ASME Journal of Applied Mechanics
Online journal at:
https://asmedigitalcollection.asme.org/appliedmechanics

Chunjiang Li
Department of Engineering Mechanics,
Zhejiang University,
Hangzhou 310027, China;
Zhejiang Lab,
Hangzhou 311121, China
e-mail: lcj@zju.edu.cn

Zhanchao Huang
Department of Engineering Mechanics,
Zhejiang University,
Hangzhou 310027, China;
School of Engineering,
Westlake University,
Hangzhou 310030, China
e-mail: huangzhanchao@westlake.edu.cn

Zhilong Huang
Department of Engineering Mechanics,
Zhejiang University,
Hangzhou 310027, China
e-mail: zlhuang@zju.edu.cn

Yong Wang[1]
Department of Engineering Mechanics,
Zhejiang University,
Hangzhou 310027, China
e-mail: ypwang@zju.edu.cn

Hanqing Jiang
School of Engineering,
Westlake University,
Hangzhou 310030, China
e-mail: hanqing.jiang@westlake.edu.cn

Automated Identification of Differential-Variational Equations for Static Systems

Data-driven equation identification for dynamical systems has achieved great progress, which for static systems, however, has not kept pace. Unlike dynamical systems, static systems are time invariant, so we cannot capture discrete data along the time stream, which requires identifying governing equations only from scarce data. This work is devoted to this topic, building a data-driven method for extracting the differential-variational equations that govern static behaviors only from scarce, noisy data of responses, loads, as well as the values of system attributes if available. Compared to the differential framework typically adopted in equation identification, the differential-variational framework, due to its spatial integration and variation arbitrariness, brings some advantages, such as high robustness to data noise and low requirements on data amounts. The application, efficacy, and all the aforementioned advantages of this method are demonstrated by four numerical examples, including three continuous systems and one discrete system.
[DOI: 10.1115/1.4063641]

Keywords: data-driven equation identification, differential-variational framework, static system, sparse optimization, dimensional analysis, elasticity, structures

1 Introduction

Data science in recent years has received remarkable progress along with the great advances in computing hardware and numerical algorithms and is qualified to be the fourth paradigm of scientific discovery [1]. The explosive advancement of data science has innovated the research means of engineering techniques and has even promoted the automation of scientific discovery.

Here, let us focus our attention on a relatively narrow but exciting subfield: data-driven equation identification, which involves automatically or at least semi-automatically extracting the equations that govern the behaviors of interest from the discrete data of outputs and inputs captured from systems. Early works in this area resort to symbolic regression [2–4] and sparse-promoting optimization [5–7]. Symbolic regression-based methods possess a high degree of flexibility but are prone to overfitting [8,9], while sparse regression-based methods possess high efficiency but are sensitive to data noise. Further developments along this direction include, but are not limited to, integral identification [7], variational identification [10], and differential equation weak-form identification [11]. Undoubtedly, all of these *apparently* different descriptions reflect *essentially* the same phenomenon. The ultimate objective of all these works is to improve computational efficiency, reduce the requirement on data amount, and improve robustness to data noise, albeit with different means.

All works mentioned above aim to extract the *explicit* relations between outputs and inputs. Another set of works principally based on neural networks–ordinary differential equations [12,13] has been parallelly established to extract the *implicit* relations. By embedding knowledge of fields, Lagrangian neural networks [14], Hamiltonian neural networks [15,16], and their various variants [17–19] have been further developed to improve the precision of identification. Compared to implicit methods, explicit methods always extract the governing equations with interpretability and,

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received September 20, 2023; final manuscript received September 25, 2023; published online October 17, 2023. Tech. Editor: Pradeep Sharma.

therefore, play a pivotal role in scientific discovery because it is only the explicit relations that are the golden criteria of scientific explanation.

Nevertheless, the preceding works are entirely confined to identifying governing equations for dynamical systems, i.e., identifying the time-evolution equations of states or other variants, while no work has been devoted to identifying governing equations for static systems. Hence, an ultimately important issue arises: *why are statics unconsciously neglected or even deliberately excluded?* As a reasonable speculation, it could be because statics is always regarded as a special case of dynamics. However, this is not true, at least in equation identification! The most notable difference between statics and dynamics lies in the difficulty of capturing discrete data from which governing equations are identified. For dynamical systems, data are endlessly generated with the elapse of time and can be readily captured by sensors [20]; for static systems, however, one load only generates one set of data, which does not change over time. This requires the equation identification of static systems must be implemented from scarce data. The variational identification method [10] immediately becomes one alternative, given its low requirement for data amount and the robustness to data noise [21], although the present version is only suitable for dynamical systems [22].

The variational identification for dynamical systems takes the variational principle, the integral-variational principle specifically, as the fundamental assumption [23]. Under the integral-variational framework, it comes down to determine two functions of integrands. As a matter of fact, there exist two classes of variational principles, that is, integral-variational principles and differential-variational principles; both of them provide the conditions distinguishing the true response from all possible responses. The former compares in a finite time interval, while the latter at some instant. Both the integral- and differential-variational principles can be adopted to describe the behaviors of dynamical systems, and they are mutually equivalent. Static systems do not evolve over time; thus, only the differential-variational principles can be adopted.

This work is devoted to this topic, namely, building a data-driven method to automatically identify the differential-variational equations for static systems only from scarce discrete data. It can be regarded as the companion piece of variational equation identification for dynamical systems [10]. The rest is organized as follows: Sec. 2 gives all details of methodology, from introducing fundamental assumption, collecting discrete data, prespecifying variations, constructing and identifying functions, up to implementing sparsification; in Sec. 3, through three continuous examples, we illustrate the application of this method and discuss its robustness to data noise and the requirement on data amount; Sec. 4 is dedicated to the equation identification of discrete static systems, with an emphasis on the difference between continuous system identification and discrete system identification. Section 5 closes this paper and gives some conclusions, discussion, and prospects.

2 Methodology

This section is devoted to establishing a data-driven method to automatically identify the differential-variational equations for general static systems. The differential-variational framework is first embedded as a fundamental assumption. The method consists of several successive steps, including *collecting* discrete data of actual responses (output) and of loads (input), *prespecifying* the variations of actual responses, *constructing* the form of the to-be-determined functions, and finally *identifying* functions and implementing sparsification.

Herein, let us first confine the discussion to the case of continuous systems; the discrete systems will be treated in Sec. 4. For a continuous static system (i.e., a time-invariant continuous system) encountered in practice, we should first pinpoint the responses concerned $\boldsymbol{q}(\mathbf{x})$, all effective loads $\mathbf{f}(\mathbf{x})$ influencing the responses, and provide all system attributes $\mathbf{s}$ if available, in which the arguments (vector $\mathbf{x}$) represent the essential spatial dimension of the system under consideration. Taking a mechanical system as a representative, the responses will be the displacements of points of material with spatial coordinates $\mathbf{x}$, the effective loads will be the forces and/or couples imparted on some positions and/or some domains, while the system attributes may include mass, length, material properties, and so forth. Our objective is to establish the explicit relations within $\boldsymbol{q}(\mathbf{x})$, $\mathbf{f}(\mathbf{x})$, and $\mathbf{s}$ only from the discrete data captured.

