# Sparse identification method of extracting hybrid energy harvesting system from observed data

Ya-Hui Sun(孙亚辉)[1,2], Yuan-Hui Zeng(曾远辉)[1], and Yong-Ge Yang(杨勇歌)[1,†]

[1] *School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou* 510520, *China*
[2] *State Key Laboratory for Strength and Vibration of Mechanical Structures, Xi'an Jiaotong University, Xi'an* 710049, *China*

(Received 15 August 2022; revised manuscript received 13 October 2022; accepted manuscript online 21 October 2022)

Hybrid energy harvesters under external excitation have complex dynamical behavior and the superiority of promoting energy harvesting efficiency. Sometimes, it is difficult to model the governing equations of the hybrid energy harvesting system precisely, especially under external excitation. Accompanied with machine learning, data-driven methods play an important role in discovering the governing equations from massive datasets. Recently, there are many studies of data-driven models done in aspect of ordinary differential equations and stochastic differential equations (SDEs). However, few studies discover the governing equations for the hybrid energy harvesting system under harmonic excitation and Gaussian white noise (GWN). Thus, in this paper, a data-driven approach, with least square and sparse constraint, is devised to discover the governing equations of the systems from observed data. Firstly, the algorithm processing and pseudo code are given. Then, the effectiveness and accuracy of the method are verified by taking two examples with harmonic excitation and GWN, respectively. For harmonic excitation, all coefficients of the system can be simultaneously learned. For GWN, we approximate the drift term and diffusion term by using the Kramers–Moyal formulas, and separately learn the coefficients of the drift term and diffusion term. Cross-validation (CV) and mean-square error (MSE) are utilized to obtain the optimal number of iterations. Finally, the comparisons between true values and learned values are depicted to demonstrate that the approach is well utilized to obtain the governing equations for the hybrid energy harvester under harmonic excitation and GWN.

**Keywords:** data-driven, hybrid energy harvester, harmonic excitation, Gaussian white noise

**PACS:** 02.50.–r, 05.40.–a, 05.40.Ca **DOI:** 10.1088/1674-1056/ac9cbf

## 1. Introduction

The vibration energy harvester (VEH) can produce electric energy from ambient vibrations[1–3] to provide wireless sensors with sustaining energy. There are many types of energy harvesting techniques, such as piezoelectric,[4,5] electromagnetic,[6] etc.[7,8] Among them, combining piezoelectric components and electromagnetic elements, hybrid energy harvesters were investigated and widely applied due to the advantage of improving the harvesting efficiency.[9] Panyam and Daqaq[10] studied a tri-stable hybrid energy harvester under harmonic excitation by using multi-scales method. Karami and Inman[11] proposed an approximation method to explore the hybrid energy harvester under harmonic excitation. Xia *et al.*[12] investigated the performances of the energy harvesters with different boundary conditions. Moreover, noises exist in the real environment and result in different dynamical behaviors.[13,14] Therefore, hybrid energy harvesters under random excitation were studied in previous researches.[15–20] For example, Zhou *et al.*[15] showed that the performance of hybrid energy harvesters under Gaussian excitation is improved by adding nonlinear components. Sengha *et al.*[16] modeled a hybrid energy harvesting system driven by harmonic excitation and colored noise. Foupouapouognigni *et al.*[17] indicated that the performance of the hybrid VEH, which was influenced by Gaussian white noise (GWN) and harmonic excitation, was improved. Sun *et al.*[18] used the stochastic averaging technique to analyze the stochastic responses of a fractional-order hybrid VEH driven by GWN. The influences of colored noise excitation on hybrid VEH were investigated by Yang and Cao.[20]

Stochastic responses of hybrid VEH in the above articles were studied in the condition that the governing equations were firstly known. However, for a complex system, it is trouble to model the governing equations precisely in practical applications. To overcome this difficulty, in recent years, data-driven modeling[21,22] was presented and applied in different areas, such as fluid dynamics, meteorology, finance, etc.[23–26] With the developments of machine learning and data science, some methods were proposed to identify elusive dynamical systems. For deterministic differential equations, Brunton *et al.*[27] proposed a data-driven method called sparse identification of nonlinear dynamics (SINDy), and demonstrated that the method is well agreement in the deterministic differential equations by using the iterative threshold algorithm. Boninsegna *et al.*[28] modified the SINDy to avoid the adjustment of the threshold parameter in Ref. [27]. However, stochastic differential equations widely exist in practical applications. For stochastic differential equations (SDEs), the methods which

†Corresponding author. E-mail: yonggeyang@163.com
© 2022 Chinese Physical Society and IOP Publishing Ltd 

were devised in Refs. [27,28] were also used to model SDEs driven by GWN. Rudy *et al.*[29] presented a deep neural network approach to estimate the coefficients and the measurement noise simultaneously, and showed good robustness of the method by increasing the noise level. Dai *et al.*[30] proposed the maximum likelihood estimation to learn the SDE under fractional Brownian motion, and the results showed a good accuracy for true values. Lu and Lermusiaux[31] devised a Bayesian learning technique to model stochastic dynamical systems. Huang and Li[32] used the SINDy to discover the equations of a four-dimensional stochastic projectile system. Wu *et al.*[33] obtained the mean residence time and escape probability of SDEs from data by using the devised approach. Additionally, based on the Koopman generator, Zhang *et al.*[34] developed an approach to extract SDEs influenced by Lévy noise from data on mean exit time, and indicated that this method can also apply to dynamical systems under GWN. Lu and Duan[35] discovered the SDEs from data with Lévy noise by utilizing extended dynamic mode decomposition, and acquired the transition probability density functions by solving the Fokker–Planck equations. Together with the Kramers–Moyal formulas and SINDy, governing laws under different Lévy noise were extracted from the observed data of stochastic dynamics equations.[36–38] Among these methods, SINDy has been widely applied in discovering governing equations from massive datasets. The method combines the least-squares and compressed-sensing to solve the sparse coefficients of the equations, so that the approximate governing equations can be obtained. It is useful to analyze the subsequent dynamical behavior of the system.

To the best of our knowledge, few authors pay attention to discovering the equations from data for the hybrid energy harvester in present. In this paper, we develop a sparse identification approach to identify the equations for the hybrid VEH. The framework of this paper is organized as follows. In Section 2, the mathematical model of the hybrid VEH with non-dimensional form is given. In Section 3, a sparse identification process is developed to solve the sparse coefficients for the hybrid VEH. In Sections 4 and 5, two examples of the hybrid VEH are taken to examine the validity of the devised method. In Section 6, some conclusions are remarked.

## 2. Hybrid VEH under noise excitation

A family of hybrid electromagnetic and piezoelectric energy harvesters are considered as shown in Fig. 1. The model is simplified as a mass–spring–damper system in Fig. 1(a) coupled with a piezoelectric circuit in Fig. 1(b) and an electromagnetic circuit in Fig. 1(c) under base acceleration.

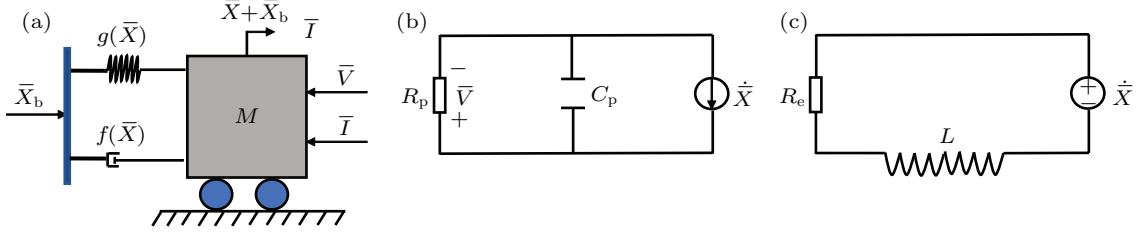



**Fig. 1.** (a) Simplified diagram of a hybrid VEH coupled with (b) a piezoelectric energy harvesting circuit and (c) an electromagnetic circuit.

The coupling equations of the hybrid energy harvesting system are given by

$$\begin{aligned}
&M\ddot{\bar{X}}(\tau)+f(\bar{X})\dot{\bar{X}}(\tau)+g(\bar{X})+\bar{\zeta}_1\bar{V}(\tau)+\bar{\zeta}_2\bar{I}(\tau)=-M\ddot{\bar{X}}_b(\tau),\\
&C_p\dot{\bar{V}}(\tau)+\frac{\bar{V}(\tau)}{R_p}=\bar{\zeta}_1\dot{\bar{X}}(\tau),\\
&L\dot{\bar{I}}(\tau)+(R_c+R_e)\bar{I}(\tau)=\bar{\zeta}_2\dot{\bar{X}}(\tau).
\end{aligned}\tag{1}$$

Here, $\bar{X}$, $\dot{\bar{X}}$ and $\ddot{\bar{X}}_b$ denote the displacement, the velocity and the base acceleration of the mass $M$, respectively. $\bar{V}$ is the electric voltage measured across the equivalent resistance load $R_p$. $C_p$ and $\bar{\zeta}_1$ are the the piezoelectric capacitance and the piezoelectric coupling coefficient. $L$, $\bar{I}$ and $\bar{\zeta}_2$ denote the inductance of the coil, the output current and the electromagnetic coupling coefficient, respectively. $R_c$ and $R_e$ successively denote the load resistances of the electromagnetic and the resistance of the coil. $f(\bar{X})$ and $g(\bar{X})$ represent the damping term and the stiffness term.

