Discovering time-varying aerodynamics of a prototype bridge by sparse identification of nonlinear dynamical systems

Shanwu Li,[1] Eurika Kaiser,[2] Shujin Laima,[1] Hui Li,[1] Steven L. Brunton,[2] and J. Nathan Kutz[3,*]
[1]*School of Civil Engineering, Harbin Institute of Technology, Harbin, China*
[2]*Department of Mechanical Engineering, University of Washington, Seattle, Washington 98195-2420, USA*
[3]*Department of Applied Mathematics, University of Washington, Seattle, Washington 98195-2420, USA*

(Received 29 May 2019; published 21 August 2019)

Vortex-induced vibrations (VIVs) have been observed on a long-span suspension bridge. The nonstationary wind in the field characterized by the time-varying mean wind speed is likely to lead to time-varying aerodynamics of the wind-bridge system during VIVs, which is different from VIVs induced by stationary or even steady wind in wind tunnels. In this paper, data-driven methods are proposed to reveal the time-varying aerodynamics of the wind-bridge system during VIV events based on field measurements on a long-span suspension bridge. First, a variant of the *sparse identification of nonlinear dynamics* algorithm is proposed to identify parsimonious, time-varying aerodynamical systems that capture VIV events of the bridge. Thus we are able to posit new, data-driven, and interpretable models highlighting the aeroelastic interactions between the wind and bridge. Second, a density-based clustering algorithm is applied to discovering the potential modes of dynamics during VIV events. As a result, the time-dependent model is obtained to reveal the evolution of the aerodynamics of the wind-bridge system over time during an entire VIV event. It is found that the level of self-excited effects of the wind-bridge system is significantly time varying with the real-time wind speed and bridge motion state. The simulations of VIVs by the obtained time-dependent models show high accuracies of the models with an averaged normalized mean-square error of 0.0023. The clustering of obtained models shows underlying distinct dynamical regimes of the wind-bridge system, which are distinguished by the level of self-excited effects.

DOI: 10.1103/PhysRevE.100.022220

I. INTRODUCTION

Through improved sensors and the emerging structural health monitoring (SHM) system, it is now possible to continuously assess modern bridge performance in real time. Not only is it critical that bridges be monitored, e.g., for structure monitoring and safety, but the rich time-series recordings provided by the sensors allow bridge engineers to gain a new understanding of the aerodynamics of prototype bridges subjected to real wind. By leveraging sparse regression techniques, the so-called *sparse identification of nonlinear dynamics* (SINDy) method provides a new paradigm for data-driven model discovery [1]. The emergence of the SINDy algorithm is allowing researchers to discover governing equations by sampling either the full or partial state space of a given system, respectively. Although nonlinear, data-driven system identification methods such as SINDy are emerging as viable techniques for a broad range of applications, the methods have yet to be applied to the complex aeroelastic interactions observed on bridges. In this paper, we leverage (i) time-series measurements from the SHM system on a long-span suspension bridge and (ii) the SINDy model discovery architecture to build data-driven models of the wind-bridge system to reveal and interpret the time-varying aerodynamics during vortex-induced vibration (VIV) events. We find that the SINDy architecture is effective in identifying parsimonious, time-varying dynamical systems which result from VIV events of the bridge. Thus we are able to posit new, data-driven and interpretable models highlighting the time-varying aerodynamics of a long-span bridge during VIV events subjected to time-varying wind.

The conventional study of bridge aerodynamics is essentially composed of theoretical analysis, wind tunnel tests, and computational fluid dynamics (CFD). The complex wind loading and fluid-structure interaction of a long-span bridge result in a variety of aerodynamical phenomena such as buffeting [2–5], VIV [6–9], and flutter [10,11]. Tremendous advances in theoretical analysis, wind tunnel tests, and computational modeling have made significant contributions to characterizing bridge aerodynamics. Wind tunnel tests with cylinders, simplified sectional models, or scaled, full aeroelastic models are combined with theoretical analysis to discover bridge aerodynamics [12–16], leading to simplified semi-empirical models and a number of corresponding aerodynamic and aeroelastic parameter identifications [17–26]. In addition, significant progress has been made in computational models, resulting in a number of CFD-based methods [27–29]. Such CFD-based models are typically idealized versions of the bridge itself. Our proposed model discovery architecture can also help augment bridge models using the data acquired in field testing for discrepancy modeling [30].