Here, we first introduce the *fundamental assumption*: for a static continuous system, the actual responses satisfy the following differential-variational equation:

$$\delta\left[\int_\Omega U_0(\mathbf{s};\,\mathbf{q},\,\nabla\mathbf{q},\,\nabla^2\mathbf{q},\,\ldots)d\mathbf{x}\right]+\int_\Omega U(\mathbf{f}(\mathbf{x});\,\mathbf{s};\,\mathbf{q},\,\nabla\mathbf{q},\,\nabla^2\mathbf{q},\,\ldots;\,\delta\mathbf{q},\,\delta\nabla\mathbf{q},\,\delta\nabla^2\mathbf{q},\,\ldots)d\mathbf{x}=0 \tag{1}$$

provided that the variations of actual responses satisfy the imposed boundary conditions. Here, Ω is the spatial domain occupied by the continuous system, U_0 is a function of responses, their various-order spatial derivatives, and system attributes, while U is a function of responses, its spatial derivatives and their variations, loads, and system attributes. U_0 and U possess the dimension of energy density. This fundamental assumption is exactly the same as the differential-variational principle of mechanics, which is one of the most ubiquitous principles in natural science. In the viewpoint of mathematics, a differential-variational equation is equivalent to a set of differential equations, with the former involving lower-order spatial derivatives. Our objective is now converted to identify two unknown functions U_0 and U from the discrete data $\boldsymbol{q}(\mathbf{x})$, $\mathbf{f}(\mathbf{x})$, and system attributes $\mathbf{s}$.

Step 1: Collecting discrete data of responses and loads. For a continuous system encountered in practice, the only way to capture the discrete data of responses and loads is through experimental measurements: arranging various-class sensors appropriately (usually equally distributed in space) and accurately recording time-invariant discrete data. Intrinsically, a continuous system occupies three spatial dimensions; in practice, however, the responses of interest may only depend on one or two spatial dimensions, which can dramatically reduce the requirement on the number of sensors or the times of testing. Different discrete data can be captured for different cases by changing the values of loads. It should be emphasized that the discrete data captured from several cases are crucial to extract all imposed boundary conditions and to specify the variations of actual responses. For case i, we denote the discrete data of responses as $\{\boldsymbol{q}_i(\mathbf{x}_j)\}$ and that of loads as $\{\mathbf{f}_i\}$, in which $\mathbf{x}_j$ represents discrete spatial coordinates.

Step 2: Prespecifying variations of actual responses. Here, we start with several sets of discrete data of responses $\{\boldsymbol{q}_i(\mathbf{x}_j)\}$ captured from several cases and numerically evaluate various-order spatial derivatives $\{\nabla\boldsymbol{q}_i(\mathbf{x}_j)\}$, $\{\nabla^2\boldsymbol{q}_i(\mathbf{x}_j)\}\ldots$. at boundaries. All imposed boundary conditions can be readily extracted by carefully inspecting the discrete data and various-order spatial derivatives at the boundaries: if some order derivative at some boundaries remains unchanged for all sets of discrete data, this specific derivative is regarded as an imposed boundary condition. The highest order evaluated at boundaries is first selected as one or two and then gradually increases until the equation identified meets the required precision. This implementation satisfies the rule of parsimony [5]. Once all imposed boundary conditions are extracted, we can easily prespecify a set of variations of responses satisfying all of them.

Step 3: Constructing the form of to-be-determined functions. Here, we illustrate how to construct the form of to-be-determined functions U_0 and U for two different situations, respectively. For the first situation, where the dimensions

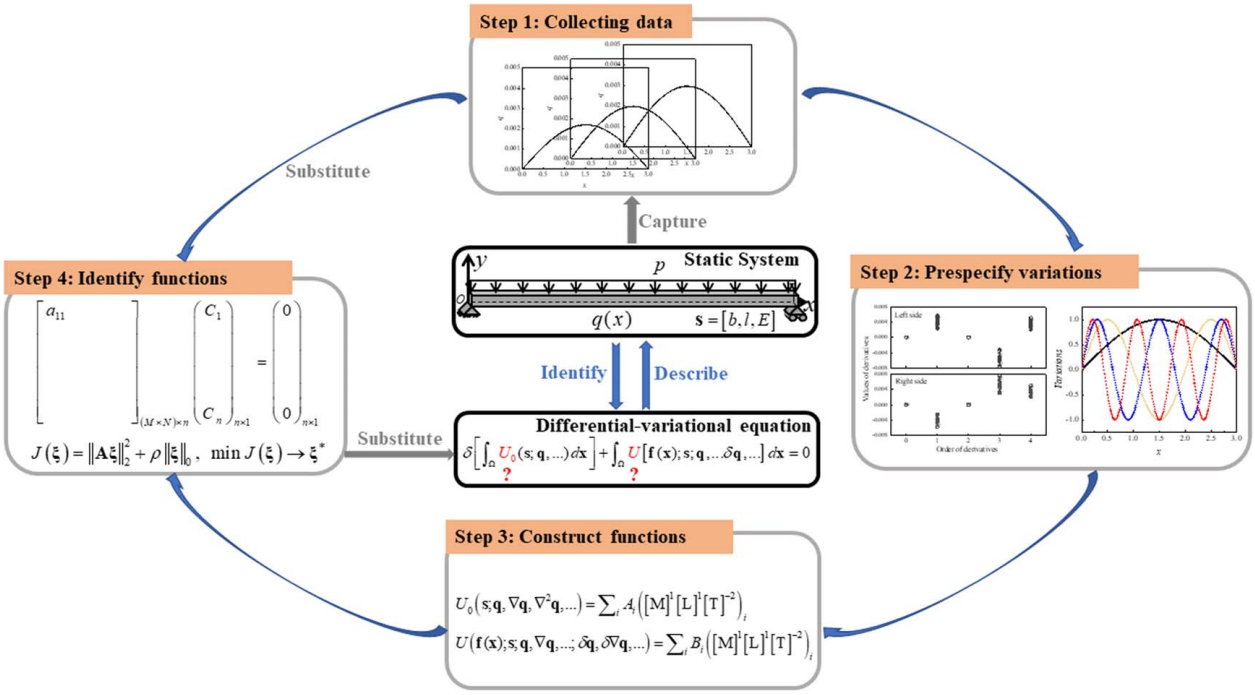

Fig. 1 Flow chart of the data-driven method for identifying differential-variational equations

of all quantities including responses, loads, and system attributes are known to us, we express their dimensions by a set of basic dimensions, mass (M), length (L), and time (T), for instance; we then construct parameter clusters with the dimension of energy density by responses $\boldsymbol{q}(\mathbf{x})$, their various-order spatial derivatives $\nabla\boldsymbol{q}$, $\nabla^2\boldsymbol{q}\ldots$., loads $\mathbf{f}(\mathbf{x})$, and system attributes $\mathbf{s}$, and express the integrand U_0 by the linear combination of these parameter clusters, that is, $U_0 = \sum_i A_i([\mathrm{M}]^1[\mathrm{L}]^1[\mathrm{T}]^{-2})_i$. Similarly, we construct parameter clusters by responses, their various-order spatial derivatives, their variations $\delta\mathbf{q}$, $\delta\nabla\mathbf{q}$, $\delta\nabla^2\mathbf{q}\ldots$., loads, and system attributes and express the integrand U as $U = \sum_i B_i([\mathrm{M}]^1[\mathrm{L}]^1[\mathrm{T}]^{-2})_i$. Here, $(\bullet)_i$ denotes the ith parameter cluster with dimension "•" constructed by associated quantities. Now, let us move to another situation in which the dimensions of all quantities, or some of them, are unknown to us. In this situation, constructing the to-be-determined functions by the rule of dimensional consistency is not feasible. Nevertheless, we can simply expand the integrands as the linear combinations of a set of preselected functions of responses, their various-order spatial derivatives, and their variations, such as power functions, trigonometric functions, exponential functions, and so forth. Undoubtedly, it is impossible to include all terms in the analysis, so the linear combinations must be truncated according to the parsimonious criterion [5], which means that only the lower-order terms are significant in practical physical systems. So far, what we should do is to determine two sets of coefficients A_i, B_i by the discrete data captured.