Then, we make the equations dimensionless by means of a transformation[18,39]

$$\begin{aligned}
&t=\omega\tau,\quad x=\frac{\bar{X}}{l_c},\quad x_b=\frac{\bar{X}_b}{l_c},\quad V=C_p\bar{V},\quad I=L\bar{I},\\
&f(x)=\frac{f(\bar{X})}{M\omega},\quad g(x)=\frac{g(\bar{X})}{M\omega^2 l_c},\\
&\zeta_1=\frac{\bar{\zeta}_1}{M\omega^2 l_c C_p},\quad \zeta_2=\frac{\bar{\zeta}_2}{M\omega^2 l_c L},\\
&\mu_1=\frac{1}{\omega R_p C_p},\quad \mu_2=\frac{R_c+R_e}{\omega L},\quad \lambda_1=\bar{\zeta}_1,\quad \lambda_2=\bar{\zeta}_2.
\end{aligned}\tag{2}$$

Subsequently, the dimensionless equations of the hybrid VEH can be expressed as follows:

$$\begin{aligned}
&\ddot{x}+f(x)\dot{x}+g(x)+\zeta_1 V+\zeta_2 I=-\ddot{x}_b,\\
&\dot{V}+\mu_1 V=\lambda_1\dot{x},\\
&\dot{I}+\mu_2 I=\lambda_2\dot{x}.
\end{aligned}\tag{3}$$

Here, $f(x)$ and $g(x)$ are the non-dimensional damping force and stiffness term. $\zeta_1$, $\zeta_2$, $\lambda_1$ and $\lambda_2$ are the non-dimensional coupling coefficients in Eq. (1). $\mu_1$ is the reciprocal of the product of resistance and capacitance. $\mu_2$ represents the ratio of resistance and inductance.

As examples of different external excitation, harmonic excitation and GWN[18,40] are considered in Sections 4 and 5. With the method of stepwise sparse regressor (SSR),[28,41] we devise the method of identified sparse regression to discover the governing equations by learning the coefficients of the formula.

## 3. Sparse identification for hybrid VEH

Drawing on the ideas of machine learning, we combine the least-square sense, SSR algorithm and cross-validation (CV). A sparse identification process is developed to learn the unknown coefficients in the equations. Assume that we have observed $N$ data points of system state time series of displacement, voltage and output current denoted by $\boldsymbol{X}$, $\boldsymbol{V}$ and $\boldsymbol{I}$, respectively. Meanwhile, data points $\boldsymbol{X}$, $\boldsymbol{V}$ and $\boldsymbol{I}$ have been dimensionless to $[x_1,x_2,x_3,\ldots,x_N]$, $[v_1,v_2,v_3,\ldots,v_N]$ and $[i_1,i_2,i_3,\ldots,i_N]$ at $[t_1,t_2,t_3,\ldots,t_N]$. We transform Eq. (3) into the following differential equations:

$$\begin{aligned}
\mathrm{d}x &= y\mathrm{d}t \triangleq b_1\mathrm{d}t,\\
\mathrm{d}y &= -\left[f(x)\dot{x}+g(x)+\zeta_1 V+\zeta_2 I\right]\mathrm{d}t+\xi(t)\mathrm{d}t\\
&\triangleq b_2\mathrm{d}t+\xi(t)\mathrm{d}t,\\
\mathrm{d}V &= (-\mu_1 V+\lambda_1 y)\,\mathrm{d}t \triangleq b_3\mathrm{d}t,\\
\mathrm{d}I &= (-\mu_2 I+\lambda_2 y)\,\mathrm{d}t \triangleq b_4\mathrm{d}t,
\end{aligned} \tag{4}$$

where $\xi(t)=F\cos(t)+W(t)$ is the harmonic excitation and Gaussian white noise. For the GWN, it satisfies

$$E[\xi(t)]=0,\quad E[\xi(t)\xi(t')]=2D\delta(t-t'), \tag{5}$$

where $\delta(t-t')$ represents the Dirac delta function and $D$ denotes the noise intensity. Then, $\xi(t)\mathrm{d}t=F\cos(t)\mathrm{d}t+\sqrt{2D}\mathrm{d}W(t)$, where $W(t)$ is the standard Winner process.

### 3.1. Learned process of drift term

In this subsection, we introduce the sparse identification of drift term $\boldsymbol{b}=[b_1,b_2,b_3,b_4]^{\mathrm{T}}$. Firstly, we approximate the first-order derivative by using the first-order difference, i.e.,

$$\begin{aligned}
\dot{x}_k &= \frac{x_{k+1}-x_k}{\Delta t},\quad \dot{y}_k=\frac{\dot{x}_{k+1}-\dot{x}_k}{\Delta t},\\
\dot{v}_k &= \frac{v_{k+1}-v_k}{\Delta t},\quad \dot{i}_k=\frac{i_{k+1}-i_k}{\Delta t}.
\end{aligned} \tag{6}$$

The derivative of the system state with respect to $t$ can be calculated. Subsequently, we obtain the datasets as follows:

$$\begin{aligned}
&\boldsymbol{X}=[x_1,x_2,x_3,\ldots,x_N],\\
&\dot{\boldsymbol{X}}=[\dot{x}_1,\dot{x}_2,\dot{x}_3,\ldots,\dot{x}_{N-1}]\triangleq[y_1,y_2,y_3,\ldots,y_{N-1}],\\
&\boldsymbol{Y}=[y_1,y_2,y_3,\ldots,y_{N-1}],\quad \dot{\boldsymbol{Y}}=[\dot{y}_1,\dot{y}_2,\dot{y}_3,\ldots,\dot{y}_{N-2}],\\
&\boldsymbol{V}=[v_1,v_2,v_3,\ldots,v_N],\quad \dot{\boldsymbol{V}}=[\dot{v}_1,\dot{v}_2,\dot{v}_3,\ldots,\dot{v}_{N-1}],\\
&\boldsymbol{I}=[i_1,i_2,i_3,\ldots,i_N],\quad \dot{\boldsymbol{I}}=[\dot{i}_1,\dot{i}_2,\dot{i}_3,\ldots,\dot{i}_{N-1}].
\end{aligned} \tag{7}$$

According to the feature of unknown coefficient term in Eq. (4), a series of basis functions called dictionary is conducted as

$$\begin{aligned}
&\boldsymbol{\Phi}(\boldsymbol{t},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{V},\boldsymbol{I})\\
&=\left[\phi_1(\boldsymbol{t},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{V},\boldsymbol{I}),\ldots,\phi_K(\boldsymbol{t},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{V},\boldsymbol{I})\right].
\end{aligned} \tag{8}$$

The learned results will be better if we select an abundant type of basis functions, while the real workload is enormous and polynomial basis functions are enough precise for most cases.

By referring to Ref. [42], $b_2$, $b_3$ and $b_4$ are estimated by using the Kramers–Moyal formula, i.e.,

$$\begin{aligned}
b_2 &= \lim_{\Delta t\to 0}\frac{y_{k+1}-y_k}{\Delta t},\quad b_3=\lim_{\Delta t\to 0}\frac{v_{k+1}-v_k}{\Delta t},\\
b_4 &= \lim_{\Delta t\to 0}\frac{i_{k+1}-i_k}{\Delta t}.
\end{aligned} \tag{9}$$

Substituting Eqs. (6)–(9) into Eq. (4), we obtain

$$\dot{y}\approx\sum_{j=1}^{K}u_{j1}\phi_j,\quad \dot{V}\approx\sum_{j=1}^{K}u_{j2}\phi_j,\quad \dot{I}\approx\sum_{j=1}^{K}u_{j3}\phi_j. \tag{10}$$

Then, the array of unknown coefficients $\boldsymbol{u}$ can be derived by

$$\boldsymbol{A}\boldsymbol{u}=\boldsymbol{B}, \tag{11}$$

where

$$\boldsymbol{A}=\begin{bmatrix}
\phi_1(t_1,x_1,\dot{x}_1,v_1,i_1) & \cdots & \phi_K(t_1,x_1,\dot{x}_1,v_1,i_1)\\
\vdots & \ddots & \vdots\\
\phi_1(t_{N-2},x_{N-2},\dot{x}_{N-2},v_{N-2},i_{N-2}) & \cdots & \phi_K(t_{N-2},x_{N-12},\dot{x}_{N-2},v_{N-2},i_{N-2})
\end{bmatrix}, \tag{12}$$

$$\boldsymbol{B}=\begin{bmatrix}
\dot{y}_1 & \dot{v}_1 & \dot{i}_1\\
\vdots & \vdots & \vdots\\
\dot{y}_{N-2} & \dot{v}_{N-2} & \dot{i}_{N-2}
\end{bmatrix},\quad \boldsymbol{u}=\begin{bmatrix}
u_{11} & u_{12} & u_{13}\\
\vdots & \vdots & \vdots\\
u_{K1} & u_{K2} & u_{K3}
\end{bmatrix}.$$

However, Eq. (11) may have no solutions due to the equations more than variables. Based on this, we use the least square sense, i.e.,

$$\arg\min_{\boldsymbol{u}} \|\boldsymbol{A}\boldsymbol{u}-\boldsymbol{B}\|_2^2. \tag{13}$$

Thus, the solution $\boldsymbol{u}$ can be obtained by

$$\boldsymbol{u}=\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}\right). \tag{14}$$

As we know, here $\boldsymbol{u}$ are not yet the sparse solutions. For enhancing the sparsity, we introduce $L_1$ regular, i.e.,

$$\hat{\boldsymbol{u}}=\arg\min_{\boldsymbol{u}}\left(\|\boldsymbol{A}\boldsymbol{u}-\boldsymbol{B}\|_2^2+\rho\|\boldsymbol{u}\|_1\right), \tag{15}$$

where $\rho>0$ is called as the penalty factor applied in restricting the weight of the sparsity constraint. Due to the meaning of the $L_1$ regular, some terms in the solutions will be equal to zero.