In recent years, the emerging SHM systems on long-span bridges provide an opportunity to study bridge aerodynamics

*kutz@uw.edu

based on the full-scale structure subjected to real natural wind, although the field monitoring is still usually limited by the spatial sparsity of measurements and type of measurements. For example, only winds and bridge vibrations can be monitored for wind and wind effects monitoring. Wind loads and wind pressures are difficult to monitor in the field. Therefore, how to mine the dimension-limited data and discover the potential knowledge hidden in the data is the key challenge in studying bridge aerodynamics by field measurements.

Emerging data-driven methods are allowing for the discovery of physical and engineering principles directly from time-series recordings. Our focus is on the SINDy architecture [1], which has been demonstrated on a diverse set of problems, including spatiotemporal [31], parametric [32], networked [33], control [34], and multiscale [35] systems. The underlying algorithms can also be made robust and can accommodate parametrized functional dependencies [36]. Importantly, the SINDy architecture promotes sparsity and parsimony which can be directly related to model selection theory [33] in order to assess the quality and robustness of the model discovered. The SINDy method is computationally efficient and the algorithms for all the innovations mentioned above are available as open source code. An alternative data-driven approach to SINDy uses symbolic regression to identify directly the structure of a nonlinear dynamical system from data [37–39]. This works remarkably well for discovering interpretable physical models, but the symbolic regression is computationally expensive and can be difficult to scale to large problems.

There are numerous alternative approaches to fitting the data with models, including nonsparsity promoting regressions to polynomial and/or special function bases [40,41]. Deep neural networks are yet another approach to data-driven models, allowing for future-state prediction of dynamical systems [42–49]. However, a key limitation of these data-driven methods is the lack of interpretability of the resulting model: They are focused on reconstruction error and fitting the data and do not provide governing equations or clearly interpretable models in terms of the original variable set. Additionally, models that are parametrized by a larger number of terms often do poorly when considering model selection from an information criteria viewpoint, since there is a linear penalty in the total number of terms using either Akaike information criteria (AIC) or Bayesian information criteria (BIC) [33].

Our aim is to use the SINDy architecture to provide interpretable dynamical models that can aid in understanding the time-varying aerodynamics of the wind-bridge system during VIV events. Using field measurements data, we discover the nonlinear aerodynamics that results from wind-bridge interaction during VIV events. In particular, we discover a parsimonious set of governing equations which are time varying. These parsimonious models are also ideal from a model selection viewpoint of AIC and/or BIC. The discovered models allow us to reveal the evolution of aerodynamics with time during VIV events and identify distinct regimes of the aerodynamics. Importantly, the sparsity patterns discovered by SINDy allow one to clearly identify four distinct physical regimes for different intensity VIV events, i.e., different dominant balance nonlinear physics that come into play depending on the strength of the VIV. Such distinct dynamical regimes are difficult to identify in models as they tend to lock-in the wind speed forcing the bridge.

The manuscript is outlined as follows: In Sec. II, VIV events are discussed in detail as they are the central concern affecting the nonlinear bridge aerodynamics. Section III details the bridge field monitoring and data acquisition of the time-series measurements used for model identification. Section IV develops the SINDy architecture for the bridge data of Sec. III. In Sec. V, we discover the distinct dynamical regimes of the bridge-wind system by clustering the discovered models. The paper is concluded in Sec. VI.

II. VORTEX-INDUCED VIBRATION OF LONG-SPAN BRIDGE

A long-span bridge may have intrinsically distinct modes of aerodynamic behavior such as buffeting, VIV, and flutter. For modern bridges, flutter must be avoided in the design stage by increasing the critical flutter wind speed, because of its unique divergent response which results from aeroelastic instability. As a consequence, only buffeting and VIV are observed in modern bridges. Unlike buffeting, VIV involves aeroelastic effects characterized by fluid-structure interactions which result in a possible negative aerodynamic damping, thus generating large vibration amplitudes. VIV occurs during periodic vortex shedding within a range of shedding frequencies near the structural natural frequency. Large-amplitude oscillations occur in this range that appear to control the shedding process in a fluid-structure interaction phenomenon known as *lock-in*.

