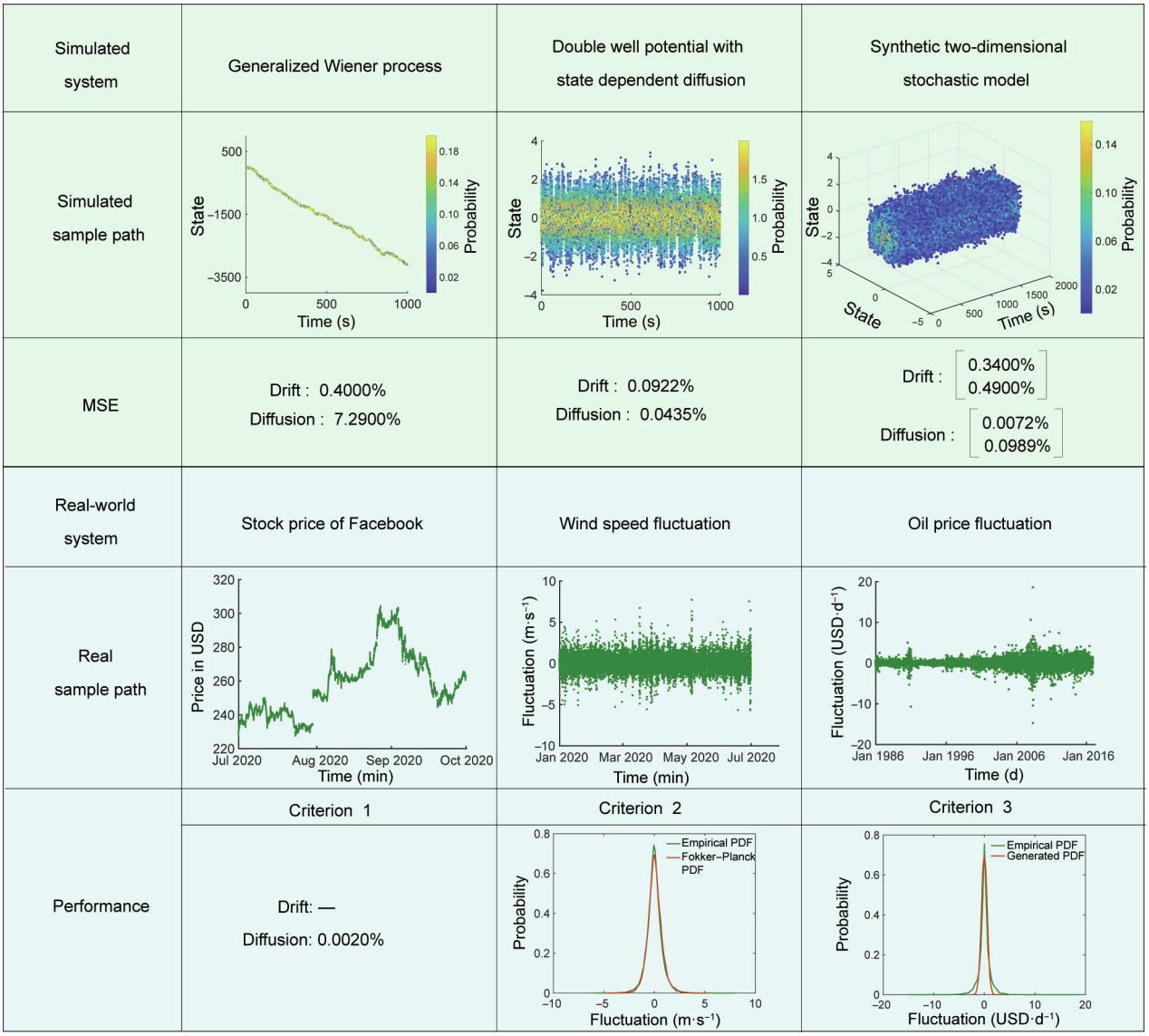

Fig. 3. Summary of the BISDE algorithm applied to numerous examples. BISDE has been tested on three simulated systems, including a two-dimensional system, and three real-world systems, where each type is marked with a specific background color.