The approach we used to solve Eq. (15) is iterative algorithm SSR together with CV. Compared with the iterative threshold algorithm,[21,27] Lasso,[43] elastic net[44] and matching pursuit,[45] SSR not only can adaptively select the number of iterations and the sparsity level, but also does not needlessly adjust the external parameters like threshold parameter $\lambda$. The purposes of CV[46] are the division of data and the selection of optimal parameter. The pipeline works are summarized in Table 1.

**Table 1.** The algorithm for sparse identification.

| **Algorithm** | Python pseudo code of sparse identification for the hybrid VEH |
|---|---|
| **Input** | Datasets of non-dimensional system state time series $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{I}$ |
| | Basis function $\boldsymbol{\Phi}(\boldsymbol{t},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{V},\boldsymbol{I})$. |
| **Calculate** | Derivative of the system state. |
| | Matrix $\boldsymbol{A}$ and $\boldsymbol{B}$. |
| | **for** $\boldsymbol{i}$ in range(the number of k-fold): |
| | Divide the data into training data and test data. |
| | **for** $\boldsymbol{j}$ in range($\boldsymbol{K}$): #$\boldsymbol{K}$ is the number of basis functions. |
| | Solve $\boldsymbol{A}\boldsymbol{u}$=$\boldsymbol{B}$ by using $\boldsymbol{u}$=$\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}\right)$ and training data. |
| | $m=\min_{\mathrm{j}}\lvert u_{\mathrm{j}}\rvert$, $\boldsymbol{u}_m$=0, $\boldsymbol{A}[:,m]=[\,]$ and $\boldsymbol{u}[m]=[\,]$. |
| | Calculate the MSE for the $\boldsymbol{u}$ by using test data. Meanwhile, we save all MSE and $\boldsymbol{u}$. |
| | Select the $\boldsymbol{u}$ corresponding to the minimum MSE as $\hat{\boldsymbol{u}}$. |
| **Output** | The optimal solutions $\hat{\boldsymbol{u}}$. |

### 3.2. Learned process of diffusion term

For GWN, the diffusion term $a_{22}$ in Eq. (4) is calculated as $2D$. Analogously, we construct the basis function $\Psi(\boldsymbol{X})$. Then, the diffusion term is approximated as

$$a_{22}=\sum_{j=1}^{L}q_j\psi_j. \tag{16}$$

The $a_{22}$ can be estimated by the Kramers–Moyal formula

$$a_{22}=\lim_{\Delta t\to 0}\frac{(y_{k+1}-y_k)^2}{\Delta t}. \tag{17}$$

Consequently, the diffusion coefficients $\boldsymbol{q}$ are acquired by

$$\hat{\boldsymbol{A}}\boldsymbol{q}=\hat{\boldsymbol{B}}, \tag{18}$$

where

$$\hat{\boldsymbol{A}}=\begin{bmatrix}\psi_1(x_1) & \cdots & \psi_L(x_1)\\ \vdots & \ddots & \vdots\\ \psi_1(x_{N-1}) & \cdots & \psi_L(x_{N-1})\end{bmatrix},$$
$$\hat{\boldsymbol{B}}=\left[(y_2-y_1)^2,\ldots,(y_{N-1}-y_{N-2})^2\right]/\Delta t. \tag{19}$$

Finally, by performing Eqs. (13)–(15) and the algorithm in Table 1, we obtain

$$\boldsymbol{q}=\left(\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{A}}\right)^{-1}\left(\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{B}}\right). \tag{20}$$

According to the steps in Table 1, we strengthen the sparse level of the solution $\boldsymbol{q}$ and the sparse solution $\hat{\boldsymbol{q}}$ is obtained.

## 4. Discover equations for a hybrid VEH under harmonic excitation

### 4.1. A hybrid VEH under harmonic excitation

Let $f(x)=c_4x^4+c_2x^2+c_0$, $g(x)=\delta_1x+\delta_3x^3+\delta_5x^5$[47,48] and $D=0$. The equations of the system can be expressed as

$$\begin{aligned}&\ddot{x}+(c_4x^4+c_2x^2+c_0)\dot{x}+\delta_1x+\delta_3x^3+\delta_5x^5\\&+\zeta_1V+\zeta_2I=F\cos(t),\\&\dot{V}+\mu_1V=\lambda_1\dot{x},\\&\dot{I}+\mu_2I=\lambda_2\dot{x},\end{aligned} \tag{21}$$

where $c_0$, $c_2$ and $c_4$ are the damping coefficients, and $\delta_1$, $\delta_3$ and $\delta_5$ are the stiff coefficients. Then, we obtain the following differential equations:

$$\mathrm{d}x = y\,\mathrm{d}t, \tag{22a}$$
$$\mathrm{d}y = -[(c_4x^4+c_2x^2+c_0)y+\delta_1x+\delta_3x^3+\delta_5x^5 +\zeta_1V+\zeta_2I]\,\mathrm{d}t+F\cos(t)\,\mathrm{d}t, \tag{22b}$$
$$\mathrm{d}V = (-\mu_1V+\lambda_1y)\,\mathrm{d}t, \tag{22c}$$

$$\mathrm{d}I = (-\mu_2 I + \lambda_2 y)\,\mathrm{d}t. \tag{22d}$$

Meanwhile, we make

$$b_{2\mathrm{h}} = -[(c_4x^4 + c_2x^2 + c_0)y + \delta_1 x + \delta_3 x^3 + \delta_5 x^5 \tag{23a}$$
$$+ \zeta_1 V + \zeta_2 I], \tag{23b}$$
$$b_{3\mathrm{h}} = -\mu_1 V + \lambda_1 y,$$
$$b_{4\mathrm{h}} = -\mu_2 I + \lambda_2 y. \tag{23c}$$

### 4.2. Sparse learning of the system under harmonic excitation

We set parameters as $c_0 = -0.5$, $c_2 = 0.5$, $c_4 = -0.1$, $\delta_1 = 1$, $\delta_3 = -3$, $\delta_5 = 1$, $\zeta_1 = 0.5$, $\zeta_2 = 0.5$, $F = 1$, $\mu_1 = 1$, $\lambda_1 = 1$, $\mu_2 = 0.5$, $\lambda_2 = 1$.

According to Eq. (22), five independent system state trajectories of $N_\mathrm{h} = 10^5$ steps each are generated as the observed data by using the method of the fourth order Runge–Kutta. We compute the derivatives $\dot{\boldsymbol{X}}$, $\dot{\boldsymbol{Y}}$, $\dot{\boldsymbol{V}}$ and $\dot{\boldsymbol{I}}$ by applying Eq. (6) with the time step $\Delta t = 0.001$, where the smaller $\Delta t$, the higher the accuracy. Then, $K_\mathrm{h} = 27$ basis function dictionary $\boldsymbol{\Phi}_\mathrm{h}(t, \boldsymbol{X}, \dot{\boldsymbol{X}}, \boldsymbol{V}, \boldsymbol{I})$ is considered. The specific composition of the dictionary reads as

$$\begin{aligned}&\boldsymbol{\Phi}_\mathrm{h}(\boldsymbol{t}, \boldsymbol{X}, \dot{\boldsymbol{X}}, \boldsymbol{V}, \boldsymbol{I})\\ &= \big[\mathbf{1}, \boldsymbol{X}, \dot{\boldsymbol{X}}, \boldsymbol{X^2}, \boldsymbol{X}\dot{\boldsymbol{X}}, \dot{\boldsymbol{X}}^2, \boldsymbol{X^3}, \boldsymbol{X^2}\dot{\boldsymbol{X}}, \boldsymbol{X}\dot{\boldsymbol{X}}^2,\\ &\quad \dot{\boldsymbol{X}}^3, \boldsymbol{X^4}, \boldsymbol{X^3}\dot{\boldsymbol{X}}, \boldsymbol{X^2}\dot{\boldsymbol{X}}^2, \boldsymbol{X}\dot{\boldsymbol{X}}^3, \dot{\boldsymbol{X}}^4\\ &\quad \boldsymbol{X^5}, \boldsymbol{X^4}\dot{\boldsymbol{X}}, \boldsymbol{X^3}\dot{\boldsymbol{X}}^2, \boldsymbol{X^2}\dot{\boldsymbol{X}}^3, \boldsymbol{X}\dot{\boldsymbol{X}}^4, \dot{\boldsymbol{X}}^5,\\ &\quad \boldsymbol{V}, \boldsymbol{I}, \sin(\boldsymbol{t}), \cos(\boldsymbol{t}), \sin(2\boldsymbol{t}), \cos(2\boldsymbol{t})\big]. \end{aligned} \tag{24}$$

The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be obtained by Eqs. (9), (10) and (12). To avoid under-fitting and over-fitting, 7-fold CV is utilized to select the specific number of iterations on which SSR needs to be run. We have a family of models with the number of iterations as parameter (SSR(0), ..., SSR(27)). By performing the algorithm in Table 1, we select the one model that is best to fit the observed datasets. Figures 2(a), 2(b) and 2(c) respectively demonstrate MSE of coefficients learning from Eqs. (22b), (22c) and (22d), where $n$ denotes the number of non-zero coefficients.

