Contents lists available at ScienceDirect
Engineering
journal homepage: www.elsevier.com/locate/eng

Research
Engineering Management—Article

数据驱动的随机微分方程辨识

王亚森[a,b]，方华臻[c]，金骏阳[d]，马贵君[a,b]，何心[a]，代星[a,d]，岳作功[e]，程骋[e]，张海涛[b,e]，浦栋麟[d]，伍冬睿[e]，袁烨[a,b,e,*]，Jorge Gonçalves[e,f,g]，Jürgen Kurths[h,i]，丁汉[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

关键词
数据驱动方法
系统辨识
稀疏贝叶斯学习
随机微分方程
随机现象

摘要

随机微分方程(SDE)是一种广泛用于描述受不同来源噪声干扰的复杂过程或现象的数学模型。由于数据固有的强随机性和系统动力学的复杂性，控制系统的SDE的辨识通常是一个挑战。辨识SDE的现有参数化方法的实用性通常受到数据资源不足的限制。本研究提出了一种通过稀疏贝叶斯学习(SBL)技术从候选基函数空间中搜索简洁但物理必需的表示形式来辨识SDE的新框架。更重要的是，利用SBL的解析可处理性开发了一种可以通过少量数据来高效地构建辨识SDE的线性回归模型的方法。本文利用股票和石油价格、轴承变化和风速的真实数据，以及包括广义维纳(Wiener)过程和朗之万方程在内的著名随机动力系统的仿真数据，验证了所提出框架的有效性。该框架旨在帮助专家从自然科学、经济学和工程领域的随机现象中提取随机数学模型，用于分析、预测和决策。

1. 引言

自然界、工业和人类社会充满了在随机噪声影响下演化的现象、过程和系统。这类系统的实例包括：布朗粒子在流体中的运动[1]、由肿瘤-免疫相互作用驱动的肿瘤演变[2]、股票价格和风的运动。随机微分方程（SDE）是建模和分析受随机噪声影响的系统的一种强有力的数学方法。对SDE的研究起源于爱因斯坦在1905年发表的论文[1]，该论文建立了布朗粒子在热浴中的波动理论。从那之后，SDE已经在众多科学和工程领域得到了广泛的应用[3–7]。传统上，可以根据数据的统计特性和专家的经验利用SDE对随机现象进行建模。然而，在许多情况下，数据的统计特性可能与充分理解的SDE的统计特性不同，因此对于发现潜在SDE的帮助较小。此外，用于控制随机现象的SDE通常未知或难以捉摸。因此，开发一种数据驱动方法来辨识随机现象的潜在SDE至关重要。然而，

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

英文原文：Engineering 2022, 17(10): 244–252
引用本文：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, Han Ding. Data-Driven Discovery of Stochastic Differential Equations. *Engineering*, https://doi.org/10.1016/j.eng.2022.02.007

系统（由SDE控制）的复杂行为和强随机性使得精确辨识SDE具有很大的挑战性。

确定SDE的关键是辨识其漂移项和扩散项。关于这一主题的工作可以大致分为两类。第一类是非参数化辨识，该工作试图从时间序列数据中为漂移项和扩散项建立一个无模型的输入输出映射。Kramers-Moyal平均法引入了基于直方图回归（HBR）为漂移项和扩散项建立适当映射[8]。然而，即使对于一维SDE，该方法也需要大量的数据，并且所需的数据量随着SDE的维数呈指数增长。实际上，这种数据需求加重了传感器成本、数据存储和计算资源的负担，这在某些情况下可能难以满足。为了利用有限的数据量提高辨识精度，基于核函数[9]和多项式[10]的回归方法可以作为针对一维SDE更有效的映射方法。此外，非参数化的贝叶斯估计提供了另一种减少数据量要求的映射方法[11–14]。最近提出采用马尔可夫链蒙特卡罗（Markov chain Monte Carlo, MCMC）方法[11–12]和高斯过程回归（GPR）[13]来辨识SDE。进一步地，具有高计算效率的稀疏GPR方法也被提出用于辨识SDE [13–14]。虽然上述方法可以充分地预测漂移项和扩散项，但是它们不能对随机动力系统建模，因为它们仅提供模型的黑箱表示。

相比之下，参数化方法侧重于辨识SDE的漂移项和扩散项的模型结构，并且更有利于揭示随机动力系统的潜在物理规律。对于具有已知模型结构的标量齐次SDE，可以结合KBR和最小二乘法来估计参数[15]。对于非线性动力系统，在假设模型结构在可能的基函数空间中稀疏的情况下，非线性动力学稀疏辨识（SINDy）框架已经被用来确定控制方程[16]。基于HBR和SINDy，已经提出了稀疏学习方法来辨识漂移项和扩散项[17–18]。然而，因为这些方法最初依赖于HBR估计漂移项和扩散项，所以它们表现出与HBR相同的缺点。

在系统辨识和信号处理领域，一种用于辨识模型结构的新兴技术为稀疏贝叶斯学习（SBL）[19–25]，其目的在于通过平衡模型的复杂性和准确性从输入输出数据中找到模型在基函数空间的简洁表示。本研究利用相对有限的时间序列数据，通过此技术来发现随机动力系统的潜在SDE。所提出的算法的实现过程可以概括为以下两个阶段。首先，利用Euler-Maruyama方法对SDE进行离散，得到漂移项和扩散项的理论表达式。然后，基于中心极限定理，将SDE的辨识问题转化为一个针对漂移项和扩散项的输入输出回归问题。虽然分箱操作可以用于估计选定点处的漂移项和扩散项的值，从而实现输入输出回归问题的构建表达，但是此操作需要大量数据并且受到维度灾难的影响。基于稀疏贝叶斯学习的解析可处理性，本文提出了一种在实际中更有效的从相对有限的时间序列数据中构建输入输出回归问题的实现方法。通过与最先进的方法相比，所提的SDE的贝叶斯辨识（BISDE）方法的强大性能和鲁棒性已经在著名的SDE上得到了证明。此外，本文所提出的BISDE算法在大量的仿真和实际系统中得到了验证。