Step 4: Identifying functions and implementing sparsification. Substituting the captured discrete data (responses $\{\boldsymbol{q}_i(\mathbf{x}_j)\}$, loads $\{\mathbf{f}_i\}$, and system attributes $\mathbf{s}$), the discrete data calculated (various-order spatial derivatives $\{\nabla\boldsymbol{q}_i(\mathbf{x}_j)\}$, $\{\nabla^2\boldsymbol{q}_i(\mathbf{x}_j)\}\ldots$.), and the variations prespecified ($\delta\boldsymbol{q}_i(\mathbf{x}_j)$, $\delta\nabla\boldsymbol{q}_i(\mathbf{x}_j)$, $\delta\nabla^2\boldsymbol{q}_i(\mathbf{x}_j)\ldots$.) into differential-variational Eq. (1), we establish a set of overdetermined linear algebraic equations $\mathbf{A}\boldsymbol{\xi} = \mathbf{0}$ by some simple calculus, in which $\mathbf{A}$ is a definite and known rectangular matrix, $\boldsymbol{\xi}$ is a column vector constituted by the to-be-determined coefficients A_i, B_i. The non-trivial solution of the overdetermined algebraic equations, i.e., the values of coefficients $\boldsymbol{\xi}$, can be calculated by the pseudo-inverse algorithm. Once the values of coefficients $\boldsymbol{\xi}$ are determined, the differential-variational equation satisfied by the static system under consideration is totally identified.

Note that the column vector $\boldsymbol{\xi}$ (i.e., the list of coefficients of parameter clusters) determined by the pseudo-inverse algorithm [24] is usually dense but not sparse. Sparsity, however, is one of the most significant characteristics of nature; thus, it is natural to construct an associated problem of sparse optimization [25–27] to solve the aforementioned overdetermined equations and acquire the sparsest solution and, consequently, the sparsest differential-variational equation. Specifically, we introduce the residual as

$$J(\boldsymbol{\xi}) = \|\mathbf{A}\boldsymbol{\xi}\|_2^2 + \rho\|\boldsymbol{\xi}\|_0 \tag{2}$$

in which $\|\cdot\|_2$ and $\|\cdot\|_0$ denote L^2-norm and L^0-norm of the argument, respectively; ρ is a weighted factor penalizing the number of non-zero components. By solving the optimization problem in Eq. (2), we obtain a sparse solution balancing the precision of identification and the complexity of the model. The optimization problem is implemented by an iterative procedure; that is, in each iterative step, we delete one term with minimal absolute value and proceed to the next iterative step until the residual abruptly increases. It is worth pointing out that the assumption of sparsity is not always valid. For instance, let us consider a relation precisely described by a sinusoidal function; the coefficients are certainly not sparse if we preselect basis functions as power functions. Thus, the appropriate selection of base functions (i.e., parameter clusters) is of great importance to sparse optimization.

The detailed flowchart of this data-driven method, which includes four successive steps, i.e., collecting data, prespecifying variations, and constructing and identifying integrands, is depicted in Fig. 1. In what follows, several representative examples are investigated at length to illustrate the application and efficacy of the established method.

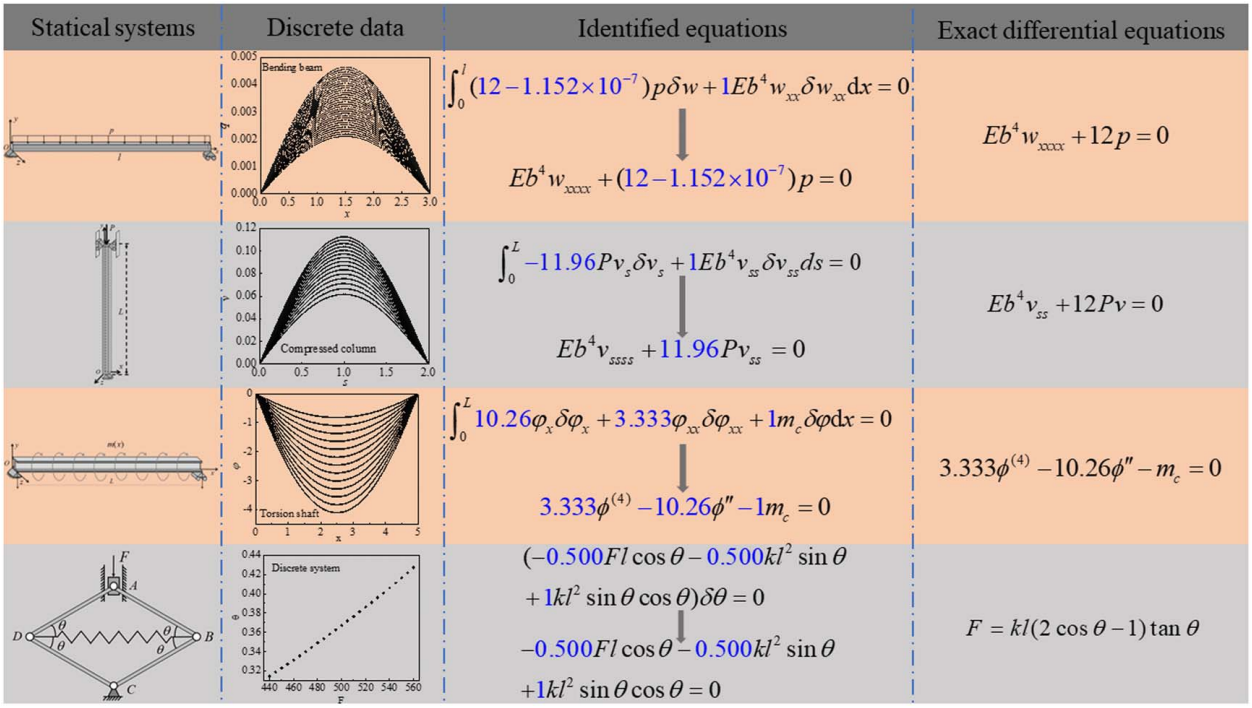

Fig. 2 Summary of governing equations identified for four typical static systems

3 Numerical Examples

As preliminary research, the continuous systems selected are confined to one-dimensional bounded systems; that is, the independent spatial variable **x** is a scalar taking values in a given interval. The continuous systems are selected from classical mechanics, including a uniform-section homogeneous beam imposed by a transverse uniform-distributed load, an elastic slender column compressed by a concentrated force along its axis, and an open thin-walled shaft torqued by a uniform-distributed couple. For all these problems, the discrete data can be captured from the analytical expressions (with deliberately added noises) but not necessarily from experimental measurements. All these examples, including physical models themselves, the discrete data adopted, the governing equations identified, and the associated exact equations, are summarized in Fig. 2.

3.1 Simply Supporting Beam Imposed by Transverse Uniform-Distributed Load. Let us first investigate a simply supported square cross-section beam with a side length of b, a length of l, and an elastic modulus of E. The beam is subjected to a uniformly distributed transverse load with an intensity of p (i.e., force per unit length) along the negative y direction, as shown in the first row and the first column in Fig. 2. The response of concerned is the transverse displacement (i.e., deflection) $w(x)$ of points of material at the middle line; the external load is measured by its intensity p; and the system attributes include the side length, length, and elastic modulus, that is, $\mathbf{s} = [b, l, E]$.

3.2 Collecting Data. Herein, the discrete data of response are directly captured from the analytical expression $w(x) = (px/24EI)(l^3 - 2lx^2 + x^3)$, which can be readily derived by solving the governing differential equation $EIw_{xxxx} + p = 0$, in which $I = b^4/12$ denotes the moment of inertia of the square cross section. We set the values of system attributes as $b = 0.1$ m, $E = 3 \times 10^9$ Pa, and $l = 3$ m, while the intensity of load takes values from 50 N/m to 110 N/m with an increment 2 N/m. Set the sampling interval as 0.01 m. Accordingly, we capture 31 sets of discrete data, with each set including 301 discrete data points. The above parameters will also be adopted below without mentioned otherwise. All the discrete data collected are depicted in the first row and the second column in Fig. 2.