Comprehensive investigations of the mechanisms responsible for VIV have been performed. Nakamura and Mizota [12] have observed the *lock-in* phenomenon by measuring the lift force and characterizing wakes of rectangular prisms with various aspect ratios oscillating transversely in a uniform flow, with the short sides normal to the flow direction in a wind tunnel. It was found that the phase angles of the frequency response components of both the lift and near-wake velocity show abrupt changes when approaching the critical reduced wind velocity for vortex shedding. This is suggested to be a key phenomenon involved when solving the problem of the vortex excitation of bluff structures. Komatsu and Kobayashi [13] characterized two types of VIV through a series of experiments on various cross sections (such as L-shaped, T-shaped, H-shaped, and rectangular cylinders) with various aspect ratios in a wind tunnel. One is a forced small-amplitude vibration caused by von Kármán vortex shedding in cylinders (T cylinders) with a separation point at the trailing edge. The other is a self-excited vibration with relatively large amplitude in cylinders (L, H, and rectangular cylinders) with a separation point at the leading edge, which occurs independently on the von Kármán vortex street. The generating mechanism in the latter case is described as a motion-induced vortex at the leading edge that synchronizes with the motion of the cylinder. The frequency of this type of vibration does not change within a certain range of wind velocities and coincides with the natural frequency of the cylinder, i.e., the *lock-in* phenomenon. Li *et al.* [16] have investigated the Reynolds number effects on the aerodynamic characteristics and VIV of a twin-box girder within a range of Reynolds number values

(5.85×10^3–1.12×10^5). They find that the transition point of the separated shear layer moves upstream, and the bubble size gradually decreases with increasing Reynolds number values. Such investigations give a strong foundation for a qualitative understanding of VIV and critical fluid-structure interactions.

In addition to understanding fundamental mechanisms, accurate VIV modeling is quite important, especially for the design of a bridge. Rigorous mathematical-physical modeling of VIV requires simultaneously solving the Navier-Stokes (N-S) equations and equations of motion of the structure. However, because of the strong nonlinearity of the N-S equations, this has proven mathematically and computationally intractable [50]. As a less-than-ideal alternative, simplified semiempirical models have been proposed based on wind tunnel tests. To date, the most widely accepted empirical model is proposed by Simiu and Scanlan [9], which is described as

$$m\left(\ddot{y} + 2\zeta\omega_1\dot{y} + \omega_1^2 y\right) = F, \tag{1}$$

with

$$F = \frac{1}{2}\rho U^2(2D) \times \left[\underbrace{Y_1(K)\left(1 - \lambda\frac{y^2}{D^2}\right)\frac{\dot{y}}{U} + Y_2(K)\frac{y}{D}}_{\text{motion induced}} + \underbrace{\frac{1}{2}\tilde{C}_L}_{\text{fluid induced}}\right], \tag{2}$$

where m is mass per unit span length; ω_1 is mechanical circular frequency; ζ is mechanical damping ratio; y is cross-flow displacement; F is aerodynamic force; ρ is air density; U is wind speed and assumed to be time independent; D is cross-flow dimension of the section; $K = \omega D/U$ is the reduced frequency of vortex shedding, where ω is vortex-shedding frequency that satisfies the Strouhal relation, $\omega D/U = 2\pi St$, outside *lock-in* regime; and St is the Strouhal number. The parameters λ, Y_1, Y_2, and $\tilde{C}_L$ have to be determined by calibration to experiments. Specifically, λ is a constant denoting the nonlinear dependence of self-excited force on displacement amplitude, $\tilde{C}_L$ is the stochastic lift force coefficient, and Y_1 and Y_2 are aerodynamic parameters which are functions of the reduced frequency of vortex-shedding K.

The total force in the model consists of two types of forces: One is induced directly by vortex shedding around the bluff body simulated by the third term with $\tilde{C}_L$ in Eq. (2), and the other is a motion-induced lift force represented by the first two terms in Eq. (2) including aerodynamic damping with Y_1 and aerodynamic stiffness with Y_2. The direct forcing term with $\tilde{C}_L$ is found to be small relative to the motion-induced force when large-amplitude oscillations are present [51]. The model (1) thus may be simplified by dropping the direct forcing term and then be nondimensionalized to:

$$\eta''(s) + 2\zeta K_1\eta'(s) + K_1^2\eta(s) = m_r Y_1[1 - \lambda\eta^2(s)]\eta'(s) + m_r Y_2\eta(s), \tag{3}$$

where $\eta = y/D$ is the nondimensionalized cross-flow displacement, $m_r = \rho D^2/m$ is mass ratio, $K_1 = \omega_1 D/U$ is the reduced natural frequency, and primes indicate derivatives with respect to the dimensionless time, $s = Ut/D$.