2. 方法

2.1. 漂移项和扩散项的数学表达式

在本研究中，考虑了n维SDE的一般形式：

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

式中，$\boldsymbol{x}(t)\in\mathbb{R}^n$是时刻$t$的状态向量；$\boldsymbol{f}\left(\boldsymbol{x}(t)\right)\in\mathbb{R}^n$是状态相关的漂移向量；$\boldsymbol{G}\left(\boldsymbol{x}(t)\right)\in\mathbb{R}^{n\times n}$是正定的扩散矩阵；$\boldsymbol{W}(t)$是$n$维标准布朗运动或维纳过程。假设$\boldsymbol{f}\left(\boldsymbol{x}(t)\right)$和$\boldsymbol{G}\left(\boldsymbol{x}(t)\right)$的每个元素都是连续函数。但是，$\boldsymbol{f}\left(\boldsymbol{x}(t)\right)$和$\boldsymbol{G}\left(\boldsymbol{x}(t)\right)$的确切模型结构未知。本研究的目的是从关于$\boldsymbol{x}(t)$相对有限的时间序列数据中辨识它们。

由于SDE的解析解通常难以得到，因此采用Euler-Maruyama离散法来计算其数值解。为了确保方程（1）中SDE解的存在性和唯一性以及Euler-Maruyama方法的可行性，需要假设漂移项和扩散项满足局部Lipschitz和Khasminskii型条件（见附录A中的第S1节）。此外，在此条件下，基于Euler-Maruyama方法的数值解依概率收敛于解析解[26–27]。

将Euler-Maruyama方法应用于方程（1）的SDE，得到

$$\begin{aligned}&\hat{\boldsymbol{x}}\left((k+1)\Delta t\right)-\hat{\boldsymbol{x}}\left(k\Delta t\right)\\&=\boldsymbol{f}\left(\hat{\boldsymbol{x}}\left(k\Delta t\right)\right)\Delta t+\boldsymbol{G}\left(\hat{\boldsymbol{x}}\left(k\Delta t\right)\right)^{\frac{1}{2}}\sqrt{\Delta t}\,\boldsymbol{\epsilon}_k\end{aligned} \tag{2}$$

式中，k是离散化时间索引；Δt是离散化时间步长；$\hat{\boldsymbol{x}}(k\Delta t)$是$\boldsymbol{x}(k\Delta t)$的数值解，且$\boldsymbol{\epsilon}_k\sim\mathcal{N}\left(0,\boldsymbol{I}_n\right)$，满足均值为零且协方差矩阵为$\boldsymbol{I}_n$（$\boldsymbol{I}_n$为$n\times n$单位矩阵）的正态分布。考虑连续时间近似进行分析更方便。因此，对于任意$t\in(k\Delta t,\ (k+1)\Delta t)$，设

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

已知$\hat{\boldsymbol{x}}(k\Delta t)$，显然，方程（2）中的$\hat{\boldsymbol{x}}\left((k+1)\Delta t\right)$满足高斯分布。这意味着，对于任意$t\geq 0$，

$$\hat{\boldsymbol{x}}(t+\Delta t)|\hat{\boldsymbol{x}}(t)\sim\mathcal{N}\left(\hat{\boldsymbol{x}}(t)+\boldsymbol{f}\left(\hat{\boldsymbol{x}}(t)\right)\Delta t,\boldsymbol{G}\left(\hat{\boldsymbol{x}}(t)\right)\Delta t\right) \tag{4}$$

根据方程（4），可以推导出漂移项和扩散项的表达式。首先，对于任意点$\xi\in\mathbb{R}^n$，通过条件期望可以得到漂移项$\boldsymbol{f}(\xi)$的表达式如下：

$$\boldsymbol{f}(\xi)=\frac{1}{\Delta t}E\left[\hat{\boldsymbol{x}}(t+\Delta t)-\hat{\boldsymbol{x}}(t)|\hat{\boldsymbol{x}}(t)=\xi\right] \tag{5}$$

式中，E表示期望算子。

与漂移项的情况类似，通过条件方差可以得到扩散项$\boldsymbol{G}(\xi)$的表达式如下：

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

式中，T表示转置算子。

虽然是根据数值解$\hat{\boldsymbol{x}}(t)$来估计漂移项和扩散项，但是可以证明：当$\Delta t\to 0$时，$\boldsymbol{f}(\hat{\boldsymbol{x}}(t))$ 和$\boldsymbol{G}(\hat{\boldsymbol{x}}(t))$的所有元素分别依概率收敛于$\boldsymbol{f}(\boldsymbol{x}(t))$和$\boldsymbol{G}(\boldsymbol{x}(t))$的相应元素。下面的命题总结了这个结果，为了简单起见，省略了元素索引的下标符号。

命题1. 假设方程（1）中的SDE满足局部Lipschitz和Khasminskii型条件，并且漂移项和扩散项的所有元素都是连续函数。那么，对于任意$\tau>0$，

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

和

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

证明：参见附录A中的第S2节。

根据中心极限定理，可以基于方程（5）通过收集的时间序列$\left\{\hat{\boldsymbol{x}}(t_i)\right\}_{i=1}^{m}$近似计算$\boldsymbol{f}(\xi)$。然而，计算$\boldsymbol{G}(\xi)$更复杂。根据方程（6），注意到计算$\boldsymbol{G}(\xi)$不仅需要数据，而且还需要$\boldsymbol{f}(\xi)$。一种可行的方法是使用辨识的漂移项来估计$\boldsymbol{f}(\xi)$的值。

2.2. 推断漂移项

在收集时间序列$\left\{\hat{\boldsymbol{x}}(t_i)\right\}_{i=1}^{m}$后，根据中心极限定理可以利用下式来估计$\boldsymbol{f}(\xi)$：

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

式中，K是等于ξ的$\hat{\boldsymbol{x}}(t_{j_s})$的数量，且$\hat{\boldsymbol{x}}(t_{j_s})\in\left\{\hat{\boldsymbol{x}}(t_i)\right\}_{i=1}^{m}$；$\boldsymbol{\varepsilon}_1$是一个向量，其中每个元素都服从均值为零、方差与$1/K$成比例的高斯分布。由于噪声可以被建模为高斯噪声，因此该方程将经验估计与SBL连接起来。因此，辨识漂移项的问题可以转化为一个输入输出回归问题。该方程还意味着可以独立地辨识漂移项的每一个元素。在不失一般性的情况下，假设$\xi(i)\in\mathbb{R}^n$，$\boldsymbol{f}_r(\xi(i))\in\mathbb{R}$，$i=1, 2, \ldots, N$分别表示第$r$个漂移项的输入输出数据，并根据方程（9）使用$K_i$对时间序列数据点估计$\boldsymbol{f}_r(\xi(i))$。设