3.3 Prespecifying Variations. Let us first estimate the lowest several-order spatial derivatives of response at boundaries from the discrete data. Various-order derivatives are calculated by the finite difference method, with the forward difference for the left boundary and the backward one for the right boundary. The calculated results are depicted in Fig. 3 for all 31 sets of discrete data (only the first four order derivatives are shown), from which the imposed boundary conditions can be readily extracted. Whether at the left boundary or the right boundary, for all cases, the zero-order derivative (i.e., the response itself) and the second-order derivative remain invariable while the others do not. Whereby the imposed boundary conditions are extracted, that is, $w(0) = w(l) = 0$ and $w_{xx}(0) = w_{xx}(l) = 0$. The variations of actual responses should

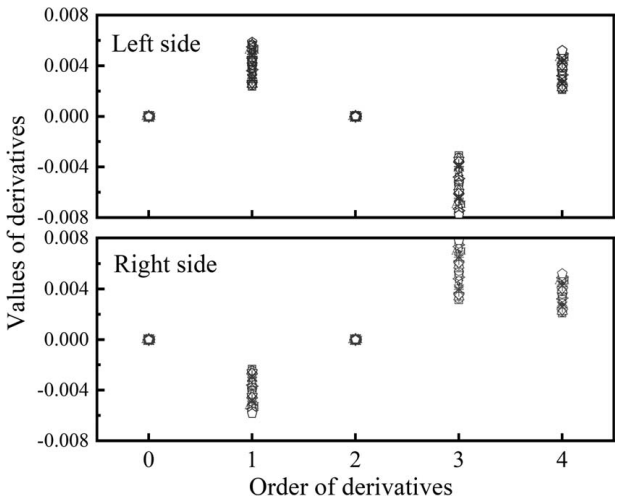

Fig. 3 Various-order spatial derivatives at boundaries for bending beam

satisfy all imposed boundary conditions, i.e., $\delta w(0)=\delta w(l)=0$ and $\delta w_{xx}(0)=\delta w_{xx}(l)=0$. We simply specify the variations as $\sin(n\pi x/l)$, in which n denotes arbitrary positive integers. For instance, by setting $n=1$, 3, 5, 7, we obtain four different variations of the actual response.

3.4 Constructing Integrands. Assume the dimensions of response, load, and system attributes are all known to us. Two integrands U_0 and U (to-be-determined functions) can be constructed with the help of dimension analysis. The dimensions of system attributes, loads, responses, their various-order derivatives, and two integrands, expressed by basic units, are shown in Table 1. Note that the variation of one function possesses the same dimension as the function itself. As a first try, only the lowest two order derivatives and their associated variations are retained. Let us first exclude all terms with negative powers and confine the total power of each term not exceeding 7. Besides, noting that the length l enters the differential-variational equation according to the upper bound of spatial integral, we exclude it in the construction of parameter clusters. As a result, two integrands are expressed as the linear combinations of parameter clusters constructed, that is,

$$\begin{aligned}U_0 = {} & A_1Eb^2 + A_2Ebw + A_3Ew^2 + A_4Eb^2w_x + A_5Ebww_x \\ & + A_6Ew^2w_x + A_7Eb^3w_{xx} + A_8Eb^2ww_{xx} + A_9Ebw^2w_{xx} \\ & + A_{10}Ew^3w_{xx} + A_{11}Eb^3w_{xx}w_x + A_{12}Eb^2ww_{xx}w_x \\ & + A_{13}Ebw^2w_{xx}w_x + A_{14}Ew^3w_{xx}w_x + A_{15}Eb^4w_{xx}^2 + A_{16}Eb^3ww_{xx}^2 \\ & + A_{17}Eb^2w^2w_{xx}^2 + A_{18}Ebw^3w_{xx}^2 + A_{19}Ew^4w_{xx}^2 \end{aligned} \tag{3a}$$

$$\begin{aligned}U = {} & B_1p\delta w + B_2pw_x\delta w + B_3pw_x^2\delta w + B_4pw_x^3\delta w + B_5pb\delta w_x \\ & + B_6pw\delta w_x + B_7pb^2\delta w_{xx} + B_8pbw\delta w_{xx} + B_9pw^2\delta w_{xx} \end{aligned} \tag{3b}$$

3.5 Identifying Integrands. Substituting the expanded expressions of integrands in Eq. (3) into the differential-variational Eq. (1), implementing the variational calculus, deleting linearly dependent terms, and rearranging terms according to variations δw, δw_x, and δw_{xx} yield

$$\begin{aligned}\int_0^l (& C_1Eb + C_2Ew^2w_{xx} + C_3Ebw_x + C_4Eww_x + C_5Ebww_{xx} \\ & + C_6Eb^2w_xw_{xx} + C_7Ebww_xw_{xx} + C_8Ew^2w_xw_{xx} + C_9Eb^3w_{xx}^2 \\ & + C_{10}Eb^2ww_{xx}^2 + C_{11}Ebw^2w_{xx}^2 + C_{12}Ew^3w_{xx}^2 + C_{13}p \\ & + C_{14}pw_x + C_{15}pw_x^2 + C_{16}pw_x^3)\delta w + (C_{17}Eb^3w_{xx} \\ & + C_{18}Ew^2 + C_{19}Eb^2ww_{xx} + C_{20}Ebw^2w_{xx} + C_{21}Ew^3w_{xx} \\ & + C_{22}pb + C_{23}pw)\delta w_x + (C_{24}Eb^4w_{xx} + C_{25}Eb^3w_x \\ & + C_{26}Eb^2ww_x + C_{27}Ebw^2w_x + C_{28}Ew^3w_x + C_{29}pb^2 \\ & + C_{30}pbw + C_{31}pw^2)\delta w_{xx}\mathrm{d}x = 0 \end{aligned} \tag{4}$$

in which the coefficients C_i are linear functions of A_i and B_i. Plugging the discrete data captured and various variations prespecified into Eq. (4), we obtain a set of overdetermined linear algebraic equations with respect to coefficients $C_i (i=1, 2,\ldots, 31)$. The total number of equations is $31\times4=124$, i.e., the number of cases (31) times the number of variations (4). Solving the overdetermined algebraic equations by the pseudo-inverse algorithm and substituting the coefficients obtained into Eq. (4) yield the differential-variational equation to be identified, that is,

$$\begin{aligned}\int_0^l (& 3.98\times10^{-10}Ew^2w_{xx} - 8.58\times10^{-11}Ebww_{xx} - 9.53\times10^{-9}Eb^3w_{xx}^2 \\ & + 5.65\times10^{-8}Eb^2ww_{xx}^2 - 3.16\times10^{-7}Ebw^2w_{xx}^2 \\ & + 1.75\times10^{-6}Ew^3w_{xx}^2 + 12.00p)\delta w + 1.28\times10^{-7}pw\delta w_x \\ & + (1Eb^4w_{xx} + 9.65\times10^{-10}pb^2 - 6.47\times10^{-7}pw^2)\delta w_{xx}\mathrm{d}x = 0 \end{aligned} \tag{5}$$

Furthermore, sparse optimization is implemented to derive the sparsest differential-variational equation, which balances the precision of prediction and the complexity of the model. The weighted factor ρ is set as 1.00×10^{-8} in this example, and the detailed iterative process, including the values of coefficients and the associated value of residual at each iterative step, is shown in Fig. 4. Associated with the sparsest solution, the sparsest differential-variational equation is finally identified, that is,

$$\int_0^l (12 - 1.152\times10^{-7})p\delta w + 1Eb^4w_{xx}\delta w_{xx}\mathrm{d}x = 0 \tag{6}$$

The associated differential equation can be readily derived by direct calculus of variations, that is, $Eb^4w_{xxxx} + (12 - 1.152\times10^{-7})\,p = 0$, which agrees very well with the exact differential equation $E(b^4/12)w_{xxxx} + p = 0$ derived from the classical theory of one-dimensional elastic bodies.