From Fig. 2(a), we can see that when $n < 8$, the MSE vibrates firstly and then decreases slowly. Subsequently, the part of $n \geq 8$ is magnified in the upper right of Fig. 2(a). The results show that when $n$ changes from 8 to 9, the mean square error plummets, and then slowly decreases until $n = 11$. We hold the opinion that the model is under-fitting in $n < 8$ and over-fitting in $n > 11$.[28] Analogously, Figs. 2(b) and 2(c) show that the number of non-zero coefficients in Eqs. (22c) and (22d) is 2. Thus, for Eq. (22b), we focus on $n = 9$ and $n = 10$. For Eqs. (22c) and (22d), $n = 2$ is the best parameter value.

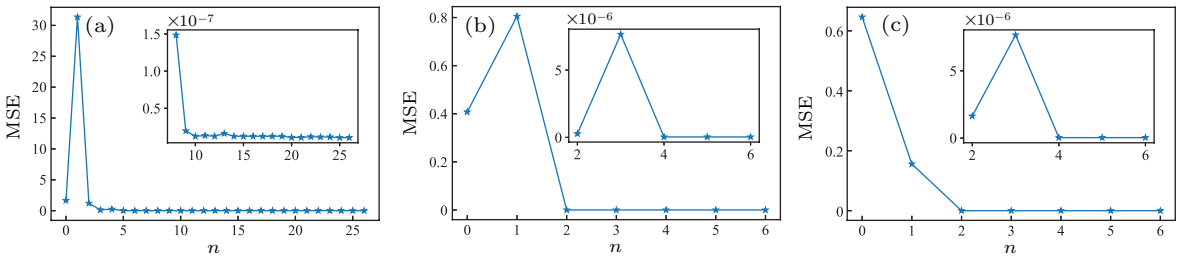



**Fig. 2.** MSE from applying SSR algorithm is plotted as a function of solution size $n$. (a) Eq. (22b); (b) Eq. (22c); (c) Eq. (22d).

**Table 2.** Identified coefficients for Eq. (22b).

| Function (Func.) | True | Learn ($n = 9$) | Learn ($n = 10$) | Func. | True | Learn ($n = 9$) | Learn ($n = 10$) |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | $\dot{X}^4$ | 0 | 0 | 0 |
| $X$ | −1 | −1.00082 | −1.00089 | $X^5$ | −1 | −1.00003 | −1.0002 |
| $\dot{X}$ | 0.5 | 0.49936 | 0.49953 | $X^4\dot{X}$ | 0.1 | 0.09753 | 0.09753 |
| $X^2$ | 0 | 0 | 0 | $X^3\dot{X}^2$ | 0 | 0 | 0 |
| $X\dot{X}$ | 0 | 0 | 0 | $X^2\dot{X}^3$ | 0 | 0 | 0 |
| $\dot{X}^2$ | 0 | 0 | 0 | $X\dot{X}^4$ | 0 | 0 | 0 |
| $X^3$ | 3 | 3.00041 | 3.00034 | $\dot{X}^5$ | 0 | 0 | 0 |
| $X^2\dot{X}$ | −0.5 | −0.49566 | −0.49572 | $V$ | −0.5 | −0.50104 | −0.50152 |
| $X\dot{X}^2$ | 0 | 0 | 0 | $I$ | −0.5 | −0.49906 | −0.49888 |
| $\dot{X}^3$ | 0 | 0 | 0 | $\sin(t)$ | 0 | 0 | −0.00055 |
| $X^4$ | 0 | 0 | 0 | $\cos(t)$ | 1 | 0.99993 | 0.99989 |
| $X^3\dot{X}$ | 0 | 0 | 0 | $\sin(2t)$ | 0 | 0 | 0 |
| $X^2\dot{X}^2$ | 0 | 0 | 0 | $\cos(2t)$ | 0 | 0 | 0 |
| $X\dot{X}^3$ | 0 | 0 | 0 | | | | |

The sparse solutions $\hat{\boldsymbol{u}}_h$ can be learned as listed in Tables 2–4. The results indicate that the coefficients learned from the algorithm are similar to the true values, and $n = 9$ is more suitable for Eq. (22b). Furthermore, Figs. 3(a) and 3(b) demonstrate the comparisons of learned values and true values with regard to damping force $f(x)$ and stiffness term $g(x)$, respectively. Figure 4

shows the comparison of time-varying displacement and time-varying voltage from the original system and the learned system. The results show that the true system and the learning system are in good agreement.

**Table 3.** Identified coefficients for Eq. (22c).

| Func. | True | Learn | Func. | True | Learn | Func. | True | Learn | Func. | True | Learn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 0 | 0 | $\dot{X}$ | 1 | 1.000240 | $X^2$ | 0 | 0 | $X\dot{X}$ | 0 | 0 |
| $\dot{X}^2$ | 0 | 0 | $V$ | −1 | −1.001459 | $I$ | 0 | 0 | | | |

**Table 4.** Identified coefficients for Eq. (22d).

| Func. | True | Learn | Func. | True | Learn | Func. | True | Learn | Func. | True | Learn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 0 | 0 | $\dot{X}$ | 1 | 1.000506 | $X^2$ | 0 | 0 | $X\dot{X}$ | 0 | 0 |
| $\dot{X}^2$ | 0 | 0 | $V$ | −0.5 | −0.499224 | $I$ | 0 | 0 | | | |

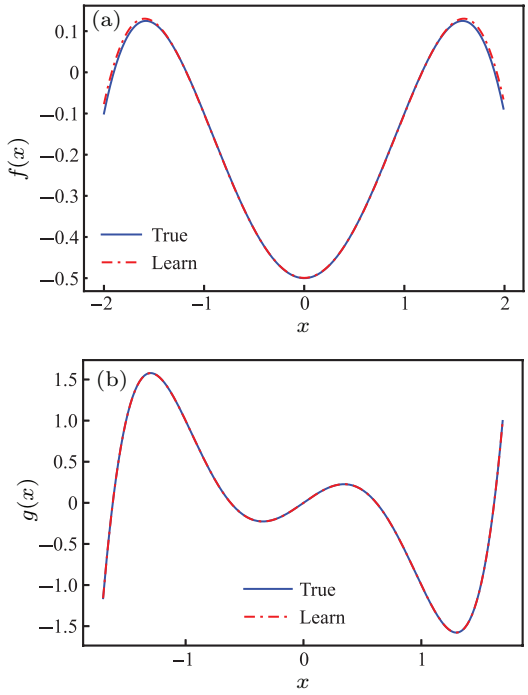



**Fig. 3.** Comparisons between the learned values and true values of (a) $f(x)$ and (b) $g(x)$.

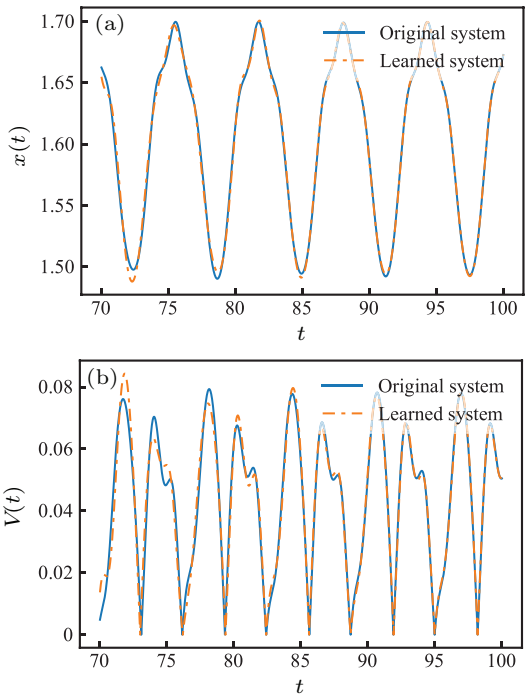



**Fig. 4.** Comparisons of (a) time-varying displacement and (b) time-varying voltage.

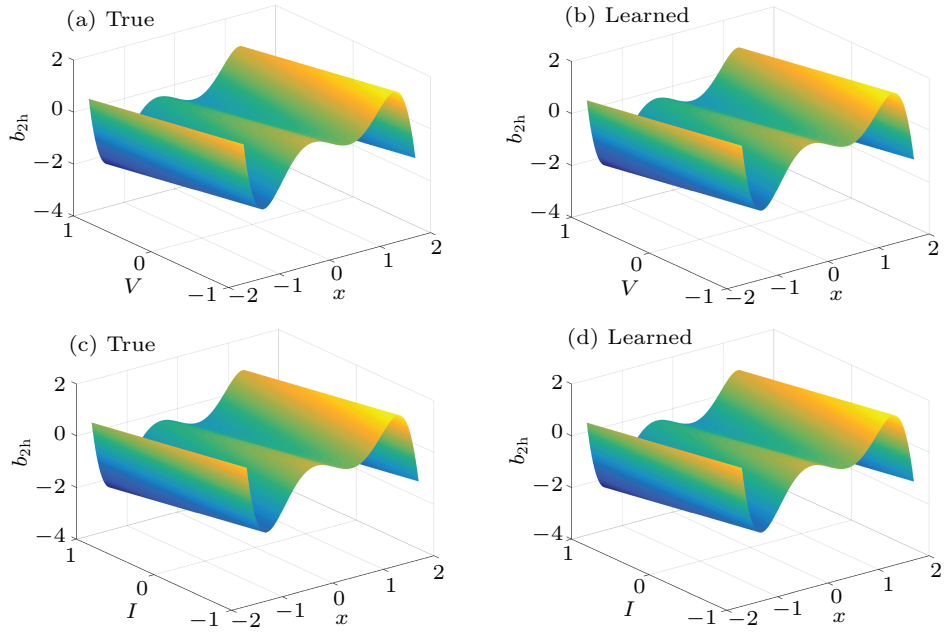



**Fig. 5.** Comparisons between true values and learned values of the $b_{2h}$. (a) and (b) $y = 0.5$, $I = 0.5$; (c) and (d) $y = 0.5$, $V = 0.5$.