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

假设第r个漂移项是属于候选函数库的一些基函数的线性组合。通常希望函数库足够大，来实现对潜在模型结构的彻底搜索和确定。对于许多实际系统，漂移项在由基函数张开的空间中是稀疏的，因为它仅包括几项基函数。此外，所考虑的随机动力系统的任何可用的先验知识都可以指导我们更有效地选择基函数。示例库包含如下的常数项、线性项和多项式项：

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

式中，$\boldsymbol{X}^2$、$\boldsymbol{X}^3$等表示更高次的多项式。例如，$\boldsymbol{X}^2$的每一列都可以被指定为$\xi(i)$ 和$\xi(j)$的元素乘积（其中i可以等于j）。剩下的问题在于估计基函数的权重。稀疏权重向量的精确辨识对于辨识第r个漂移项的模型结构至关重要。

为了估计权重向量，可以近似地求解从方程（9）推导出的以下回归方程：

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

式中，$\boldsymbol{\Phi}\in\mathbb{R}^{N\times M}$是构建的字典矩阵；$M$是基函数的数量；$\boldsymbol{\theta}=\left[\theta_1\,\theta_2\,\ldots\,\theta_M\right]^{\mathrm{T}}$是权重向量。假设噪声向量$\boldsymbol{\varepsilon}$遵循高斯分布$\mathcal{N}(0,\boldsymbol{\Psi})$，其中$\boldsymbol{\Psi}$是第$i$个元素为$\sigma^2/K_i$的对角矩阵，$\sigma^2$是一个标量的方差参数。首先，考虑漂移项的模型结构在所选基函数张开的空间中是稀疏的，对权重θ_i施加具有均值为零和方差为γ_i的稀疏促进的高斯先验。因此，$\boldsymbol{\theta}$表示具有初始概率分布$p(\boldsymbol{\theta};\boldsymbol{\gamma})$的随机向量，其中$\boldsymbol{\gamma}=\left[\gamma_1\,\gamma_2\,\ldots\,\gamma_M\right]^{\mathrm{T}}$。基于最大后验（MAP）原则，$\boldsymbol{\theta}$的后验分布

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

的均值被选为$\boldsymbol{\theta}$的估计值。这里，$p(\boldsymbol{Y}|\boldsymbol{\theta};\boldsymbol{\Psi})$是由方程（12）得出的似然函数。然而，$\boldsymbol{\theta}$的估计涉及$\boldsymbol{\gamma}$。为了设置$\boldsymbol{\gamma}$的合理值，可以最大化第II类似然函数$p(\boldsymbol{Y};\boldsymbol{\Psi},\boldsymbol{\gamma})=\int p(\boldsymbol{Y}|\boldsymbol{\theta};\boldsymbol{\Psi})p(\boldsymbol{\theta};\boldsymbol{\gamma})\mathrm{d}\boldsymbol{\theta}$。因此，在获得其最优值（表示为$\boldsymbol{\gamma}^*$）之后，可以得到

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

式中，$\boldsymbol{\Gamma}^*=\mathrm{diag}(\gamma^*)$。

对于一个感兴趣的系统，可能难以获得等于ξ的多个测量$\hat{\boldsymbol{x}}(t_i)$。近似地估计条件期望$\boldsymbol{f}(\xi)$的一个有用技巧是将落在$\xi$的小邻域内的任何数据点$\hat{\boldsymbol{x}}(t_{j_s})$视为$\xi$；或者，具

体而言，

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

式中，$\boldsymbol{\delta}=\left[\delta_1\,\delta_2\cdots\delta_n\right]^{\mathrm{T}}$是用于控制邻域大小的超参数向量。参考文献[8,17–18]已经证明了该分箱技术的有效性。以这种方式，对于n维SDE，数据点被划分进$\prod_{j=1}^{n}\left(\left\{\max\left[\hat{x}_j\left(t_i\right)\right]-\min\left[\hat{x}_j\left(t_i\right)\right]\right\}/2\delta_j\right)$个区间中，其中，$\hat{x}_j(t_i)$是$\hat{\boldsymbol{x}}(t_i)$的第$j$个元素。因此，如果想要保持总体上的近似精度，该方法会受到维数灾难的影响，因为SDE维数的增加会导致区间和数据的数量呈指数级增长。此外，也很难平衡区间的数量和近似的准确性。为了解决上述问题，本文开发了一种更有效的策略来构建回归方程以辨识漂移项的模型结构。首先，采用方程（12）中回归方程的等效实现来获得相同的权重向量。

定理1. 通过下式辨识的权重向量

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

即$\tilde{\boldsymbol{\theta}}$，与方程（12）中辨识的权重向量$\boldsymbol{\theta}$相同，式中

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

$$\tilde{\boldsymbol{f}}_r\left(\xi\left(j\right)\right)=\left[\frac{1}{\Delta t}\left(\hat{\boldsymbol{x}}_r\left(t_{j_1+1}\right)-\hat{\boldsymbol{x}}_r\left(t_{j_1}\right)\right)\Big|_{\hat{\boldsymbol{x}}\left(t_{j_1}\right)=\xi(j)},\cdots,\frac{1}{\Delta t}\left(\hat{\boldsymbol{x}}_r\left(t_{j_{K_j}+1}\right)-\hat{\boldsymbol{x}}_r\left(t_{j_{K_j}}\right)\right)\Big|_{\hat{\boldsymbol{x}}\left(t_{j_{K_j}}\right)=\xi(j)}\right] \tag{18}$$

$$\tilde{\boldsymbol{\Phi}}=\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}}\otimes\boldsymbol{\Phi}_1\\ \vdots\\ \boldsymbol{a}_N^{\mathrm{T}}\otimes\boldsymbol{\Phi}_N\end{bmatrix},\quad \boldsymbol{a}_i=\underbrace{\left[1\cdots1\right]}_{K_i} \tag{19}$$

式中$\otimes$是克罗内克积；$\hat{\boldsymbol{x}}_r$是$\hat{\boldsymbol{x}}$的第r项；$\boldsymbol{\Phi}_i$是$\boldsymbol{\Phi}$的第i行；$\tilde{\boldsymbol{\varepsilon}}$服从高斯分布$\mathcal{N}\left(0,\sigma^2\boldsymbol{I}\right)$，这里$\boldsymbol{I}$是一个$\sum_{i=1}^{N}K_i\times\sum_{i=1}^{N}K_i$单位矩阵。