Comparison with state-of-the-art method: To demonstrate that BISDE can identify SDEs from relatively limited quantities of time-series data, we compared it with the state-of-the-art method in Ref. [17], which is referred to as SDE_SINDy hereinafter for convenience. Fig. 4 presents a visual comparison of BISDE and SDE_SINDy on the simulated systems. When the amount of data was relatively limited, SDE_SINDy failed to identify the underlying model structures for all cases, whereas BISDE yielded a near-perfect identification result.

Financial economics—stock prices: Stock prices are mutually influenced by many economic, financial, and political factors. The dynamics of stock prices can be considered a stochastic process because random noise introduces uncertainty when predicting future stock prices. In the 1970s, the weak-form efficient market hypothesis proposed by Eugene Fama, a recipient of the 2013 Nobel Prize in Economics, stated that future stock prices cannot be predicted by analyzing historical data [35]. Based on this hypothesis, stock prices are usually assumed to follow a Markov process. Lengthy historical data sequences are unhelpful in determining the dynamics of stock prices as these dynamics change over time [3]. Therefore, we collected the stock price data of Facebook every minute over three months from 1 July 2020 to 30 September 2020 (Figs. 5(a) and (b)).

After applying the proposed BISDE algorithm (Fig. 5(c)), the identified geometric Brownian motion model describing Facebook's stock price behavior is shown in Fig. 5(d). Geometric Brownian motion is the most widely accepted model for describing stock price behaviors [3], which is one of the assumptions used to derive the Black–Scholes–Merton formula to price European call and put options [36,37]. Surprisingly, the identified volatility (0.4039) almost coincides with the estimated volatility (0.4087) using the method suggested in Ref. [3], demonstrating the accuracy of the identified model. Compared with the one-year annual percentage yields (APY) provided by different banks (Fig. 5(e)), one