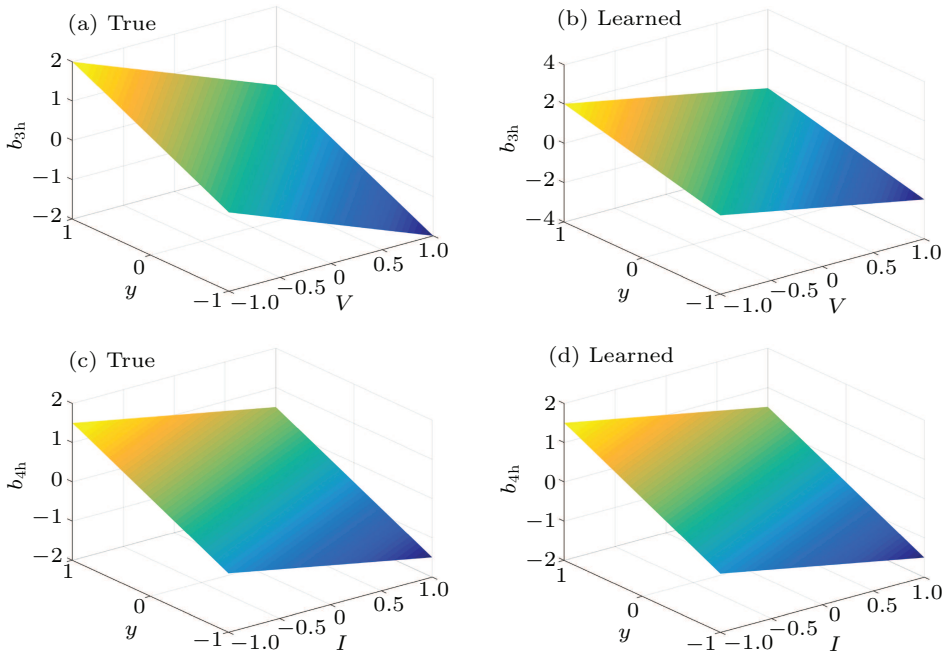



**Fig. 6.** Comparisons between true and learned functions. (a) and (b) $b_{3h}$; (c) and (d) $b_{4h}$.

**Table 5.** The learned results of 10 sets data with different $\Delta t$ and different $N_h$.

| Parameters | True | $\Delta t = 0.001$, $N_h = 2\times10^5$ | $\Delta t = 0.001$, $N_h = 10^5$ | $\Delta t = 0.01$, $N_h = 10^5$ |
|---|---|---|---|---|
| | | Mean (Std.) | Mean (Std.) | Mean (Std.) |
| $\delta_1$ | 1 | 1.000802(0.00019522) | 1.00082(0.00020036) | 1.00611(0.0014855) |
| $c_0$ | −0.5 | −0.499361(0.00015668) | −0.499356(0.00015668) | −0.49452(0.0013342) |
| $\delta_3$ | −3 | −3.00039(0.00009571) | −3.00041(0.000099989) | −3.00117(0.0002840) |
| $c_2$ | 0.5 | 0.495663(0.0010569) | 0.495655(0.0010569) | 0.45675(0.010522) |
| $\delta_5$ | 1 | 1.000031(0.0000077152) | 1.0000346(0.0000084255) | 0.99953(0.000112) |
| $c_4$ | −0.1 | −0.097525(0.0006024) | −0.097524(0.0006024) | −0.075233(0.006024) |
| $\zeta_1$ | 0.5 | 0.50101(0.00024634) | 0.50104(0.00025415) | 0.51296(0.003154) |
| $\zeta_2$ | 0.5 | 0.49908(0.00022365) | 0.49906(0.00022843) | 0.48973(0.002498) |
| $F$ | 1 | 0.999974(0.48542) | 0.99992(0.485485) | 0.99872(0.486192) |
| $\lambda_1$ | 1 | 1.00024(0.48655) | 1.00025(0.486560) | 1.002298(0.487063) |
| $\mu_1$ | 1 | 1.00295(0.00071926) | 1.00298(0.00072429) | 1.02915(0.007092) |
| $\lambda_2$ | 1 | 1.000497(0.486625) | 1.000498(0.486625) | 1.00480(0.487672) |
| $\mu_2$ | 0.5 | 0.49852(0.00035954) | 0.49851(0.00036207) | 0.48546(0.0035364) |

Since Eq. (23b) is four-dimensional, the figure of equation is impossible to plot intuitively. Thus, we plot the equation as two-dimensional by fitting two state variables. In Figs. 5 and 6, comparisons of the true and learned $b_{2h}$, $b_{3h}$ and $b_{4h}$ are shown to indicate the learning results. Here, Figs. 5(a) and 5(c) represent the true function $b_{2h}$; Figs. 5(b) and 5(d) represent the learning function $b_{2h}$. The two rows of the figures denote the case with (i) $y = 0.5$ and $I = 0.5$; (ii) $y = 0.5$ and $V = 0.5$, respectively. Figures 6(a) and 6(b) are the true functions of $b_{3h}$ and $b_{4h}$, respectively. Figures 6(c) and 6(d) are the learning functions of $b_{3h}$ and $b_{4h}$, respectively. The results show that the learning functions agree well with the true functions.

By referring to Refs. [49–51], we know that as the fluctuations of the estimated parameters are small with the increase of data length, the estimated parameters converge approximately to the true value of the parameters. Then, 10 sets data with different $\Delta t$ and $N_h$ are learned by the sparse identification algorithm. The results from Table 5 demonstrate that for the same time step $\Delta t$, the longer the data length, the smaller the standard (Std.) deviation. For the same data length $N_h$, the standard deviation decreases as the $\Delta t$ decreases. Thus, with an increase of data length and a small time step, the identified coefficients can be close to the true values of the coefficients.

## 5. Discover equations for a hybrid VEH under GWN and harmonic excitation

### 5.1. A hybrid VEH under GWN and harmonic excitation

In this section, the parameter values are the same as those in Subsection 4.2, except for the parameter $2D = 0.01$. We consider a hybrid energy harvester under both GWN and harmonic excitation, i.e.,

$$\begin{aligned}&\ddot{x}+(c_4x^4+c_2x^2+c_0)\dot{x}+\delta_1x+\delta_3x^3+\delta_5x^5\\&+\zeta_1V+\zeta_2I=F\cos(t)+W(t),\\&\dot{V}+\mu_1V=\lambda_1\dot{x},\end{aligned}$$

$$\dot{I}+\mu_2 I=\lambda_2\dot{x}. \tag{25}$$

Then, Eq. (25) can be written as the Itô stochastic differential equations

$$\begin{aligned} \mathrm{d}x &= y\mathrm{d}t, && (26a)\\ \mathrm{d}y &= -\left[(c_4x^4+c_2x^2+c_0)y+\delta_1 x+\delta_3x^3+\delta_5x^5\right. \\ &\quad \left.+\zeta_1V+\zeta_2I\right]\mathrm{d}t+F\cos(t)\mathrm{d}t+\sqrt{2D}\mathrm{d}W(t), && (26b)\\ \mathrm{d}V &= (-\mu_1V+\lambda_1 y)\,\mathrm{d}t, && (26c)\\ \mathrm{d}I &= (-\mu_2 I+\lambda_2 y)\,\mathrm{d}t. && (26d) \end{aligned}$$

Accordingly, the diffusion term $a_{22}=2D$, and the drift term $\boldsymbol{b}$ of Eq. (26) can be obtained as

$$\boldsymbol{b}=\begin{bmatrix} b_{1\mathrm{G}}\\ b_{2\mathrm{G}}\\ b_{3\mathrm{G}}\\ b_{4\mathrm{G}}\end{bmatrix}=\begin{bmatrix} y\\ -\left[(c_4x^4+c_2x^2+c_0)y+\delta_1x+\delta_3x^3+\delta_5x^5+\zeta_1V+\zeta_2I\right]+F\cos(t)\\ -\mu_1V+\lambda_1y\\ -\mu_2I+\lambda_2y\end{bmatrix}. \tag{27}$$

### 5.2. Pre-processing of drift term and diffusion term

For the drift term $\boldsymbol{b}$, $K_{\mathrm{G}}=18$ basis function dictionary $\boldsymbol{\Phi}_{\mathrm{G}}(\boldsymbol{t},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{V},\boldsymbol{I})$ is given by

$$\boldsymbol{\Phi}_{\mathrm{G}}(\boldsymbol{t},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{V},\boldsymbol{I})=[\boldsymbol{1},\boldsymbol{X},\dot{\boldsymbol{X}},\boldsymbol{X^2},\boldsymbol{X}\dot{\boldsymbol{X}},\boldsymbol{X^3},\boldsymbol{X^2}\dot{\boldsymbol{X}},\boldsymbol{X^4},\boldsymbol{X^3}\dot{\boldsymbol{X}},\boldsymbol{X^5},\boldsymbol{X^4}\dot{\boldsymbol{X}},\boldsymbol{X^6},\boldsymbol{V},\boldsymbol{I},\sin(\boldsymbol{t}),\cos(\boldsymbol{t}),\sin(2\boldsymbol{t}),\cos(2\boldsymbol{t})]. \tag{28}$$

By using the sparse identification algorithm in Table 1, the drift coefficients $\hat{\boldsymbol{u}}_{\mathrm{G}}$ can be solved by

$$\boldsymbol{A}\boldsymbol{u}=\boldsymbol{B}, \tag{29}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are obtained by Eqs. (9), (10) and (12).