证明：参见附录A中的第S2节。

备注1. 对于任何置换矩阵$\boldsymbol{P}$，有$\boldsymbol{P}\tilde{\boldsymbol{Y}}=\boldsymbol{P}\tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\theta}}+\boldsymbol{P}\tilde{\boldsymbol{\varepsilon}}$。因此，$\tilde{\boldsymbol{Y}}$的元素可以按时间顺序重新排列。

定理1给出了一种通过考虑方程（12）的替代回归方程（16）来辨识$\boldsymbol{\theta}$的构造性方法，其优点显著。首先，方程（16）意味着“导数”可以直接用作输出，从而避免了对分箱操作的需要减少了所需的数据量并避免实际中的维数灾难。其次，与分箱操作不同，当构造字典矩阵和输出向量的时候，$\hat{\boldsymbol{x}}\left(t_{j_s}\right)\in(\xi-\delta,\xi+\delta)$的值被保留而不是被替换为$\xi$，这避免了相应的近似误差。最后，实验也表明所提方法减少了所需数据量并且对高维SDE具有鲁棒性。

2.3. 推断扩散项

一旦成功地辨识出漂移项，则可以建立回归方程辨识扩散项。根据方程（6），考虑

$$\frac{\sum_{s=1}^{K}\left\{\begin{array}{c}\left[\hat{\boldsymbol{x}}\left(t_{j_s+1}\right)-\hat{\boldsymbol{x}}\left(t_{j_s}\right)-f\left(\hat{\boldsymbol{x}}\left(t_{j_s}\right)\right)\Delta t\right]\cdot\\ \left[\hat{\boldsymbol{x}}\left(t_{j_s+1}\right)-\hat{\boldsymbol{x}}\left(t_{j_s}\right)-\boldsymbol{f}\left(\hat{\boldsymbol{x}}\left(t_{j_s}\right)\right)\Delta t\right]^{\mathrm{T}}|\hat{\boldsymbol{x}}\left(t_{j_s}\right)=\xi\end{array}\right\}}{K\Delta t}\approx\boldsymbol{G}\left(\xi\right)+\boldsymbol{\varepsilon}_2 \tag{20}$$

式中，$\boldsymbol{\varepsilon}_2$是一个矩阵，其中每个元素均服从均值为零、方差与$1/K$成比例的高斯分布。辨识的漂移项可以用于估计相应的精确漂移项的值。遵循与漂移项类似的辨识路线，可以建立回归方程来辨识扩散项，并利用SBL方法求解。注意，这里基函数的选择可以不同于用于辨识漂移项的基函数。

2.4. 模型验证

接下来，通过仿真和实验对所提出的BISDE算法进行了评估。对于仿真系统，由于确切的模型结构已知，所以可以通过比较测量值处的真实漂移项（扩散项）和拟合漂移项（扩散项）生成的数据之间的均方误差（MSE）来评估辨识模型的性能。对于真实的实验数据，由于缺乏基准信息，设计了三个准则来评估辨识模型的性能。

准则1： 如果存在一个被广泛接受的模型，可以利用该模型来评估辨识模型。如果发现它们之间具有良好的一致性，则通过交叉验证验证了该模型的有效性。否则，明显的不一致性意味着有可能确定一个更有效的模型，其性能可通过准则2或准则3进行评估。

准则2： 如果实际的动力过程近似平稳，则可以将通过求解平稳Fokker-Planck方程得到的辨识模型的解析概率密度函数（PDF）与时间序列的经验PDF进行比较。它们之间的良好匹配则表明辨识模型的可靠性。非平稳过程可以通过构造代数或对数增量或其他方法转换为平稳过程[28]。

准则3： 对于平稳过程，当PDF无法解析求解或数值估计时，可以给定真实数据的初始值，利用辨识的模型生成仿真数据，然后通过比较仿真数据的PDF与真实数据的PDF来评估辨识模型的性能。

3. 结果

3.1. 利用仿真数据发现朗之万方程

首先将提出的BISDE算法应用于朗之万方程（见

附录A中的第S3.1节），该方程在物理学中起着重要作用[29–30]，具体如图1所示。尽管其数学形式很简单，但物理学家花了将近一百年的时间才发现该方程。具体而言，朗之万方程描述了布朗粒子在所有时间尺度上的动力学特性，克服了参考文献[1]中爱因斯坦理论的缺点。考虑一个浸入流体中的布朗粒子[图1（a）]。由于热运动，布朗粒子与各个方向的液体分子产生碰撞从而形成速度遵循朗之万方程的随机运动。

通过采用时间步长$\Delta t = 0.04$的均匀离散区间[0, 1000]获得数据。基函数由常数项、关于$\boldsymbol{x}$高达15次的多项式和指数为从$\boldsymbol{x}$到$10\boldsymbol{x}$的指数函数组成。在该示例中，相同的基函数被用于漂移项和扩散项辨识。如果与漂移项或扩散项相关的先验信息（如对称性和周期性）是已知的，则应指定每一项的基函数。图1（b）、（c）表明BISDE成功地以非常高的精度辨识了朗之万方程。该朗之万方程的示例表明了BISDE可以帮助物理学家从数量相对有限的时间序列数据中辨识随机现象的潜在SDE的能力。

3.2. 利用实验数据发现轴承振动的动力学

接下来，展示了美国凯斯西储大学（Case Western Reserve University, CWRU）轴承数据集滚动轴承振动的动力学发现。随着现代工业的快速发展，旋转机械在制造系统和家用电器中得到了广泛的应用。虽然滚动轴承在这些机械中得到了广泛和不可或缺的应用，但它也被列为与机械缺陷相关的首要部件[31–33]。轴承故障将减小机器寿命和性能，降低工件的质量，极端情况下还会造成安全隐患甚至人员伤亡。因此，轴承故障诊断已成为工程界的热门话题。一般而言，振动信号被认为是评估轴承缺陷的最有用数据，因为轴承中的任何故障都可能影响振动动力学[34]。因此，在无故障和有故障阶段确定轴承振动信号的动力学特性，可以提供关于潜在轴承缺陷的知识。