Here, we introduce a relative error to quantify the precision of identification, which is defined as

$$\text{Error} = \frac{\sum_i (C_i - C_i^{\text{exact}})^2}{\sum_i (C_i^{\text{exact}})^2} \tag{7}$$

in which C_i are the coefficients identified, while C_i^{exact} are the exact coefficients derived from classical theory. Now, let us investigate the influence of data noise, sampling interval, and number of cases on the precision of identification.

Figure 5 depicts the relation between the relative error and data noise. The noise added in discrete data of response is set as a Gaussian random variable with zero mean and standard deviation σ and is measured by its standard deviation. Undoubtedly, the relative error almost monotonically increases with the standard deviation of noise. From the noisy data with $\sigma=1\times10^{-5}$, as shown in Fig. 6, the proposed method successfully identifies the equation with high precision; the equation identified is $\int_0^l (12 - 5.09\times10^{-3})\,p\delta w + 1Eb^4w_{xx}\delta w_{xx}\mathrm{d}x = 0$ with the relative error $\text{Error} = 1.802\times10^{-7}$. This method fails as the standard deviation of noise exceeds 1×10^{-4}. The relations of the relative error to the sampling interval and the number of cases are depicted in Figs. 7 and 8, respectively. Obviously, with the increase of sampling interval, the data amount decreases, and the relative error monotonically increases. One convincing reason is that for the situation with a larger sampling interval, the spatial integration evaluated by discrete summation is with lower accuracy. Even from quite scarce data with a sampling interval of 0.07 m, as shown in Fig. 9, this method successfully identifies the equation with high precision.

Table 1 Dimensions of system attributes, load, response, its various-order derivatives, and two integrands for bending beam

Basic unit	$E(=s_1)$	$b(=s_2)$	$l(=s_3)$	$w(\delta w)$	$w_x(\delta w_x)$	$w_{xx}(\delta w_{xx})$	p	U_0	U
[M]	1	0	0	0	0	0	1	1	1
[L]	−1	1	1	1	0	−1	0	1	1
[T]	−2	0	0	0	0	0	−2	−2	−2

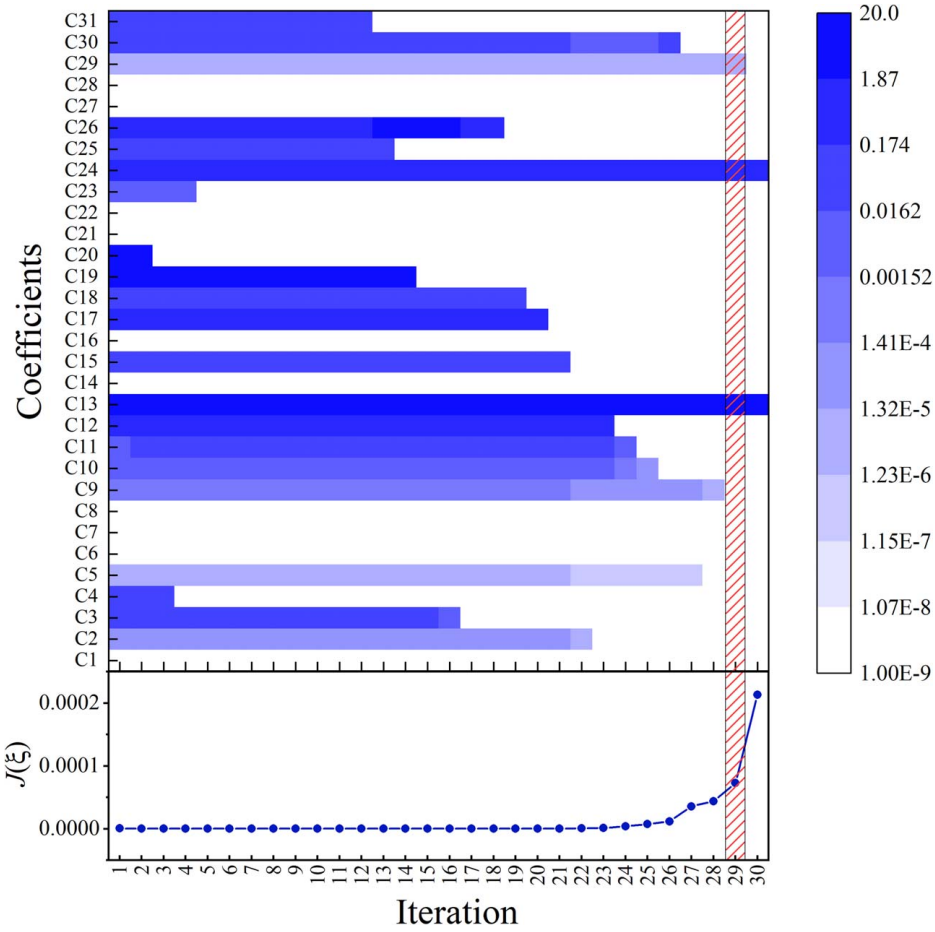

Fig. 4 Iterative process of sparse optimization for bending beam

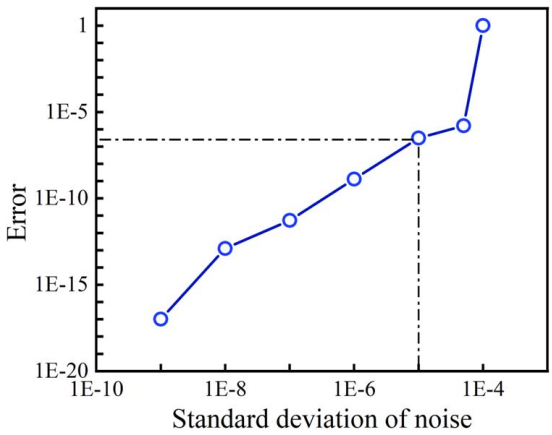

Fig. 5 Relative error versus standard deviation of noise for bending beam

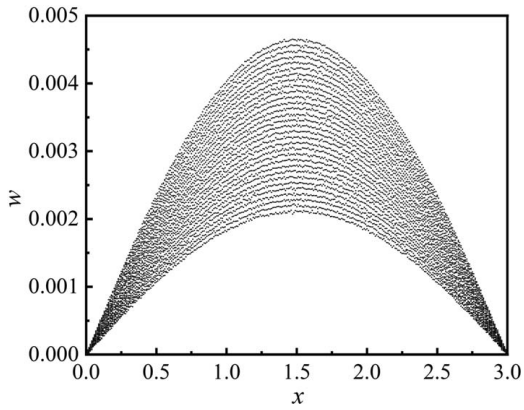

Fig. 6 Noisy data with given standard deviation of noise $\sigma = 1 \times 10^{-5}$ for bending beam

Judged from Fig. 8, with the increase in the number of cases, the relative error decreases, while the precision of identification increases on the whole. There are 31 to-be-determined coefficients while the number of variations is set as 4; thus, the least number of cases should be set as 8 to construct a set of overdetermined algebraic equations to determine these 31 coefficients.