Analogously, for the diffusion term $a_{22}=2D$, $L=4$ basis function dictionary $\boldsymbol{\Psi}(\boldsymbol{X})$ is constructed as

$$\boldsymbol{\Psi}(\boldsymbol{X})=[\boldsymbol{1},\boldsymbol{X},\boldsymbol{X}^2,\boldsymbol{X}^3]. \tag{30}$$

According to Subsection 3.2, the sparse solution $\hat{\boldsymbol{q}}$ can be calculated.

### 5.3. Sparse learning of the system under GWN and harmonic excitation

On the basis of Eq. (26), the fourth-order Runge–Kutta[52] is utilized to generate one hundred independent system state trajectories of $N_{\mathrm{G}}=10^5$ steps. Then, $\dot{\boldsymbol{X}}$, $\dot{\boldsymbol{Y}}$, $\dot{\boldsymbol{V}}$ and $\dot{\boldsymbol{I}}$ are approximated numerically by Eq. (6). Through Eqs. (28) and (29), $\boldsymbol{A}$ and $\boldsymbol{B}$ in Eq. (29) are established. Similarly, $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ are obtained via Eqs. (30) and (19).

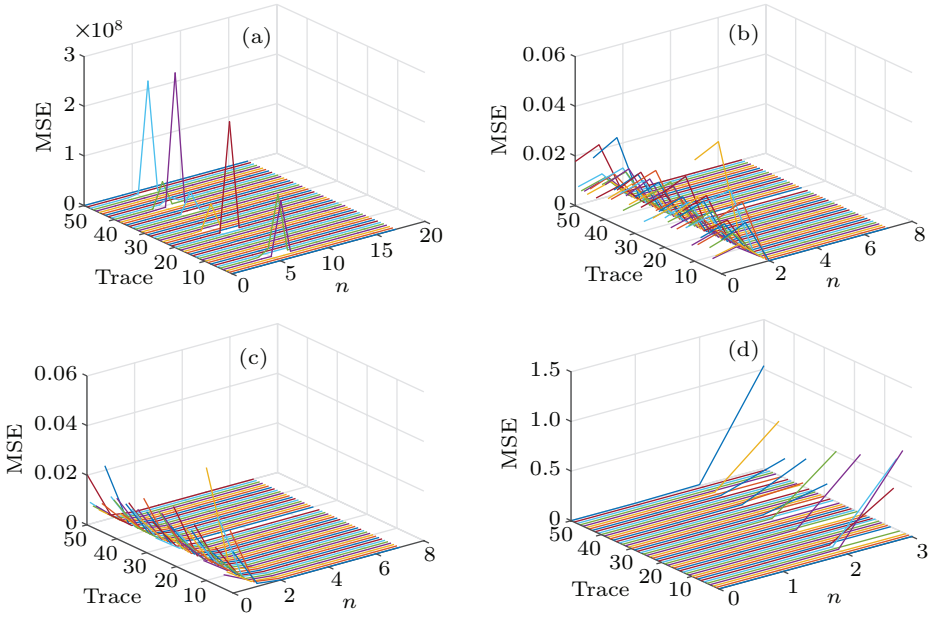



**Fig. 7.** The mean MSE of CV for fifty trajectories and each non-zero coefficients $n$. (a) Function $b_{2\mathrm{G}}$; (b) function $b_{3\mathrm{G}}$; (c) function $b_{4\mathrm{G}}$; (d) function $a_{22}$.

Likewise, we use 10-fold CV to find the optimal number of iterations of SSR. For the one hundred trajectories and each the number of non-zero coefficients $n$, the mean MSE of 10-fold CV is calculated by performing the algorithm in Table 1. Here, Figs. 7(a), 7(b), 7(c) and 7(d) show the mean MSE of the function $b_{2\mathrm{G}}$, $b_{3\mathrm{G}}$, $b_{4\mathrm{G}}$ and $a_{22}$ with every trajectories and the

number of iterations. Then, to judge the number of iterations, we choose the optimal trajectory which has a minimum MSE in all average MSE. Figures 8(a), 8(b), 8(c) and 8(d) are the MSE of the functions $b_{2G}$, $b_{3G}$, $b_{4G}$ and $a_{22}$. From Fig. 8(a), the MSE has large fluctuations in $n < 4$ and small fluctuations in $n \geq 4$. The part of $n \geq 7$ is amplified in the upper right of Fig. 8(a). It can be seen that the MSE decreases sharply from $n = 8$ to $n = 9$ and fluctuates after $n > 10$. Meanwhile, the minimum MSE of the trajectory is acquired in $n = 10$. Thus, we conclude that the best number of iterations is $K_G - 9$ or $K_G - 10$ in drift function $b_{2G}$, i.e., $n = 9$ or $n = 10$. Due to the feature of $b_{2G}$, we only consider $n = 9$. Similarly according to the above analysis, because of the minimum MSE corresponding to the $n$, the number of optimal non-zeros coefficients of the functions $b_{3G}$, $b_{4G}$ and $a_{22}$ are $n = 2$, $n = 2$ and $n = 1$, respectively. Differently, in Fig. 8(d), we choose $n = 1$ because $n = 0$ does not satisfy the existence of white noise in Eq. (25).

In the following, from the sparse identification algorithm in Table 1, we learn the coefficients of $b_{2G}$, $b_{3G}$, $b_{4G}$ and $a_{22}$ using $K_G - 9$, $K_G - 2$, $K_G - 2$ and $L - 1$ iterations, respectively. With the application of the optimal trajectory and 10-fold CV, the learned results are shown in Tables 6, 7, 8 and 9 correspond to $b_{2G}$, $b_{3G}$, $b_{4G}$ and $a_{22}$, respectively.

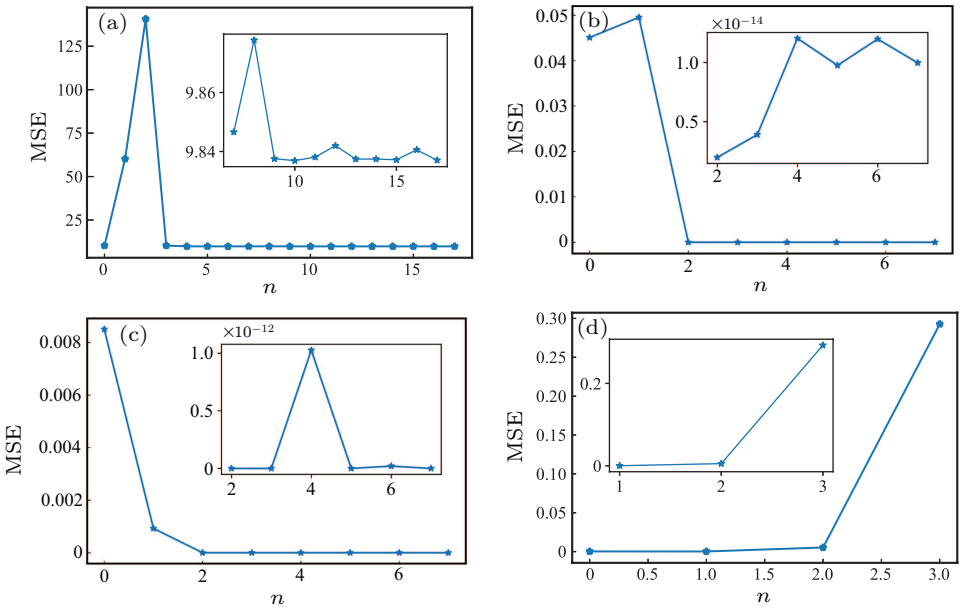



**Fig. 8.** The MSE of optimal trajectory with non-zero coefficients $n$. (a) Function $b_{2G}$; (b) function $b_{3G}$; (c) function $b_{4G}$; (d) function $a_{22}$.