A solution for the bridge dynamics is then sought in the form:

$$\eta(s) = A(s)\cos[Ks - \psi(s)]. \tag{4}$$

The VIV of a bridge is generally considered as *quasilinear*, i.e., the system has a small amount of nonlinearity where $A(s)$ and $\psi(s)$ are slowly varying functions of dimensionless time s. The solution $\eta(s)$ can then be replaced by two separate solutions for $A(s)$ and $\psi(s)$, which are given as follows:

$$A'(s) = -\frac{1}{8}\alpha A(s)[A^2(s) - \beta^2], \tag{5a}$$

$$\psi' = \frac{1}{2K}\left[m_r Y_2 + \left(K^2 - K_1^2\right)\right], \tag{5b}$$

where $\alpha = m_r Y_1\lambda$, $\beta = (2/\sqrt{\lambda})[1 - (2\zeta K_1)/(m_r Y_1)]^{1/2}$. It should be noted that the wind speed U is assumed to be time independent and is thus reduced during nondimensionalization. Actually, this model is proposed from the wind tunnel test where the wind speed is stationary or even steady during VIV. However, the real wind in the field is usually obviously nonstationary with a time-varying mean wind speed which may get out of VIV wind speed range for a while during an entire VIV event and thus lead to the change of aerodynamics with time. Therefore, this model fails to reveal the possible time-varying aerodynamics of the bridge subjected to time-varying wind speed. Our aim is to find time-dependent models (ordinary differential equations) to highlight and reveal the time-varying aerodynamics.

III. FIELD MEASUREMENTS AND DATA PREPROCESSING

The long-span suspension bridge investigated in this study crosses a narrow water channel that lies between two islands. An SHM system, including wind and vibration monitoring, was implemented in 2009 and has since continuously recorded measurements in real time.

At each side of the bridge section, the wind speed and direction are monitored with anemometers. In particular, Young Model 81000 three-dimensional ultrasonic anemometers with a sampling frequency of 32 Hz are located at 1/4, 1/2, and 3/4 center span (locations are indicated by S1, S2, and S3 in Fig. 1, respectively) on both the upstream and downstream sides. These anemometers are installed on lighting columns at a height of 6 m above the bridge deck surface. The wind data used in this study are all from the inflow anemometers, which can measure natural winds without interference from bridge components. Vertical vibration of the bridge deck is monitored by GT02 force-balance triaxial accelerometers with a sampling frequency of 50 Hz at S1, S2, and S3.

VIV events of this bridge captured by wind and vibration histories were identified using cluster analysis in a previous study [52]. In the present study, we first process the

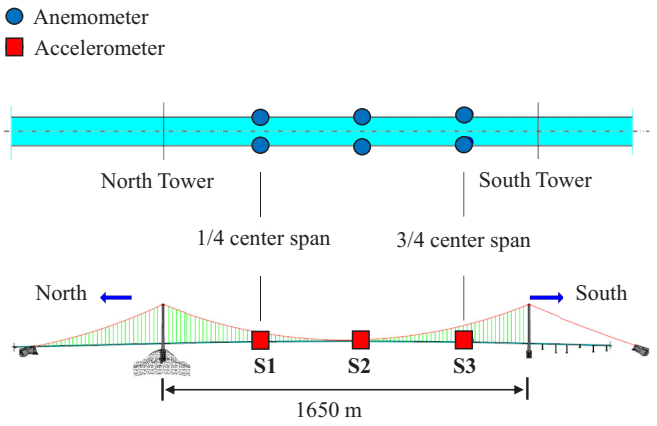

FIG. 1. Field monitoring on the bridge. Anemometers and accelerometers are installed at 1/4, 1/2, and 3/4 center span. Anemometers are installed on both sides of the bridge section.

original data to identify potential key factors accounting for the bridge aerodynamics during VIVs. First, the wind data are preprocessed (see Fig. 2). Histories of the horizontal wind speed V and wind direction θ are obtained from the measured horizontal wind components. The crosswind speed, which is the component perpendicular to the spanwise direction, is obtained by $\tilde{U} = V|\sin(\theta)|$. The time-varying mean wind speed U is determined by applying a low-pass filter to crosswind speed $\tilde{U}$.