3.6 Case II: Simply Supported Column Imposed by Compressive Load Along Its Axis. Next, we turn to another continuous system, that is, a simply supported column with a square section subjected to a compressive force along its axis, as shown in the second row and the first column in Fig. 2. What we concerned, i.e., the response, is selected as the transverse displacement v of material points at the middle line as the function of arc length s measured from one end; the input is the compressive load P; while the system attributes involve elastic modulus E, length L, and length of the side of square section b. The discrete data of response can be captured by numerically solving the differential equation $((d^2v/ds^2)/\sqrt{1-(dv/ds)^2}) + (Pv/EI) = 0$ [28] with the moment of inertia of the cross section $I = b^4/12$. For simplicity, the discrete data of response are collected from the following approximate analytical solution $v(s) = (2\sqrt{2}L/\pi)\sqrt{(P/F_{cr})-1} \times \sin(\pi s/L)$ with the Eulerian critical load $F_{cr} = \pi^2 EI/L^2$. The values of system attributes

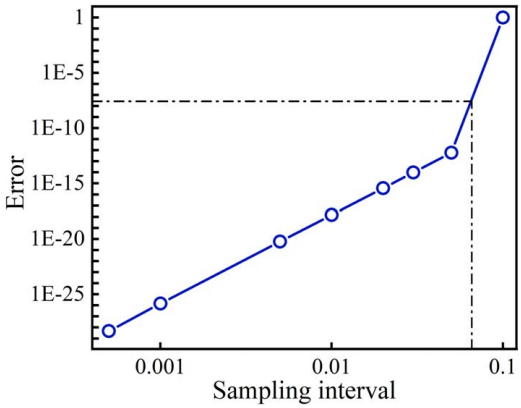

Fig. 7 Relative error versus sampling interval for bending beam

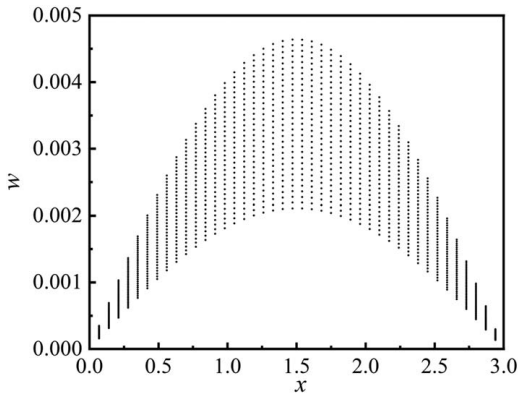

Fig. 8 Relative error versus number of cases for bending beam

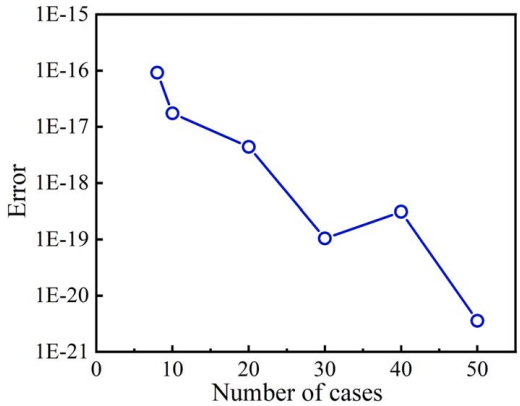

Fig. 9 Discrete data with given sampling interval 0.07 m for bending beam

are set as $E=2\times10^9$ Pa, $L=2$ m, and $b=0.05$ m. The load P, i.e., the compressive force, is equally taken values from the interval [2573, 2580]N with an increment of 0.5N. Note that the values of the compressive force are larger than the critical load (in this case, $F_{cr}=2570$ N) to guarantee the buckling of the column. The sampling interval is set as 0.001 m. Thus, we capture 15 sets of discrete data (i.e., 15 cases), with each case including 2001 data points. All discrete data captured are shown in the second row and the second column in Fig. 2 and in Supplemental Figure 1 available in the Supplemental Materials on the ASME Digital Collection.

The detailed identifying processes, including extracting all imposed boundary conditions, prespecifying variations of actual response, constructing linear expansions of integrands, identifying coefficients, and further sparsification, are shown in the Supplemental Materials available in the Supplemental Materials. The sparsest differential-variational equation identified is expressed as

$$\int_0^L -11.96Pv_s\delta v_s + 1Eb^4v_{ss}\delta v_{ss}ds = 0 \tag{8}$$

The associated differential equation is readily derived as $Eb^4v_{ssss}+11.96Pv_{ss}=0$. If we expand the radical expression in the denominator by Taylor's series, the exact differential equation can be rewritten as

$$\frac{d^2v}{ds^2}\left[1+\frac{1}{2}\left(\frac{dv}{ds}\right)^2+\frac{3}{8}\left(\frac{dv}{ds}\right)^4+\ldots\right]+\frac{Pv}{EI}=0 \tag{9}$$

Obviously, the differential-variational equation identified corresponds to the linearized equation $(d^2v/ds^2)+(Pv/EI)=0$ with $I=b^4/12$, and the differential equation associated with the differential-variational equation identified is just as the second-order spatial derivative of the linearized equation.

The influences of data noise (measured by the standard deviation of noise), the sampling interval, and the number of cases on the precision of identification (measured by the relative error of effective coefficients) are shown in Supplemental Figures 4, 6, and 8, respectively available in the Supplemental Materials. The relative error almost monotonically increases with the standard deviation of noise. From the discrete data with strong noise (set the standard deviation of noise being 5×10^{-4}, for instance, as shown in Supplemental Figure 5 available in the Supplemental Materials), the proposed method gives the identified result with high precision. From Supplemental Figure 6 available in the Supplemental Materials, the relative error increases with the sampling interval on the whole. From 15 sets of discrete data with one set only including 41 data points ($15\times41=615$ data points in total), as shown in Supplemental Figure 7 available in the Supplemental Materials, the identified result still keeps high precision. Undoubtedly, with the increase in the number of cases, the discrete data that can be utilized increase, and the relative error decreases monotonically.

3.7 Case III: Constrained Torsion of Simply Supported Open Thin-Walled Shaft. As a third representative example, let us consider the constrained torsion of an open thin-walled shaft. *Thin-walled* means that the wall thickness of the cross section is much less than the apparent dimension of the cross section; *Open* means the central lines of walls of the cross section do not constitute a simple closed curve. The shaft in the problem is shown in the third row and the first column in Fig. 2, while its I-type cross section with all details is shown in Supplemental Figure 9 available in the Supplemental Materials . The shaft is simply supported at two endpoints and subjected to a uniformly distributed couple m_c along its axis. The response of interest is the angle of twist of the cross section φ; the load is the uniformly distributed couple m; the system attributes include elastic modulus E, shear modulus G, length of the shaft L, and parameters h, b, t describing the geometrical characteristic of the cross section. The discrete data of the angle of twist are captured from the analytical solution of the governing differential equation $GI_t\varphi''-EI_\omega\varphi^{(4)}=m_c$ with $I_t=\sum_{i=1}^{n}(1/3)h_it_i^3$, $I_\omega=(1/24)b^3h^2t$ [29], that is, $\varphi=\frac{-m_cL}{2\alpha GI_t}[\alpha x-\sinh(\alpha x)]-\frac{m_c\sinh(\alpha x)}{\alpha^2GI_t}\left[\frac{\alpha L}{2}-\tanh\left(\frac{\alpha L}{2}\right)\right]+\frac{m_c}{\alpha^2GI_t}\left[\frac{\alpha^2x^2}{2}-\cosh(\alpha x)+1\right]$ with $\alpha^2=GI_t/EI_\omega$. By setting different values of the moment of the couple, 13 sets of discrete data (as shown in Supplemental Figure 10 available in the Supplemental Materials) are captured for subsequent analysis.