CWRU数据集是一个用于探索正常和故障轴承动力学特性的开源数据集。原始试验台及其横截面图分别如图2（a）、（b）所示。试验台由电机、编码器或扭矩传感器及测力计组成。数据是以48 kHz的功率在三种不同的状态下采集的：①正常轴承（NB）；②内圈故障（IRF）；③滚珠故障（BF）。将单点故障引入驱动端和风扇端轴承，故障直径分别为7 mil、14 mil和21 mil（1 mil = 0.0254 mm）。将每个故障轴承重新安装到电机中，并在电机负荷为1~3 hp（马力；1 hp = 0.7457 kW）下进行测试。

本文重点介绍了1 hp下正常和故障风扇端轴承动力学辨识的结果；更全面的比较可参见附录A中的第S3.2节。因为有足够的数据量，所以分箱操作可以为漂移项和扩散项提供合理的估计。此外，此估计值可以作为BISDE的辨识结果的验证基准。为了说明本文提出的方法可以减少所需的数据量，BISDE只使用了大约十分之一的数据集来发现潜在的动力学方程。图2（c）~（h）给出了在不同的操作条件下分箱操作估计的漂移项和扩散项与通过BISDE辨识的相应漂移项和扩散项之间的比较；使用MSE评估辨识模型的性能。实验结果表明所辨识的漂移项和扩散项可以准确地捕获轴承振动信号的动力学行为。不同电机负载下的正常轴承的辨识模型可以帮助操作员或从业者进行故障的早期诊断，以防止灾难性后果并减少维护费用。

3.3. 发现经典的随机模型和应用

最后，应用所提出的BISDE算法确定了多个经典的和现实世界中的SDE。仿真模型基于常见的物理系统和随

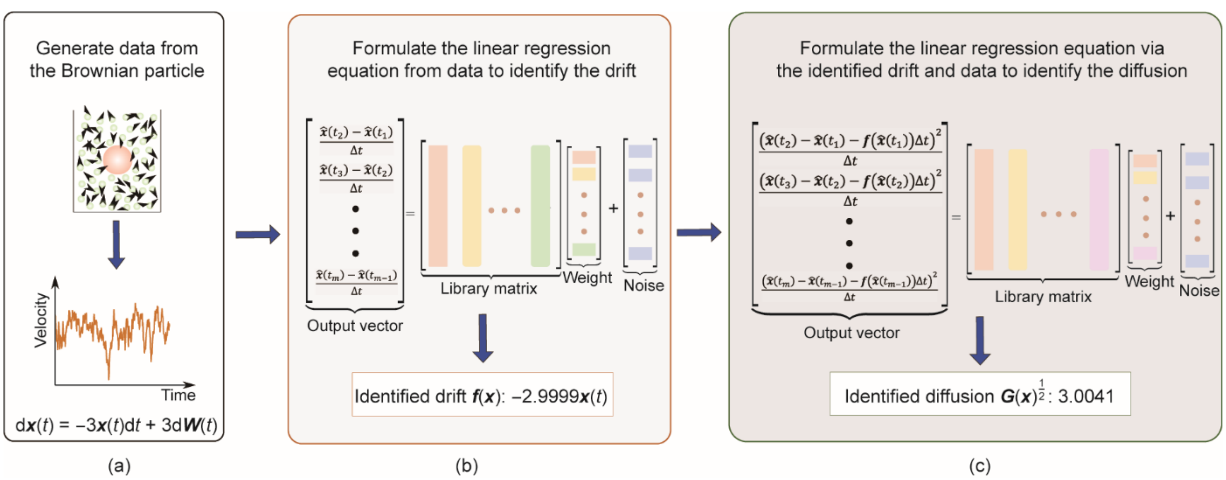

图1. 朗之万方程的BISDE的流程示意图。（a）浸没在流体中的布朗粒子（红点），其速度满足朗之万方程；（b）、（c）分别建立线性回归方程，从候选基函数张开的空间中辨识漂移项和扩散项。