Besides, the vibration displacements are obtained by double integrations of acceleration in the frequency domain. The power spectral density (PSD) of displacement is further obtained in Fig. 3. It is found that the vibration amplitude changes slowly with time; the VIV frequency (0.3252 Hz) is almost identical to the natural frequency of the bridge (0.32507 Hz), indicating that $\psi(s)$ is much smaller than Ks in Eq. (4) and does not lead to a frequency change. We thus only need to focus on the time-varying amplitude $A(s)$ in Eq. (4). Accordingly, the ordinary differential equation (ODE) of displacement amplitude A described by Eq. (5) is the key equation describing the VIV aerodynamics. Our aim is to find some time-dependent ODE of A to replace Eq. (5) to highlight the potential time-varying aerodynamics during VIVs subjected to time-varying wind speed.

In normal situations, vehicle and wind are the main causes of bridge vibrations in the field. Vehicle effect on the studied bridge is almost always present except for the closure of the bridge when the typhoon is passing. It is difficult to remove the vehicle-induced vibration from the measured total effect. However, the vehicle-induced vibration is generally much smaller than the wind-induced vibration under strong winds and VIV, which is a kind of resonance. Figure 4 shows the comparison of vehicle-induced vibrations at different times (around 00:00, 06:00, 12:00, and 18:00) and vehicle-wind-induced vibration under strong wind. It should be noted that these four samples are not from the same day because it is rare that the wind speeds at these four times on the same day are all close to zero. It can be found that the vehicle-induced vibration is much smaller than vehicle-wind-induced vibration under strong winds which is smaller than VIV

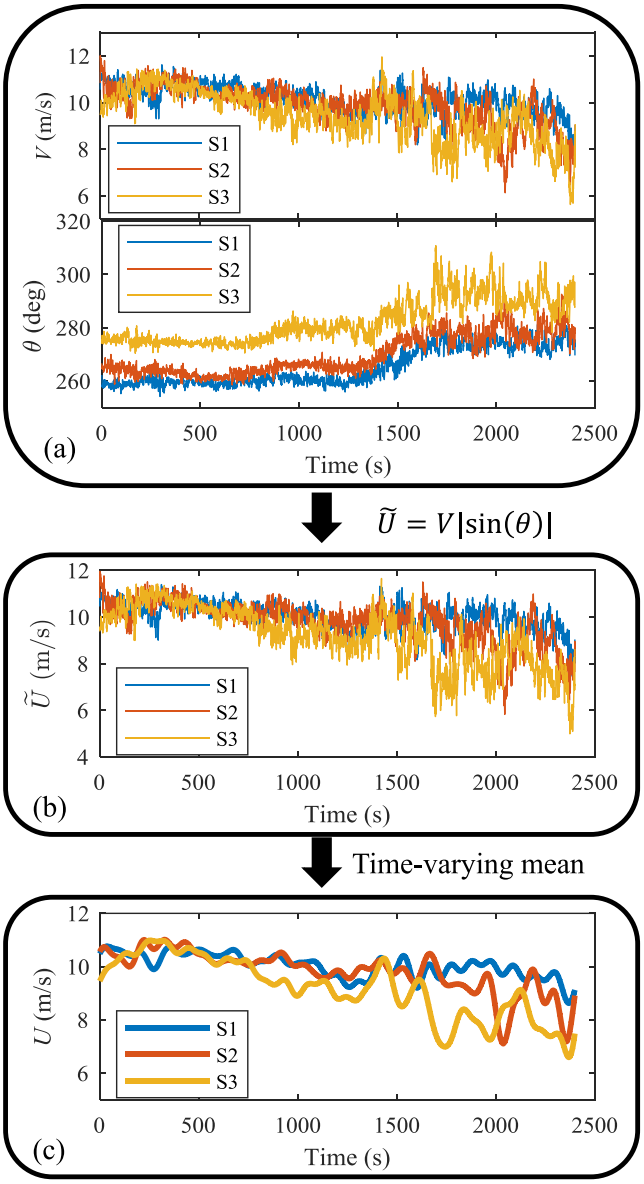

FIG. 2. Preprocessing of wind data. S1, S2, and S3 indicate 1/4 span, midspan, and 3/4 span, respectively, as shown in Fig. 1. (a) Horizontal instantaneous wind speed V and wind direction θ are obtained from original measurements of wind speed; 90° and 270° indicate the perpendicular direction to the spanwise direction. (b) The wind speed component perpendicular to the spanwise direction is determined. (c) The time-varying mean wind speed is estimated by applying a low-pass filter.

[see Fig. 3(b)]. We thus believe that the vehicle effects have little impact on the study of VIV.