Considering that there are too many terms with the same dimension (such as the dimension of length), we have abandoned dimension analysis and directly construct the to-be-determined integrands

by the linear combinations of power functions of the angle of twist, their various-order spatial derivatives, and the associated variations. Please refer to the Supplemental Materials available in the Supplemental Materials for all details of the implementation. The sparsest differential-variational equation identified is

$$\int_0^L 10.26\varphi_x\delta\varphi_x + 3.333\varphi_{xx}\delta\varphi_{xx} + 1m_c\delta\varphi\,\mathrm{d}x = 0 \tag{10}$$

The associated differential equation is readily derived as $3.333\phi_{xxxx} - 10.26\phi_{xx} + m_c = 0$. For the values of system attributes, as shown in the Supplemental Materials available in the Supplemental Materials, the exact governing equation is expressed as $10.26\varphi'' - 3.333\varphi^{(4)} = m_c$. The differential-variational equation identified agrees very well with the exact governing equation.

Regarding the relation of the relative error to data noise, sampling interval, and number of cases, similar conclusions are drawn in Supplemental Figures 13–17 available in the Supplemental Materials.

4 Applications in Discrete Static Systems

Unlikely continuous systems with an infinite degree-of-freedom, discrete systems possess only a finite degree-of-freedom. From this viewpoint, discrete static systems can be regarded as a degeneration of continuous static systems. As a matter of fact, however, in the process of development of natural science, the pioneers first observed the phenomena and established the fundamental theories of static systems and then *generalized* them to continuous static systems. In the differential-variational description, as the crucial distinguishment, the continuous systems involve spatial integration while the discrete systems do not. In addition, for continuous systems, the variations of actual responses must satisfy all imposed boundary conditions and keep sufficient smoothness. In contrast, for discrete systems, the variations must be compatible with each other. If we successfully choose generalized coordinates, i.e., a set of coordinates that are mutually independent and sufficient to reconstitute all possible configurations of the system in question, then the variations can be prespecified arbitrarily.

The fundamental assumption is converted to the following: for a static discrete system, the actual responses described by generalized coordinates satisfy the differential-variational equation of the form

$$\delta U_0(\mathbf{s};\mathbf{q}) + U(\mathbf{f};\mathbf{s};\mathbf{q};\delta\mathbf{q}) = 0 \tag{11}$$

Our objective is to identify the explicit expression of the above equation (i.e., two functions U_0, U with respect to system attributes $\mathbf{s}$, loads $\mathbf{f}$, responses $\mathbf{q}$ and its variations $\delta\mathbf{q}$, with the dimension of energy) only from the discrete data captured.

We will use a static system of single degree-of-freedom (a link mechanism in the horizontal plane) as a representative to illustrate all details of identifying the differential-variational equations of discrete static systems. The link mechanism consists of a mass block that can slide along a vertical track, four links of equal length, and a linear spring with an initial length equal to the length of the links, as shown in the fourth row and the first column in Fig. 2. Under a vertical force imposed on the mass block, the system subsists equilibrium at some configuration. The configuration of the system can be uniquely determined by the angle θ between links and the horizontal line; thus, it is reasonable to select this angle as the response of interest; the vertical force F is set as the load; and system attributes include the length l of links (also the initial length of spring), linear stiffness k of spring, and the mass of block m; thus, $\mathbf{s} = [l, k, m]$.

We set the length $l = 0.5$ m, linear stiffness $k = 3000$ N/m, and mass of block $m = 2$ kg. The vertical force is taken values from the range [440, 560] N with a fixed increment of 5 N. Then, 25 values of the angle are captured for 25 different values of the vertical force from the analytical relation $F = kl(2\cos\theta - 1)\tan\theta$, as shown in the fourth row and the second column in Fig. 2. The variations of the actual response (i.e., the angle) can be prespecified as arbitrary non-zero small constants. Next, we construct two to-be-determined functions using or not using the dimension analysis.

First, the dimensions of all quantities are assumed to be known to us. The dimensions of system attributes, load, response, and to-be-determined functions expressed by basic units are shown in Table 2. By the rule of dimensional consistency, we construct parameter clusters with the dimension of energy and expand two to-be-determined functions as the linear combinations of associated parameter clusters, that is,

$$U_0 = A_1kl^2 + A_2kl^2\sin\theta + A_3kl^2\cos\theta + A_4kl^2\sin^2\theta + A_5kl^2\cos^2\theta \tag{12a}$$

$$U = B_1Fl\delta\theta + B_2Fl\sin\theta\delta\theta + B_3Fl\cos\theta\delta\theta + B_4Fl\sin^2\theta\delta\theta + B_5Fl\cos^2\theta\delta\theta \tag{12b}$$

Substituting the expanded expressions in Eqs. (12*a*) and (12*b*) into the differential-variational Eq. (11) and implementing simple calculus of variations yield

$$(C_1Fl + C_2Fl\sin\theta + C_3Fl\cos\theta + C_4Fl\sin^2\theta + C_5Fl\cos^2\theta + C_6kl^2\cos\theta + C_7kl^2\sin\theta + C_8kl^2\sin\theta\cos\theta)\delta\theta = 0 \tag{13}$$

By plugging the discrete data captured (25 values of the angle and their associated values of the vertical force) into the above equation, we establish a set of overdetermined linear algebraic equations (25 in total) regarding undetermined coefficients $C_i(i = 1, \cdots 8)$. Solving these equations and implementing sparsification yield the sparsest differential-variational equation, that is,

$$(-0.500Fl\cos\theta - 0.500kl^2\sin\theta + 1kl^2\sin\theta\cos\theta)\delta\theta = 0 \tag{14}$$

If the dimensions of all quantities, at least part of them, are assumed to be unknown to us, we can simply construct two functions by the simple and low-order functions of angle θ; for instance, two functions can be expressed as the following expanded expressions:

$$U_0 = A_1\theta + A_2\theta^2 + A_3\sin\theta + A_4\cos\theta + A_5\sin^2\theta + A_6\cos^2\theta \tag{15a}$$

$$U = B_1F\delta\theta + B_2F\theta\delta\theta + B_3F\theta^2\delta\theta + B_4F\sin\theta\delta\theta + B_5F\cos\theta\delta\theta + B_6F\sin^2\theta\delta\theta + B_7F\cos^2\theta\delta\theta \tag{15b}$$

According to similar procedures, the sparsest differential-variational equation is identified, that is,

$$(3.333\times10^{-4}F\cos\theta + 0.500\sin\theta + 1\sin\theta\cos\theta)\delta\theta = 0 \tag{16}$$

The sole difference between these two identified equations consists of the following: the former (Eq. (14)) explicitly includes

Table 2 Dimensions of system attributes, load, response, and two functions for link mechanism

Basic unit	$m(=s_1)$	$l(=s_2)$	$k(=s_3)$	$\theta(\delta\theta)$	F	U_0	U
[M]	1	0	1	0	1	1	1
[L]	0	1	0	0	1	2	2
[T]	0	0	−2	0	−2	−2	−2

system attributes while the latter (Eq. (16)) does not. Obviously, the identified equations in Eqs. (14) and (16) are almost the same one with the exact relation between response (the angle) and load (the force), i.e., $F = kl(2\cos\theta - 1)\tan\theta$.

5 Conclusions and Discussion

This work is devoted to establishing a data-driven method for automatically identifying the differential-variational equations governing the static behaviors of continuous or discrete systems. The procedure starts from a few sets of discrete data of responses and loads, as well as values of system attributes if available, identifies all imposed boundary conditions and prespecifying variations of actual responses, constructs parameter clusters and to-be-determined functions, solves the overdetermined equations and implements sparsification, and finally extracts the sparsest differential-variational equations. Four numerical examples, including three continuous systems and one discrete system, were adopted to demonstrate its applications and efficacy.