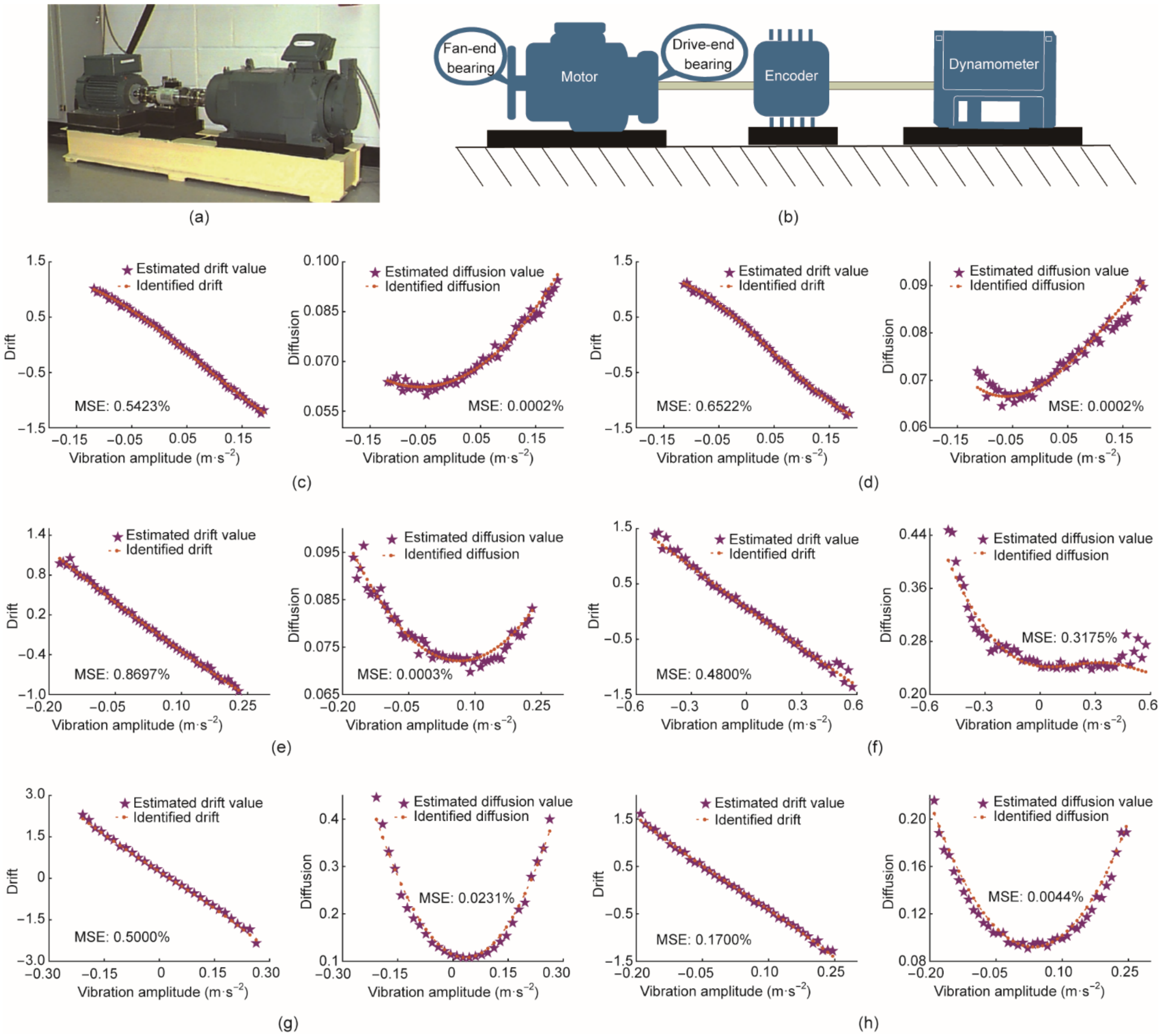

图2. 发现正常和故障风扇端轴承的动力学特性。（a）CWRU数据集轴承实验平台；（b）原始试验台的横截面图，包括电机（左）、扭矩传感器/编码器（中）和测力计（右）；（c）、（d）分别为1 hp和2 hp的正常轴承的辨识结果；（e）、（f）分别具有7 mil故障直径的滚珠故障和内圈故障的辨识结果（电机负荷：1 hp）；（g）、（h）分别具有14 mil故障直径的滚珠故障和内圈故障的辨识结果（电机负荷：1 hp）。

机过程（见附录A中的第S3.3~S3.5节）。通过一个二维仿真模型验证了BISDE算法在有限数据量下辨识多维SDE的能力。实际系统包括股票价格波动、风速和油价（见附录A中的第S3.6~S3.8节）。为这种随机动力系统建立一个辨识框架，可以帮助从业者改进系统设计并为不同场景开发更有效的系统管理策略。更详细的说明见下文和附录A中的第S3节。此外，数据和代码实现可参见https://github.com/HAIRLAB/BISDE。

图3总结了待辨识的仿真和实际系统。每个类别的三个示例都用特定的背景颜色标记。第一行和第四行分别显示了仿真和实际系统。第二行显示了用颜色表示其概率密度值的模拟样本路径。第五行由于缺乏基准信息仅显示实际样本路径。第三行和第六行评估辨识模型的性能。对于仿真系统，可以比较测量点处由真实和辨识的漂移项/扩散项生成的数据点之间的MSE，以评估其性能。对于实际系统，采用准则1、准则2和准则3来评估辨识模型的性能。

与最先进方法的比较：为了证明BISDE可以从相对有限数量的时间序列数据中辨识SDE，将BISDE与参考文献[17]中最先进方法（为方便起见，以下称为SDE_SINDy）进行了比较。图4显示了BISDE和SDE_SINDy在仿真系统上的直观比较。当数据量相对有限时，SDE_SINDy未能辨识所有情况下的潜在模型结构，而BISDE产生近乎完美的辨识结果。