Equation (5) needs to be generalized from wind tunnel tests to field measurements by carefully considering two key points: (i) the wind condition during an entire VIV event is nonstationary with time-varying mean wind speed for real VIVs while stationary or even steady in wind tunnel tests and (ii) the spatial dimension of the wind-bridge system for field measurements, which depends on the constellation of the sensors, is higher than the one-dimensional section model typically used in wind tunnel tests. To build our data-driven

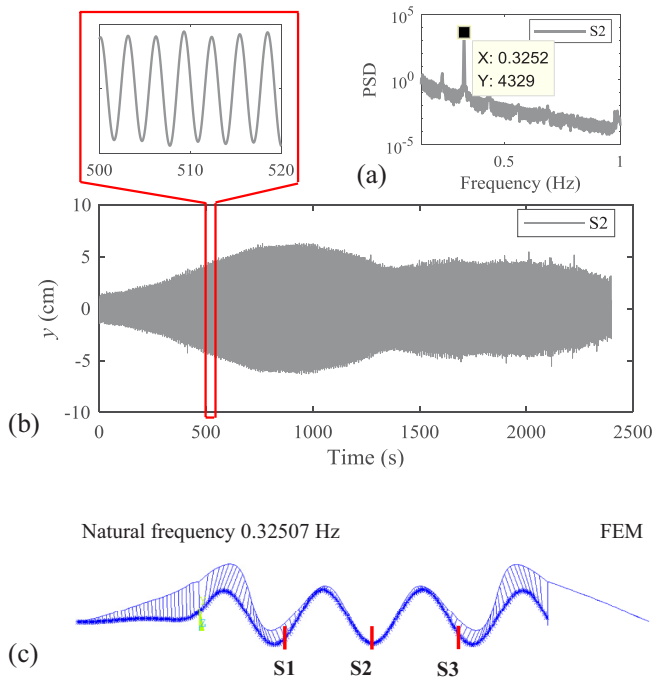

FIG. 3. Time-frequency analysis of measured acceleration for a VIV event. (a) The PSD of the vibration displacement history. (b) Displacement history of a VIV event. (c) The mode shape and the natural frequency of the bridge obtained by an accompanying numerical simulation using FEM.

model and account for these considerations, we extract the envelope of the vibration displacement to obtain the time-varying displacement amplitude A and its time derivative $\dot{A}$ (see Fig. 5).

IV. DATA-DRIVEN MODEL DISCOVERY: SPARSE IDENTIFICATION OF TIME-VARYING AERODYNAMICS OF A LONG-SPAN BRIDGE

We use data-driven model discovery methods to extract improved characterizations of the nonlinear bridge aerodynamics. Our aim is to make maximal use of the time-series data generated by the bridge sensors.

A. The SINDy algorithm

The primary method used for our model discovery is the SINDy algorithm, which leverages advances in machine learning and sparse regression to discover nonlinear dynamical systems from data [1]. SINDy solves an overdetermined linear system of equations by sparsity-promoting regularization. The basic algorithmic structure of SINDy has been modified to discover parametrically dependent systems [32], resolve multiscale physics [35], infer biological networks [33], discover spatiotemporal systems [31], and identify nonlinear systems with control [34,53].

Consider a dynamical system of the form

$$\dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}), \tag{6}$$

where the function $\boldsymbol{f}(\cdot)$ is unknown but assumed to have only a few dominant contributing terms. The SINDy algorithm posits a large set of potential candidate functions that comprise $f(\cdot)$ and then uses a sparsity-promoting regression to determine the dominant terms. The relevant active terms in the dynamics can be solved for using an ℓ_1-regularized regression that penalizes the number of active terms. The general framework for SINDy is shown in Fig. 6(b).