This method is built on a fundamental assumption, i.e., the actual response of a static system satisfies a differential-variational equation, while two functions as the determinative of responses possess the dimension of energy; thus, it undoubtedly belongs to the field of theory-guided data science. The differential-variational framework provides some advantages, some of them listed below. Compared to differential equations typically adopted in equation identification, differential-variational equations include spatial integration and only involve lower-order spatial derivatives. The implementation of spatial integration and the evaluation of only lower-order spatial derivatives dramatically improve the robustness of data noise. Furthermore, the arbitrariness of variation selection helps to increase, even infinitely, the scale of overdetermined equations, helping the sufficient utilization of information hidden in data and the improvement of precision of identification.

Nevertheless, there also remain some issues to be addressed. As a prerequisite of this method, the response of a system in question must be determined in advance. This is a difficult work but is always overlooked. Taking a beam as an example, as mentioned above, it essentially occupies a finite three-dimensional space, while its response (i.e., the vector of displacement) includes three components along three mutually orthogonal directions. In this work, however, the response is directly designated as one of these three components, that is, the transverse displacement (the deflection) guided by the classical theory of mechanics of materials. Assuming we select inappropriate or even incomplete quantities as responses, this method may generate complex (non-sparse) or even erroneous results. Selecting appropriate, especially low-dimensional, responses involves another important topic in science and engineering, that is, dimension reduction [30]. Our ultimate objective is to automatically select the low-dimensional described quantities of responses and automatically extract the governing equations relating the responses to loads, that is, achieve the total automation of equation identification. This will be our future work.

Acknowledgment

Z.H. acknowledges the National Natural Science Foundation of China under Grant No. 12132013. Y.W. acknowledges the National Natural Science Foundation of China under Grant Nos. 12372036 and 11872328. H.J. acknowledges the support from Westlake University.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Hey, T., Tansley, S., and Tolle, K. M., 2009, *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Microsoft Research, Redmond, WA.
[2] Bongard, J., and Lipson, H., 2007, "Automated Reverse Engineering of Nonlinear Dynamical Systems," Proc. Natl. Acad. Sci. U.S.A, **104**(24), pp. 9943–9948.
[3] Schmidt, M., and Lipson, H., 2009, "Distilling Free-Form Natural Laws From Experimental Data," Science, **324**(5923), pp. 81–85.
[4] Quade, M., Abel, M., Shafi, K., Niven, R. K., and Noack, B. R., 2016, "Prediction of Dynamical Systems by Symbolic Regression," Phys. Rev. E, **94**(1), p. 012214.
[5] Brunton, S. L., Proctor, J. L., and Kutz, J. N., 2016, "Discovering Governing Equations From Data by Sparse Identification of Nonlinear Dynamical Systems," Proc. Natl. Acad. Sci. U. S. A., **113**(15), pp. 3932–3937.
[6] Wang, W.-X., Yang, R., Lai, Y.-C., Kovanis, V., and Grebogi, C., 2011, "Predicting Catastrophes in Nonlinear Dynamical Systems by Compressive Sensing," Phys. Rev. Lett., **106**(15), p. 154101.
[7] Schaeffer, H., and McCalla, S. G., 2017, "Sparse Model Selection via Integral Terms," Phys. Rev. E, **96**(2), p. 023302.
[8] Udrescu, S.-M., and Tegmark, M., 2020, "AI Feynman: A Physics-Inspired Method for Symbolic Regression," Sci. Adv., **6**(16), p. eaay2631.
[9] Udrescu, S.-M., Tan, A., Feng, J., Neto, O., Wu, T., and Tegmark, M., 2020, "AI Feynman 2.0: Pareto-Optimal Symbolic Regression Exploiting Graph Modularity," Advances in Neural Information Processing Systems, Vancouver, Canada.
[10] Huang, Z., Tian, Y., Li, C., Lin, G., Wu, L., Wang, Y., and Jiang, H., 2020, "Data-Driven Automated Discovery of Variational Laws Hidden in Physical Systems," J. Mech. Phys. Solids, **2020**(137), p. 103871.
[11] Reinbold, P. A., Gurevich, D. R., and Grigoriev, R. O., 2020, "Using Noisy or Incomplete Data to Discover Models of Spatiotemporal Dynamics," Phys. Rev. E, **101**(1), p. 010203.
[12] Chen, R. T., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K., 2018, "Neural Ordinary Differential Equations," Advances in Neural Information Processing Systems, Montréal, Canada.
[13] Dupont, E., Doucet, A., and Teh, Y. W., 2019, "Augmented Neural ODEs," Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver Canada, Dec. 8, Curran Associates Inc., Article 282.
[14] Cranmer, M., Greydanus, S., Hoyer, S., Battaglia, P., Spergel, D., and Ho, S., 2020, "Lagrangian Neural Networks," arXiv preprint arXiv:2003.04630.
[15] Mattheakis, M., Sondak, D., Dogra, A. S., and Protopapas, P., 2022, "Hamiltonian Neural Networks for Solving Equations of Motion," Phys. Rev. E, **105**(6), p. 065305.
[16] Greydanus, S., Dzamba, M., and Yosinski, J., 2019, "Hamiltonian Neural Networks," Advances in Neural Information Processing Systems 32 (NeurIPS 2019), Vancouver Canada.
[17] Finzi, M., Wang, K. A., and Wilson, A. G., 2020, "Simplifying Hamiltonian and Lagrangian Neural Networks via Explicit Constraints," Advances in Neural Information Processing Systems 33 (NeurIPS 2020), Virtual-only Conference.
[18] Li, Z., Ji, J., and Zhang, Y., 2021, "From Kepler to Newton: Explainable AI for Science Discovery," arXiv preprint arXiv:2111.12210.
[19] Liu, Z., Wang, B., Meng, Q., Chen, W., Tegmark, M., and Liu, T.-Y., 2021, "Machine-Learning Nonconservative Dynamics for New-Physics Detection," Phys. Rev. E, **104**(5), p. 055302.
[20] Mukhopadhyay, S. C., Jayasundera, K. P., and Postolache, O. A., 2018, *Modern Sensing Technologies*, Springer, New York.
[21] Li, C., Huang, Z., Wang, Y., and Jiang, H., 2021, "Rapid Identification of Switched Systems: A Data-Driven Method in Variational Framework," Sci. China Technol. Sci., **64**(1), pp. 148–156.
[22] Li, C., Huang, Z., Huang, Z., Wang, Y., and Jiang, H., 2022, "Digital Twins in Engineering Dynamics: Variational Equation Identification, Feedback Control Design and Their Rapid Update," Nonlinear Dyn., **2023**(111), pp. 1–16.
[23] Whittaker, E. T., 1964, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, CUP Archive.
[24] Brunton, S. L., and Kutz, J. N., 2022, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, Cambridge University Press.
[25] Tibshirani, R., 1996, "Regression Shrinkage and Selection via the Lasso," J. R. Stat. Soc., B: Stat. Methodol., **58**(1), pp. 267–288.
[26] Boninsegna, L., Nüske, F., and Clementi, C., 2018, "Sparse Learning of Stochastic Dynamical Equations," J. Chem. Phys., **148**(24), p. 241723.
[27] Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H., 2009, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York.
[28] Timoshenko, S., 1970, *Theory of Elastic Stability 2e*, Tata McGraw-Hill Education.
[29] Gjelsvik, A., 1981, *The Theory of Thin Walled Bars*, Wiley, New York.
[30] Huang, Z., Li, C., Huang, Z., Wang, Y., and Jiang, H., 2021, "AI-Timoshenko: Automatedly Discovering Simplified Governing Equations for Applied Mechanics Problems From Simulated Data," ASME J. Appl. Mech., **88**(10), p. 101006.