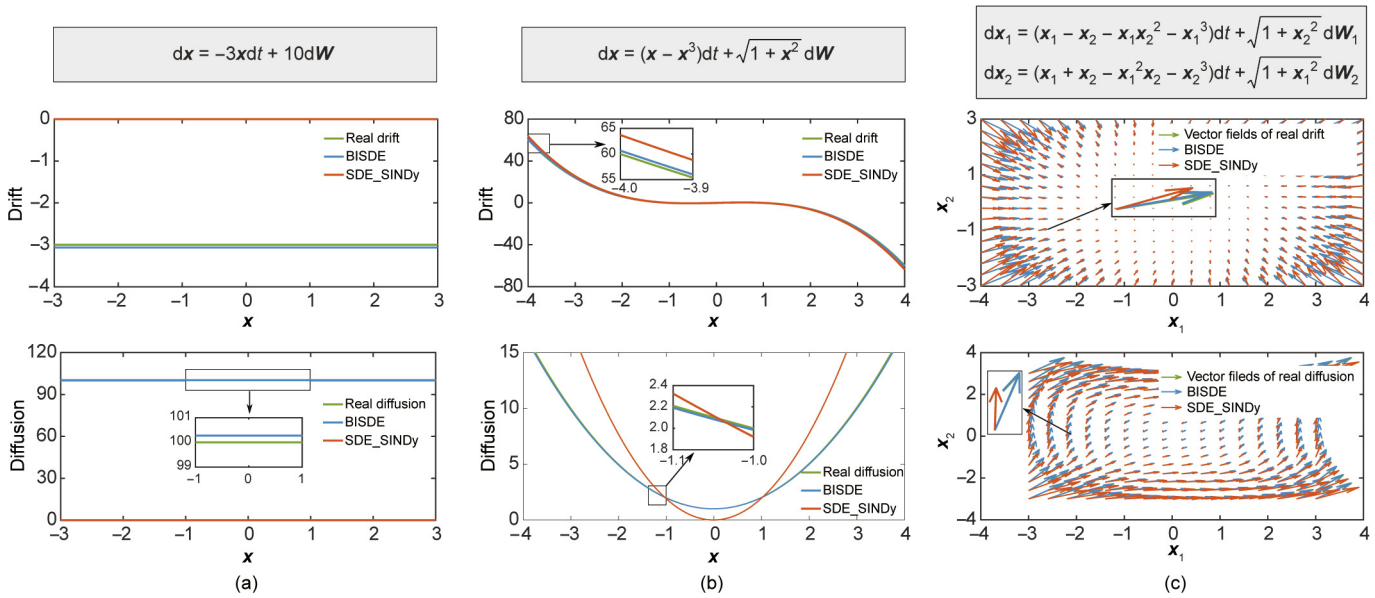

Fig. 4. Comparison of BISDE and SDE_SINDy on the simulated systems. The identification results of BISDE and SDE_SINDy on (a) the generalized Winner process, (b) double well potential with state dependent diffusion, and (c) synthetic two-dimensional stochastic model.

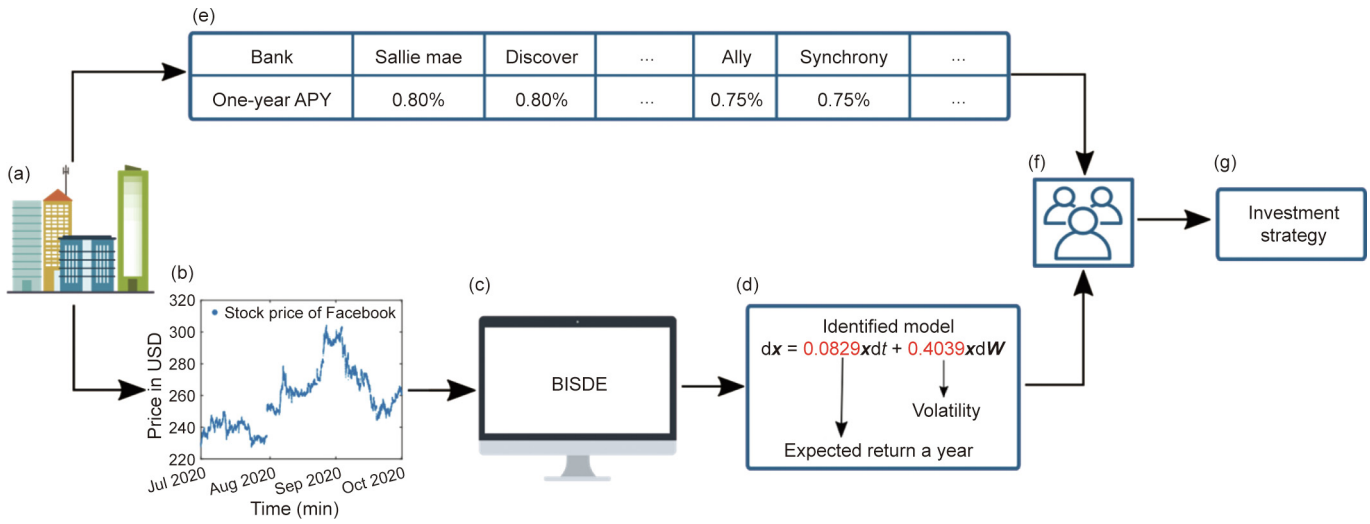

Fig. 5. Discovering the stock price behavior of Facebook. (a) Financial institutions (data acquisition); (b) stock price of Facebook over three months; (c) stock price data analyses; (d) identified SDE used to describe the stock price behavior of Facebook; (e) summary of the APY on one-year certificates of the deposit of different US banks in August 2020; (f) investors or technical analysts; (g) investment strategy made by investors by integrating investment information.

can infer that the expected return on buying Facebook stock is several times the return on saving money in the given banks. Based on the identified model and one-year APY, investors can choose a strategy to balance the expected return and the uncertainty or risk of the stock to achieve better income (Figs. 5(f) and (g)). Overall, this application reveals that BISDE is a powerful tool for identifying the dynamics of stocks. The identified model can provide insights into stocks for investors and help determine price of stock options for stock exchanges [3].