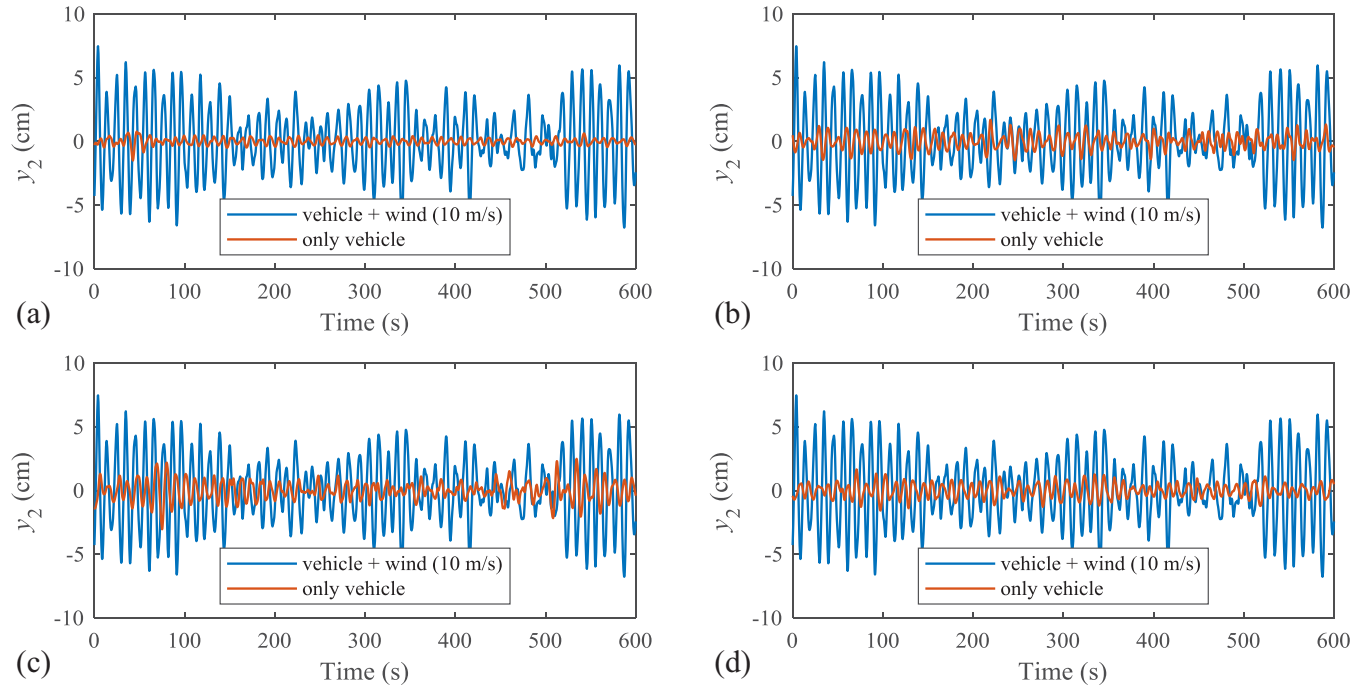

FIG. 4. Comparison of vehicle-induced vibrations (mean wind speed is close to zero) around (a) 00:00, (b) 06:00, (c) 12:00, (d) 18:00 and vehicle-wind-induced vibration under strong wind (mean wind speed is 10 m/s).

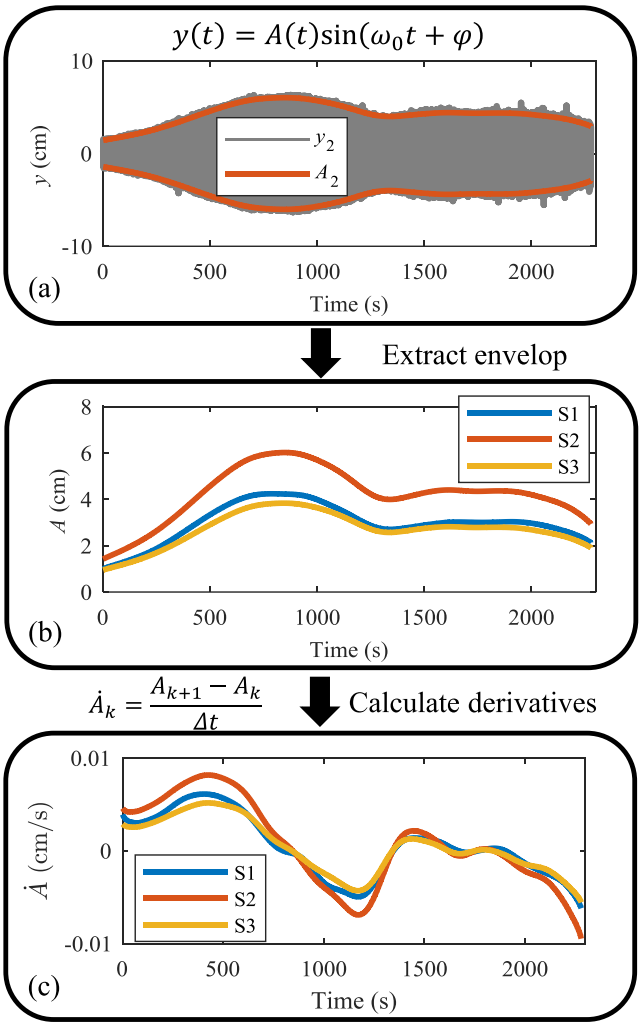

FIG. 5. Preprocessing of vibration data. (a) The time-varying displacement amplitude A is obtained by extracting the envelop from the displacement history y which is obtained by integration of acceleration $\ddot{y}$ in the frequency domain. (b) Vibration amplitudes are obtained for all the three sensor locations. (c) Time derivatives of the amplitudes are obtained.