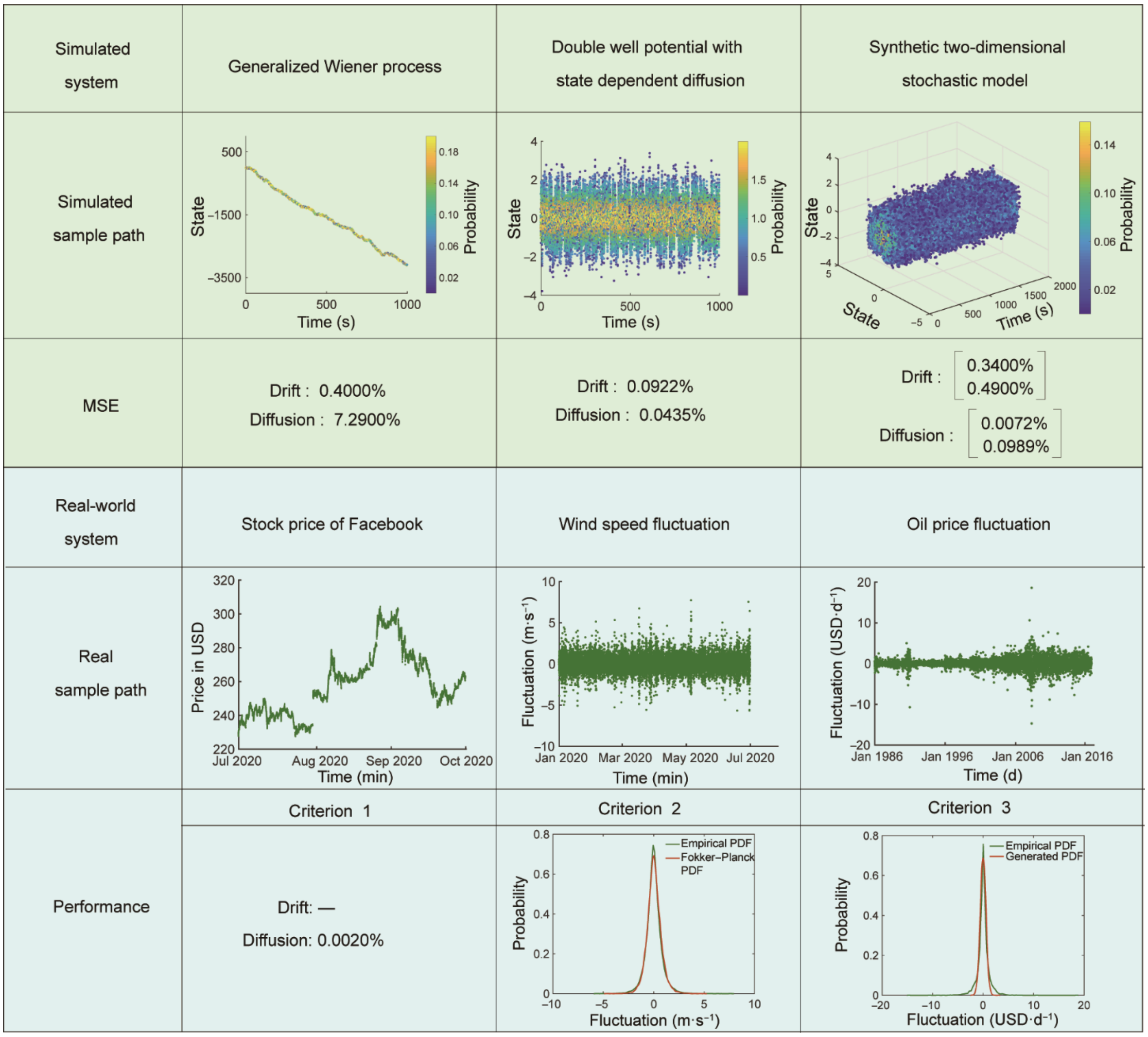

图3. BISDE算法应用于诸多实例的总结。BISDE已经在三个仿真系统和三个实际系统上进行了测试，包括一个二维仿真系统，其中每种类型都用特定的背景颜色标记。

金融经济学–股票价格：股票价格受许多经济、金融和政治因素的相互影响。股票价格的动态变化可以被认为是一个随机过程，因为随机噪声在预测未来股票价格时引入了不确定性。20世纪70年代，获得2013年诺贝尔经济学奖的Eugene Fama提出了弱式有效市场假说，指出未来股票价格不能通过分析历史数据来预测[35]。基于这一假设，通常假定股票价格遵循马尔可夫过程。冗长的历史数据序列对确定股票价格的动态变化是无益的，因为这些动态变化会随时间的推移而变化[3]。因此，本文仅收集了2020年7月1日至2020年9月30日三个月内每分钟的Facebook股票价格数据[图5（a)、(b）]。

应用提出的BISDE算法后[图5（c）]，图5（d）显示了描述Facebook股票价格行为的几何布朗运动模型。几何布朗运动是用于描述股票价格行为的最广泛的模型[3]，也是用于推导Black-Scholes-Merton公式以定价欧式看涨和看跌期权的假设之一[36–37]。令人惊讶的是，辨识出的波动率（0.4039）与使用参考文献[3]中建议的方法估计的波动率（0.4087）几乎一致，证明辨识模型的准确性。通过与不同银行提供的一年期年收益率（APY）相比[图5（e）]，可以推断出购买Facebook股票的预期回报是在指定银行存钱回报的数倍。基于辨识模型和一年期年收益率，投资者可以选择一种策略来平衡预期回报和股票的不确定性或风险，以获得更好的收益[图5（f)、(g）]。总的来说，该应用揭示了BISDE是一个用于辨识股票动态行为的强有力工具。辨识模型可以为投资者提供对股票的见解，并帮助证券交易所确定股票期权的价格[3]。

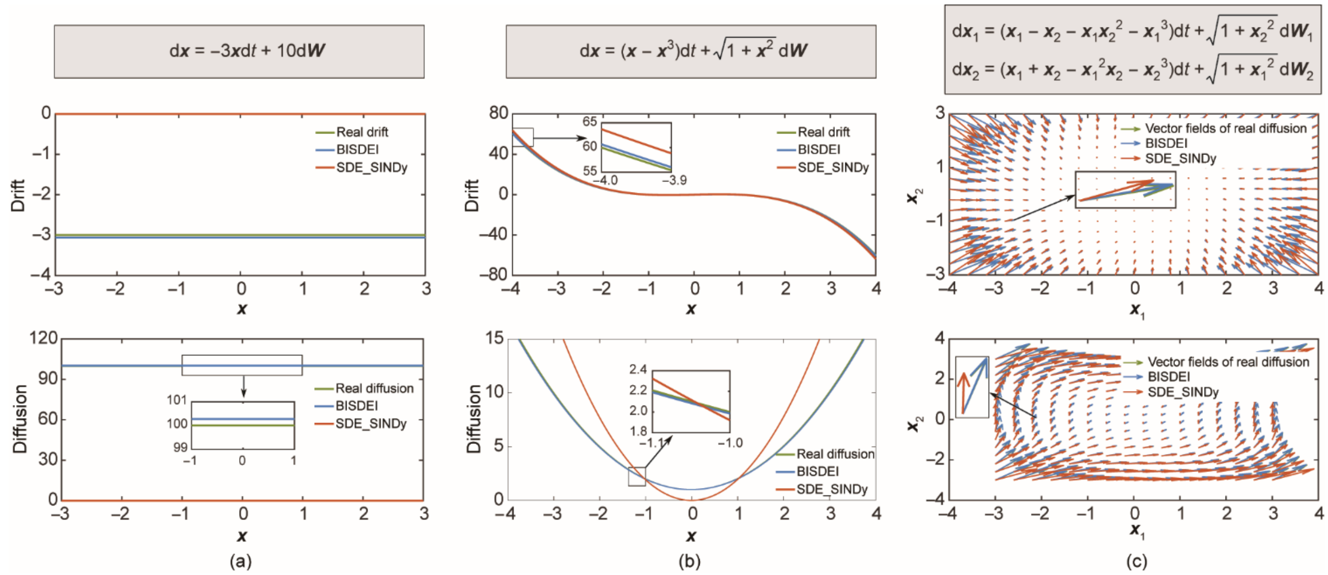

图4. BISDE和SDE_SINDy在仿真系统上的比较。BISDE和SDE_SINDy在广义Winner过程（a）、具有状态相关的双井势扩散过程（b），以及合成二维随机模型（c）上的辨识结果。