Power systems—wind speed fluctuations: As a widely distributed, sustainable, and renewable energy source, wind plays an important role in power grids across many countries. Globally, wind has accounted for approximately 5% of total power generation up to the end of 2020. Because wind speeds are uncontrollable and fluctuating under both spatial and temporal scales, the power production of a wind farm may vary from one minute to the next even when the total yearly production remains almost constant. Wind speed fluctuations affect the nominal power output, sometimes eroding power quality and reliability, as well as causing extreme gust and wind turbine fatigue loading [7,38,39]. Therefore, determining the dynamics of wind speed fluctuations is essential for power production and load design to ensure the safety and economy of wind energy resources.

To illustrate the applicability of BISDE for this problem, we collected wind speed data for the first half of 2020 from Greta Point Cws, Wellington, New Zealand, located at 174.80574°E, 41.30243°S, and 3 m above the mean sea level [40]. Wind speed fluctuations were obtained by calculating differences in wind speed data, resulting in the conversion of a non-stationary series (wind speed) to a stationary series (wind speed fluctuations). The identified model can be seen as the Ornstein–Uhlenbeck process with an added quadratic state-dependent term in the diffusion. Next, by solving the stationary Fokker–Planck equation, an analytical PDF for the identified model is derived. Its performance was verified by the high similarity between the analytic PDF and the empirical PDF generated by the measurements (Fig. 3). BISDE successfully identified an SDE to describe wind speed fluctuations. By embedding the identified model into wind turbine models, we can perform dynamic studies and thus maintain control of wind turbines [41,42].

Energy economies—oil price fluctuations: Despite the growing importance of renewable energy, oil continues to be the dominant resource in most countries. In contrast to almost all other commodities, oil has a considerable impact on the macroeconomy owing to its irreplaceability and high degree of liquidity in the international marketplace [43,44]. Oil price fluctuations cause significant losses or profits for both exporting and importing countries. Furthermore, oil price fluctuations may lead to inflation, increases in transportation costs, and changes in business policies, affecting not only industries but the daily lives of individuals and families. Because oil prices are affected by many independent and interrelated factors, an accurate model is challenging to formulate.

We applied the BISDE algorithm to resolve this challenging problem. To this end, we used crude oil prices collected from the US Energy Information Administration from the beginning of 1986 to the middle of 2017. First, we calculated the daily changes in oil prices. After determining the oil price fluctuations, we applied BISDE to discover the underlying dynamics. To illustrate the quality of the identified model, we generated a sample path based on the identified SDE with the same initial value as that of the real data. Then, we compared the empirical PDFs of the real and simulated data. The identified model yields a PDF very close to that of the real data in the high-probability region (Fig. 3). Given current oil price fluctuations, the identified model can help predict the next fluctuation and address the corresponding uncertainties in terms of firms and policymakers, which are useful for avoiding unnecessary losses.

4. Discussion

This study provides a novel parametric algorithm called BISDE for the discovery of SDE-governed systems. This algorithm has several significant advantages. First, unlike non-parametric methods which cannot provide interpretable models [13], our algorithm identifies the model structures of the drift and diffusion terms to determine the underlying mechanisms of stochastic dynamical systems. Second, existing state-of-the-art data-driven algorithms project multidimensional systems to low-dimensional ones to reduce the required quantities of data [17,18]. In contrast, BISDE can directly discover original multidimensional systems with a limited amount of data. This capability was verified by the identification of a two-dimensional SDE and can be generalized naturally to higher-dimensional SDEs. Finally, although BISDE employs the Euler–Maruyama discretization method as in previous studies [10,13,14], we specified the limit conditions of the drift and diffusion terms, performed convergence analyses between the to-be-identified and real terms, and bridged non-parametric estimation with parametric identification. Overall, BISDE is a viable method for identifying the underlying dynamics of stochastic dynamical systems with relatively limited quantities of time-series data, and it possesses the potential to help researchers model natural phenomena and engineered systems perturbed by random noise from different sources.