**Table 6.** Identified coefficients for $b_{2G}$.

| Func. | True | Learn ($n = 9$) | Func. | True | Learn ($n = 9$) |
|---|---|---|---|---|---|
| 1 | 0 | 0 | $X^5$ | −1 | −1.002978 |
| $X$ | −1 | −0.994891 | $X^4\dot{X}$ | 0.1 | 0.097383 |
| $\dot{X}$ | 0.5 | 0.494865 | $X^6$ | 0 | 0 |
| $X^2$ | 0 | 0 | $V$ | −0.5 | −0.518883 |
| $X\dot{X}$ | 0 | 0 | $I$ | −0.5 | −0.481516 |
| $X^3$ | 3 | 3.003853 | $\sin(t)$ | 0 | 0 |
| $X^2\dot{X}$ | −0.5 | −0.490058 | $\cos(t)$ | 1 | 1.010846 |
| $X^4$ | 0 | 0 | $\sin(2t)$ | 0 | 0 |
| $X^3\dot{X}$ | 0 | 0 | $\cos(2t)$ | 0 | 0 |

**Table 7.** Identified coefficients for $b_{3G}$.

| Func. | True | Learn | Func. | True | Learn | Func. | True | Learn | Func. | True | Learn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | $\dot{X}$ | 1 | 0.9995 | $X\dot{X}$ | 0 | 0 | $V$ | −1 | −0.999501 |
| $X$ | 0 | 0 | $X^2$ | 0 | 0 | $X^2\dot{X}$ | 0 | 0 | $I$ | 0 | 0 |

**Table 8.** Identified coefficients for $b_{4G}$.

| Func. | True | Learn | Func. | True | Learn | Func. | True | Learn | Func. | True | Learn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | $\dot{X}$ | 1 | 0.999750 | $X\dot{X}$ | 0 | 0 | $V$ | 0 | 0 |
| $X$ | 0 | 0 | $X^2$ | 0 | 0 | $X^2\dot{X}$ | 0 | 0 | $I$ | −0.5 | −0.499875 |

**Table 9.** Identified coefficients for $a_{22}$.

| Func. | True | Learn | Func. | True | Learn | Func. | True | Learn | Func. | True | Learn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $2D$ | 0.01 | 0.010011 | $X$ | 0 | 0 | $X^2$ | 0 | 0 | $X^3$ | 0 | 0 |

We show intuitively the learned results in Figs. 9–11. Figure 9 demonstrates the fitting situation of the functions $f(x)$ and $g(x)$. Figure 10 represents the comparison of the original system and the learned system in time-varying displacement $x(t)$ and voltage $V(t)$. It can be seen that compared with the true coefficients, the learned coefficients have slight deviation within an acceptable range.

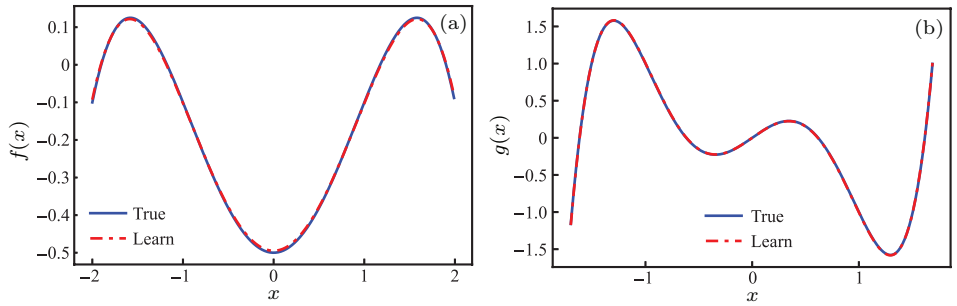



**Fig. 9.** Comparisons between the learned values and true values of (a) $f(x)$ and (b) $g(x)$.

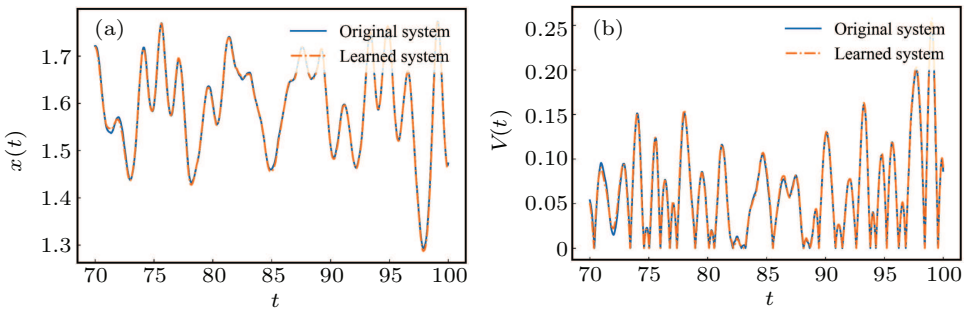



**Fig. 10.** Comparisons of (a) time-varying displacement and (b) time-varying voltage.

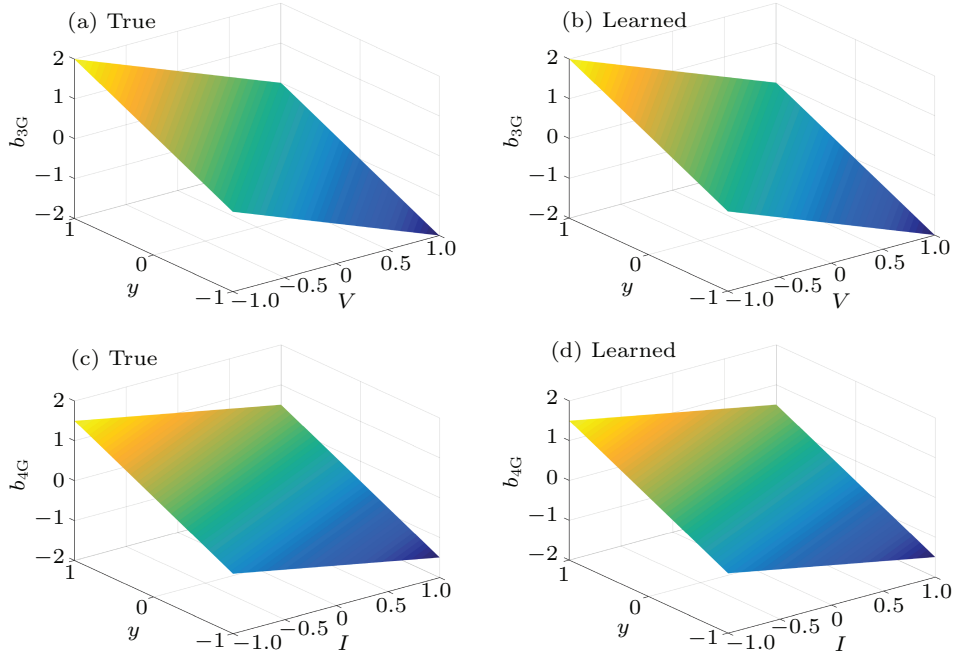



**Fig. 11.** Comparisons between true and learned functions. (a) and (b) $b_{3G}$; (c) and (d) $b_{4G}$.

Then, the results of comparison between true values and learned values with respect to $b_{3G}$ and $b_{4G}$ are shown in Fig. 11. Figures 11(a) and 11(b) are the true functions of $b_{3G}$ and $b_{4G}$, respectively. Figures 11(c) and 11(d) are the learning functions of $b_{3G}$ and $b_{4G}$, respectively. The true and learned results of function $b_{2G}$ are displayed in three parts $x$–$V$, $x$–$I$ and $x$–$t$. We

demonstrate the equation in two-dimensional due to the dimension of $b_{2G}$ more than two. Figures 12(a)–12(e) represent the true results. Figures 12(b)–12(f) represent the learned results. The three rows of the figures describe the case with (i) $y = 0.5$, $I = 0.5$ and $t = 10$; (ii) $y = 0.5$, $V = 0.5$ and $t = 10$; (iii) $y = 0.5$, $V = 0.5$ and $I = 0.5$, respectively. It can be seen that the coefficients learned from the sparse identification algorithm have good enough accuracy.

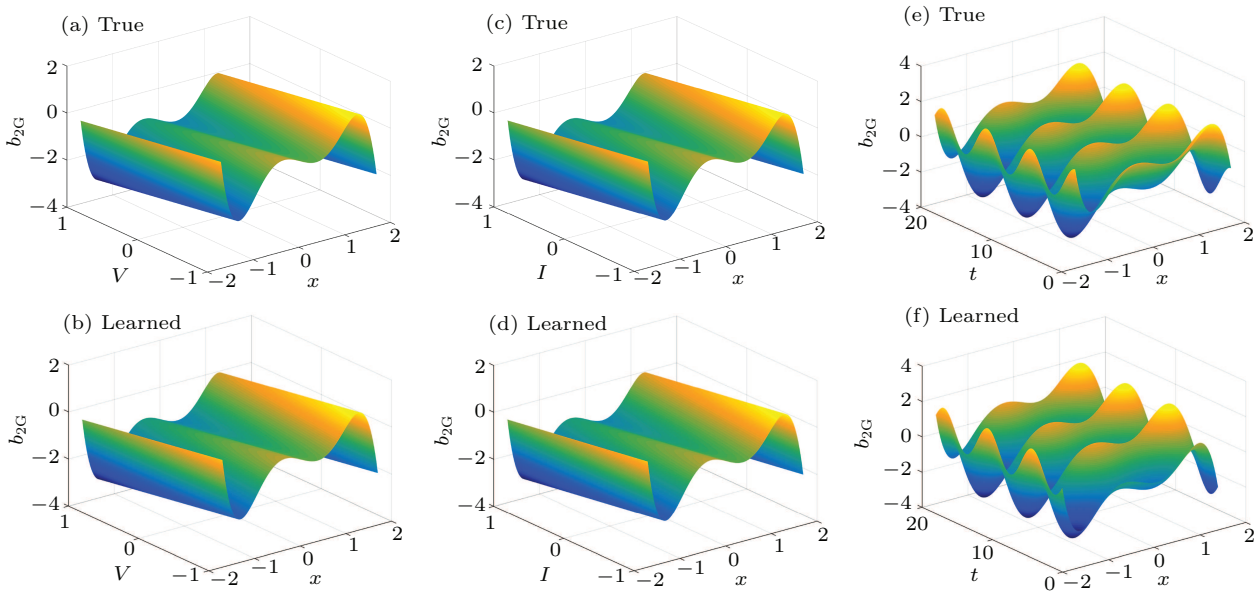



**Fig. 12.** Comparisons between true values and learned values of the $b_{2G}$. (a) and (b) $y = 0.5$, $I = 0.5$, $t = 10$; (c) and (d) $y = 0.5$, $V = 0.5$, $t = 10$; (e) and (f) $y = 0.5$, $V = 0.5$, $I = 0.5$.