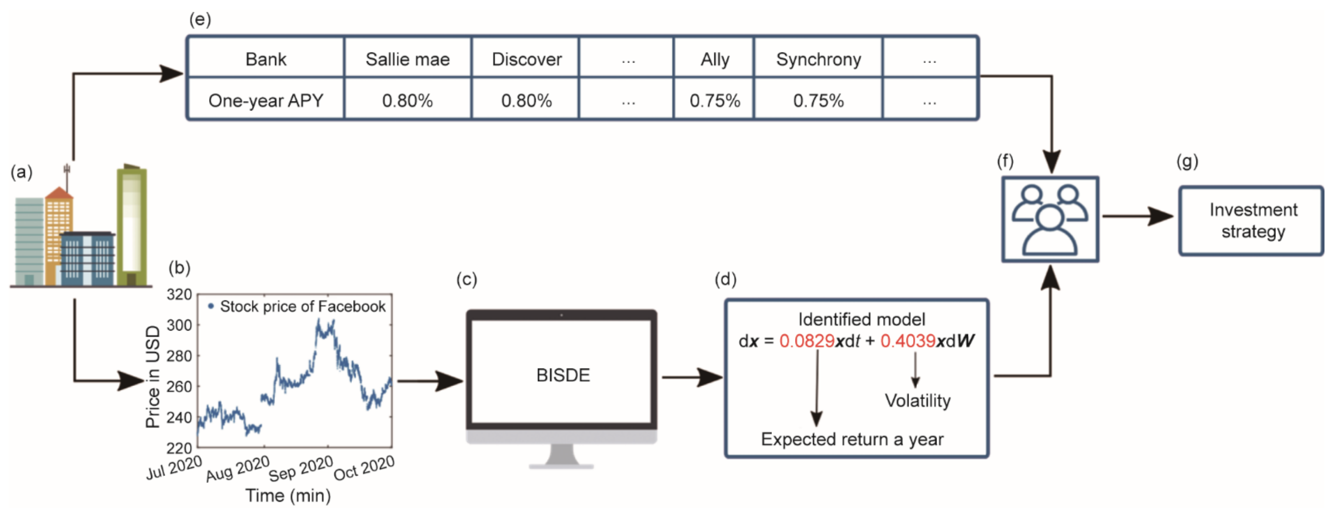

图5. 发现Facebook的股票价格行为。（a）金融机构（数据采集）；（b）Facebook三个月的股票价格；（c）股票价格数据分析；（d）用于描述Facebook股票价格行为的辨识的SDE；（e）2020年8月美国不同银行一年期年收益率（APY）汇总；（f）投资者或技术分析师；（g）投资者通过整合投资信息制定的投资策略。

电力系统–风速波动：作为一种分布广泛、可持续的再生能源，风能在许多国家的电网中发挥着重要作用。截至2020年年底，全球风力发电量约占总发电量的5%。由于风速在空间和时间尺度下均不可控且波动，因此即使年总发电量几乎保持不变，但风电场每一分钟的发电量都不一样。风速波动会影响额定功率输出，有时会破坏电能质量和可靠性，并导致极端风力涡轮机疲劳载荷[7,38–39]。因此，确定风速波动的动力学特性对于电力生产和负荷设计至关重要，以确保风能资源的安全性和经济性。

为了说明BISDE对该问题的适用性，本研究收集了2020年上半年新西兰惠灵顿Greta Point Cws（位于174.80574°E、41.30243°S、平均海平面以上3 m）的风速数据[40]。通过计算风速数据的差异获得风速波动，从而将非平稳序列（风速）转换为平稳序列（风速波动）。可以将辨识的模型看作一个含有二次状态依赖项的Ornstein-Uhlenbeck过程。然后，通过求解稳态Fokker-Planck方程，可以得到辨识模型的解析PDF。通过解析的PDF和由测量值生成的经验PDF之间的高相似性验证了模型的性能（图3）。BISDE成功地辨识了用于描述风速波动的SDE。通过将辨识模型嵌入风力涡轮机模型中，可以进行动态研究，从而实现对风力涡轮机的控制[41–42]。

能源经济–油价波动：尽管可再生能源的重要性日益增加，但石油仍然是大多数国家的主导资源。与其他所有商品相比，石油对宏观经济具有相当大的影响，这是因为

石油在国际市场上具有不可替代性和高度流动性[43–44]。油价波动会给出口国和进口国造成重大损失或带来重大利润。此外，油价波动可能导致通货膨胀、运输成本增加和商业政策变化，不仅影响工业，而且影响个人和家庭的日常生活。由于油价受多个独立和相互关联的因素影响，因此建立准确的模型极具挑战性。

本研究应用BISDE算法来解决这个具有挑战性的问题。为此，使用从美国能源情报署（US Energy Information Administration）收集的1986年年初至2017年年中的原油价格。首先，计算了油价的每日变化。在确定油价波动后，应用BISDE发现潜在的动力学行为。为了说明辨识模型的有效性，基于所辨识的SDE生成了一个样本路径，该样本路径具有与真实数据相同的初始值。然后，比较了实际数据和仿真数据的经验PDF。辨识模型产生的PDF非常接近高概率区域中真实数据的PDF（图3）。在已知当前油价波动的情况下，辨识模型可以帮助预测下一次油价波动，并从企业和决策者的角度刻画相应的不确定性，有助于避免不必要的损失。

4. 讨论

本研究提供了一种称为BISDE的新参数化算法，用于发现SDE控制的系统。该算法具有几个显著的优点。第一，与不能提供可解释模型的非参数化方法[13]不同，该算法可以辨识漂移项和扩散项的模型结构，以确定随机动力系统的潜在机理。第二，现有的最先进辨识SDE的数据驱动算法将多维系统投影到低维系统，以减少所需的数据量[17–18]。而BISDE则可以在有限的数据量下直接发现原始的多维系统。这种能力通过辨识二维SDE得到了验证，并且可以自然地推广到高维SDE。第三，虽然BISDE和之前的研究一样采用了Euler-Maruyama离散化方法[10,13–14]，但本研究探讨了漂移项和扩散项的限制条件，进行了待辨识项和实际项之间的收敛性分析，并将非参数估计与参数辨识连接起来。总体而言，BISDE是一种利用相对有限的时间序列数据来辨识随机动力系统潜在动力学行为的可行方法，并且具有帮助研究人员对自然现象和受不同来源随机噪声干扰的工程系统进行建模的潜力。

尽管BISDE具有诸多优势，但仍存在一些问题。首先，先验知识有助于选择基函数，使得模型结构的辨识更快且更准确。然而，缺乏先验信息，难以建立函数库。更糟糕的是，真实的模型可能会被多项式、核函数和其他形式的函数近似逼近[22]。此外，在不知道确切模型结构的情况下，建立非平稳过程的验证方法也是一个问题，因为在许多情况下，需要直接探索状态变量的内在的动力学行为。最后，两个不同的SDE可能产生相同的PDF（见附录A中的第S4.1节），在这种情况下，它们虽然具有相同的统计特性，但涉及两种完全不同的机制，这可能会误导研究人员做出错误的决断。

致谢

本工作得到国家重点研发计划(2018YFB1701202)、国家自然科学基金项目(92167201和51975237)、中央高校基础研究基金项目(华中科技大学:2021JYCXJJ028)的资助。

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, StLück, 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. 3th 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.