Despite the advantages of BISDE, several questions remain. First, we note that prior knowledge is helpful in choosing basis functions, making the identification of model structures faster and more accurate. However, the lack of prior information makes it difficult to build libraries. To make matters potentially worse, true models may be approximated using polynomials, kernels, and other functional forms [22]. There is also the question of establishing a verification method for non-stationary processes without knowing the exact model structures, because in many cases, we need to explore the intrinsic characteristics of state variables directly. Finally, two different SDEs may produce the same PDF (Section S4 in Appendix A), in which case they would possess the same statistical properties but involve two completely different mechanisms, which may mislead researchers.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFB1701202), the National Natural Science Foundation of China (92167201 and 51975237), and the Fundamental Research Funds for the Central Universities, Huazhong University of Science and Technology (2021JYCXJJ028).

Authors' contribution

Ye Yuan conceived of and supervised the project. Yasen Wang developed the algorithm and conducted experiments. All authors discussed the results and prepared and revised the manuscript accordingly.

Compliance with ethics guidelines

Yasen Wang, Huazhen Fang, Junyang Jin, Guijun Ma, Xin He, Xing Dai, Zuogong Yue, Cheng Cheng, Hai-Tao Zhang, Donglin Pu, Dongrui Wu, Ye Yuan, Jorge Gonçalves, Jürgen Kurths, and Han Ding declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2022.02.007.