Above all, the hybrid energy harvesting system identified by sparse identification is consistent enough with the real system.

## 6. Conclusion

In this paper, a sparse identification approach was developed to acquire the governing equations of the hybrid energy harvesting system from the simulated sample state data. Two examples were taken to verify the feasibility and effectiveness of the method.

To begin with, a hybrid energy harvester under harmonic excitation was the first example. Through approximating derivatives by the first-order difference and constructing the basis functions dictionary, we obtained the expressions of differential equations of the system, which are equal to the linear combination of basis functions. Then, for the number of non-zero coefficients $n$ and each sample trajectory, 7-fold cross-validation (CV) was applied to prevent under-fitting and over-fitting by observing the variations of MSE. Removing the situations of under-fitting and over-fitting, we selected $n = 9$, $n = 2$ and $n = 2$ for the differential equations $\dot{y}$, $\dot{V}$ and $\dot{I}$, respectively. After solving $\boldsymbol{Au} = \boldsymbol{B}$ by using the sparse identification algorithm, we learned the unknown coefficients, and discussed the degree of fitting. The results showed that the method is applied to solve the coefficients which are sufficiently accurate to the true functions, and all the learned coefficients are greatly converge to the true coefficients with the increase of data length under a small time step. Thus, this method can be well utilized to the deterministic hybrid energy harvesting system.

A hybrid energy harvester under both harmonic excitation and Gaussian white noise was the second example. Firstly, we received the approximated equations of the drift term and diffusion term based on the Kramers–Moyal formulas. According to the basis function dictionary, we calculated the iterative expressions of the drift term and diffusion term, respectively. Together with the 10-fold CV, the sparse identification algorithm was used to obtain the number of optimal non-zero coefficients $n = 9$, $n = 2$, $n = 2$ and $n = 1$ corresponding to the functions of $b_{2G}$, $b_{3G}$, $b_{4G}$ and $a_{22}$, respectively. According to the comparison of learned functions and true functions, the results demonstrated that the method is well applied to the hybrid energy harvesting system with an acceptable deviation. Meanwhile, compared with the first example, the second example depends on more sufficient data to reduce the effect of the noise.

Thus, measuring the time-series data of the system state, we can build the equations for the hybrid energy harvester. Then, the learned system can be applied to explore the subsequent dynamical behavior with the aim of the improvement of performance of the energy harvester.

## Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant Nos. 12002089 and 11902081) and

Project of Science and Technology of Guangzhou (Grant No. 202201010326).

## References


[1] Liu Q, Xu Y and Kurths J 2020 *Commun. Nonlinear Sci. Numer. Simul.* **84** 105184
[2] Yang Y G, Xu W, Sun Y H and Gu X D 2015 *Chin. Phys. B* **25** 020201
[3] Zhang Y, Jin Y and Li Y 2021 *Phys. D* **422** 132908
[4] Daqaq M F 2012 *Nonlinear Dyn.* **69** 1063
[5] Seuaciuc-Osório T and Daqaq M F 2010 *J. Sound Vib.* **329** 2497
[6] Foong F M, Thein C K, Ooi B L and Yurchenko D 2019 *Mech. Syst. Signal Process.* **116** 129
[7] Dragunov V P, Dorzhiev V Y, Ostertak D I and Atuchin V V 2018 *Sensors Actuators, A Phys.* **272** 259
[8] Ibrahim A, Ramini A and Towfighian S 2018 *J. Sound Vib.* **416** 111
[9] Siddique A R M, Mahmud S and Heyst B V 2015 *Energy Convers. Manag.* **106** 728
[10] Panyam M and Daqaq M F 2017 *J. Sound Vib.* **386** 336
[11] Karami M A and Inman D J 2011 *J. Sound Vib.* **330** 5583
[12] Xia H, Chen R and Ren L 2015 *Sensors Actuators, A Phys.* **234** 87
[13] Jiang J, Xu W, Han P and Niu L 2022 *Chin. Phys. B* **31** 060203
[14] Wang X, Duan J, Li X and Song R 2018 *Appl. Math. Comput.* **337** 618
[15] Zhou X Y, Gao S Q, Liu H P and Guan Y W 2017 *Smart Mater. Struct.* **26** 015008
[16] Sengha G G, Fokou Kenfack W, Siewe Siewe M, Tabi C B and Kofane T C 2020 *Commun. Nonlinear Sci. Numer. Simul.* **90** 105364
[17] Foupouapouognigni O, Nono Dueyou Buckjohn C, Siewe Siewe M and Tchawoua C 2018 *Phys. A* **509** 346
[18] Sun Y H, Yang Y G, Zhang Y and Xu W 2021 *Chaos* **31** 013111
[19] Mokem Fokou I S, Nono Dueyou Buckjohn C, Siewe Siewe M and Tchawoua C 2018 *Commun. Nonlinear Sci. Numer. Simul.* **56** 177
[20] Yang T and Cao Q 2019 *Mech. Syst. Signal Process.* **121** 745
[21] Brunton S L and Kutz J N 2019 *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* (Cambridge: University Press) pp. 229-275
[22] Chang L J, Mo Y F, Ling L M and Zeng D L 2022 *Chin. Phys. B* **31** 060201
[23] Schwantes C R and Pande V S 2015 *J. Chem. Theory Comput.* **11** 600
[24] Gagne D J, Christensen H M, Subramanian A C and Monahan A H 2020 *J. Adv. Model. Earth Syst.* **12** e2019MS001896
[25] Canhoto A I 2021 *J. Bus. Res.* **131** 441
[26] Kutz J N, Brunton S L, Brunton B W and Proctor J L 2016 *Dynamic mode decomposition: data-driven modeling of complex systems* (Society for Industrial and Applied Mathematics) pp. 195–206
[27] Brunton S L, Proctor J L, Kutz J N and Bialek W 2016 *Proc. Natl. Acad. Sci. USA* **113** 3932
[28] Boninsegna L, Nüske F and Clementi C 2018 *J. Chem. Phys.* **148** 241723
[29] Rudy S H, Nathan Kutz J and Brunton S L 2019 *J. Comput. Phys.* **396** 483
[30] Dai M, Duan J, Liao J and Wang X 2021 *Appl. Math. Comput.* **397** 125927
[31] Lu P and Lermusiaux P F 2021 *Phys. D* **427** 133003
[32] Huang Y and Li Y 2022 *Chin. Phys. B* **31** 070501
[33] Wu D, Fu M and Duan J 2019 *Chaos* **29** 093122
[34] Zhang Y, Duan J, Jin Y and Li Y 2020 *Chaos* **30** 1
[35] Lu Y and Duan J 2020 *Chaos* **30** 1
[36] Li Y and Duan J 2022 *J. Stat. Phys.* **186** 1
[37] Li Y, Lu Y, Xu S and Duan J 2022 *J. Stat. Mech. Theory Exp.* 023405
[38] Li Y and Duan J 2021 *Phys. D* **417** 132830
[39] Xu M, Jin X, Wang Y and Huang Z 2014 *Nonlinear Dyn.* **78** 1451
[40] Huang D, Zhou S and Litak G 2019 *Commun. Nonlinear Sci. Numer. Simul.* **69** 270
[41] Champion K, Lusch B, Nathan Kutz J and Brunton S L 2019 *Proc. Natl. Acad. Sci. USA* **116** 22445
[42] Risken H 1996 *Fokker-Planck Equation* (Berlin: Springer-Heidelberg) pp. 63–95
[43] Tibshirani R 2011 *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** 273
[44] Zou H and Hastie T 2005 *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67** 301
[45] Mallat S G and Zhang Z 1993 *IEEE Trans. Signal Process.* **41** 3397
[46] Hastie T, Tibshirani R and Friedman J H 2001 *The elements of statistical learning: data mining, inference, and prediction* (New York: Springer) p. 214
[47] Guo S L, Yang Y G and Sun Y H 2021 *Chaos, Solitons and Fractals* **151** 111231
[48] Zhang Y, Duan J, Jin Y and Li Y 2021 *Nonlinear Dyn.* **106** 2829
[49] Pavliotis G A and Stuart A M 2007 *J. Stat. Phys.* **127** 741
[50] Samson A and Thieullen M 2012 *Stoch. Process. their Appl.* **122** 2521
[51] Lu F, Lin K and Chorin A 2016 *Comm. App. Math Comp. Sci.* **11** 187
[52] Wang Z Q, Xu Y and Yang H 2016 *Sci. China Technol. Sci.* **59** 371