References

[1] Einstein A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann Phys 1905;322(8):549–60. German.
[2] Bose T, Trimper S. Stochastic model for tumor growth with immunization. Phys Rev E 2009;79:051903.
[3] Hull JC. Options, futures, and other derivatives. 9th ed. Boston: Pearson; 2015.
[4] Wilkinson DJ. Stochastic modelling for quantitative description of heterogeneous biological systems. Nat Rev Genet 2009;10(2):122–33.
[5] Chong KL, Shi JQ, Ding GY, Ding SS, Lu HY, Zhong JQ, et al. Vortices as Brownian particles in turbulent flows. Sci Adv 2020;6(34):eaaz1110.
[6] Rigas G, Morgans AS, Brackston RD, Morrison JF. Diffusive dynamics and stochastic models of turbulent axisymmetric wakes. J Fluid Mech 2015;778 (R2):1–10.
[7] Calif R. PDF models and synthetic model for the wind speed fluctuations based on the resolution of Langevin equation. Appl Energy 2012;99:173–82.
[8] Friedrich R, Siegert S, Peinke J, Lück St, Siefert M, Lindemann M, et al. Extracting model equations from experimental data. Phys Lett A 2000;271:217–22.
[9] Lamouroux D, Lehnertz K. Kernel-based regression of drift and diffusion coefficients of stochastic processes. Phys Lett A 2009;373:3507–12.
[10] Rajabzadeh Y, Rezaie AH, Amindavar H. A robust nonparametric framework for reconstruction of stochastic differential equation models. Phys A 2016;450:294–304.
[11] Papaspiliopoulos O, Pokern Y, Roberts GO, Stuart AM. Nonparametric estimation of diffusions: a differential equations approach. Biometrika 2012;99:511–31.
[12] Van der Meulen F, Schauer M, Van Zanten H. Reversible jump MCMC for nonparametric drift estimation for diffusion processes. Comput Stat Data Anal 2014;71:615–32.
[13] Batz P, Ruttor A, Opper M. Approximate Bayes learning of stochastic differential equations. Phys Rev E 2018;98:022109.
[14] Garcia CA, Otero A, Felix P, Jesus P, Marquez DG. Nonparametric estimation of stochastic differential equations with sparse Gaussian processes. Phys Rev E 2017;96:022104.
[15] Bandi FM, Phillips PCB. A simple approach to the parametric estimation of potentially nonstationary diffusions. J Econom 2007;137:354–95.
[16] Brunton SL, Proctor JL, Kutz JN. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc Natl Acad Sci USA 2016;113:3932–7.
[17] Boninsegna L, Nuske F, Clementi C. Sparse learning of stochastic dynamical equations. J Chem Phys 2018;148:241723.
[18] Callaham JL, Loiseau JC, Rigas G, Brunton SL. Nonlinear stochastic modeling with Langevin regression. Proc R Soc A Math Phys Eng Sci 2021;477:20210092.
[19] Tipping ME. Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 2001;1:211–44.
[20] Wipf DP, Rao BD. Sparse Bayesian learning for basis selection. IEEE Trans Signal Process 2004;52:2153–64.
[21] Pan W, Yuan Y, Goncalves J, Stan GB. A sparse Bayesian approach to the identification of nonlinear state–space systems. IEEE Trans Auto Control 2016;61:182–7.
[22] Yuan Y, Tang X, Zhou W, Pan W, Li X, Zhang HT, et al. Data driven discovery of cyber physical systems. Nat Commun 2019;10:1–9.
[23] Ping Z, Li X, He W, Yang T, Yuan Y. Sparse learning of network-reduced models for locating low frequency oscillations in power systems. Appl Energy 2020;262:114541.
[24] Zhou W, Ardakanian O, Zhang HT, Yuan Y. Bayesian learning-based harmonic state estimation in distribution systems with smart meter and DPMU data. IEEE Trans Smart Grid 2020;11:832–45.
[25] Yuan Y, Zhang H, Wu Y, Zhu T, Ding H. Bayesian learning-based model-predictive vibration control for thin-walled workpiece machining processes. IEEE-ASME Trans Mechatron 2017;22:509–20.
[26] Mao X. Numerical solutions of stochastic differential delay equations under the generalized Khasminskii-type conditions. Appl Math Comput 2011;217:5512–24.
[27] Mao X. The truncated Euler–Maruyama method for stochastic differential equations. J Comput Appl Math 2015;290:370–84.
[28] Ghasemi F, Sahimi M, Peinke J, Friedrich R, Jafari GR, Tabar MRR. Markov analysis and Kramers–Moyal expansion of nonstationary stochastic processes with application to the fluctuations in the oil price. Phys Rev E 2007;75:060102.
[29] Langevin P. Sur la théorie du mouvement Brownien. C R Acad Sci 1908;146:530–3. French.
[30] Coffey W, Kalmykov YP. The Langevin equation: with applications to stochastic problems in physics, chemistry and electrical engineering. 3rd ed. Singapore: World Scientific; 2012.
[31] Shao H, Jiang H, Zhang X, Niu M. Rolling bearing fault diagnosis using an optimization deep belief network. Meas Sci Technol 2015;26:115002.
[32] Yuan Y, Ma G, Cheng C, Zhou B, Zhao H, Zhang HT, et al. A general end-to-end diagnosis framework for manufacturing systems. Natl Sci Rev 2020;7:418–29.
[33] Cheng C, Ma G, Zhang Y, Sun M, Teng F, Ding H, et al. A deep learning-based remaining useful life prediction approach for bearings. IEEE-ASME Trans Mechatron 2020;25(3):1243–54.
[34] Safizadeh MS, Latifi SK. Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell. Inf Fusion 2014;18:1–8.
[35] Fama EF. Efficient capital markets: a review of theory and empirical work. J Financ 1970;25:383–417.
[36] Black F, Scholes M. The pricing of options and corporate liabilities. J Polit Econ 1973;81:637–54.
[37] Merton RC. Theory of rational option pricing. Bell J Econ Manag Sci 1973:141–83.
[38] Zárate-Miñano R, Anghel M, Milano F. Continuous wind speed models based on stochastic differential equations. Appl Energy 2013;104:42–9.
[39] Zárate-Miñano R, Milano F. Construction of SDE-based wind speed models with exponentially decaying autocorrelation. Renew Energy 2016;94:186–96.
[40] National Institute of Water and Atmospheric Research Limited. CliFlo: NIWA's National Climate Database [Internet]. Auckland: National Institute of Water and Atmospheric Research Limited; [cited 2020 Dec 8]. Available from: http://cliflo. niwa.co.nz/.
[41] Kusiak A, Li W, Song Z. Dynamic control of wind turbines. Renew Energy 2010;35:456–63.
[42] Melício R, Mendes VMF, Catalão JPS. Transient analysis of variable-speed wind turbines at wind speed disturbances and a pitch control malfunction. Appl Energy 2011;88:1322–30.
[43] Chang Y, Wong JF. Oil price fluctuations and Singapore economy. Energ Policy 2003;31:1151–65.
[44] Lizardo RA, Mollick AV. Oil price fluctuations and US dollar exchange rates. Energy Econ 2010;32:399–408.