Sensor measurements are used to collect time-series data which are arranged in the data matrix:

$$\mathbf{X} = [\mathbf{x}(t_1)\ \ \mathbf{x}(t_2)\ \ \cdots\ \ \mathbf{x}(t_m)]^T, \tag{7}$$

where the superscript "T" denotes the matrix transpose. The matrix $\mathbf{X}$ is $m \times n$, where n is the dimension of the state $\mathbf{x} \in \mathbb{R}^n$ and m is the number of measurements of the state in time. Similarly, the matrix of derivatives,

$$\dot{\mathbf{X}} = [\dot{\mathbf{x}}(t_1)\ \ \dot{\mathbf{x}}(t_2)\ \ \cdots\ \ \dot{\mathbf{x}}(t_m)]^T, \tag{8}$$

is collected or computed from the state data in $\mathbf{X}$. Accurate derivatives are critical for model identification, and the total-variation regularized derivative [54] is used as a numerically robust method to compute derivatives from noisy data.

A library of candidate nonlinear functions is constructed from $\mathbf{X}$. This takes the general form

$$\boldsymbol{\Theta}(\mathbf{X}) = [\mathbf{1}\ \ \mathbf{X}\ \ \mathbf{X}^2\ \ \cdots\ \ \mathbf{X}^d\ \ \cdots\ \ \sin(\mathbf{X})\ \ \cdots], \tag{9}$$

where $\mathbf{X}^d$ denotes the matrix containing all possible column vectors obtained from time series of the dth degree polynomials in the state vector $\mathbf{x}$. For example, for a system with two states $\mathbf{x} = [x_1,\ \ x_2]^T$, the quadratic terms are given by the matrix $\mathbf{X}^2 = [x_1^2(\mathbf{t}),\ \ (x_1x_2)(\mathbf{t}),\ \ x_2^2(\mathbf{t})]$, where $\mathbf{t}$ is a vector of times at which the state is measured. Thus, the vector $\mathbf{x}$ is a symbolic *variable*, while the matrix $\mathbf{X}$ is a *data* matrix.

It is now possible to relate the time derivatives in $\dot{\mathbf{X}}$ to the candidate nonlinearities in $\boldsymbol{\Theta}(\mathbf{X})$ by:

$$\dot{\mathbf{X}} = \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\Xi}, \tag{10}$$

where each column $\boldsymbol{\xi}_k$ in $\boldsymbol{\Xi}$ is a vector of coefficients that determines which terms are active in the kth row in Eq. (6). Sparsity promoting algorithms are used to ensure that most of the entries of the column $\boldsymbol{\xi}_k$ are zero. SINDy promotes sparsity by sequential least-squares thresholding, which has recently been shown to converge under suitable conditions [55,56].

By identifying the sparse coefficient vectors $\boldsymbol{\xi}_k$, a model of the nonlinear dynamics may be constructed:

$$\dot{x}_k = \boldsymbol{\Theta}(\mathbf{x})\boldsymbol{\xi}_k, \tag{11}$$

where x_k is the kth element of $\mathbf{x}$ and $\boldsymbol{\Theta}(\mathbf{x})$ refers to a row vector whose elements are symbolic functions of $\mathbf{x}$, as opposed to the data matrix $\boldsymbol{\Theta}(\mathbf{X})$.

Using sparse regression to identify active terms in the dynamics from the candidate library $\boldsymbol{\Theta}(\mathbf{X})$ is a convex optimization. The alternative is to apply a separate constrained regression on every possible subset of nonlinearities and then to choose the model that is both accurate and sparse. This brute-force search is intractable, and the SINDy method makes it possible to select the sparse model in this combinatorially large set of candidate models.

B. Time-varying SINDy

The potential for the SINDy algorithm to discover dominant balance physics has been demonstrated on a diverse set of problems [33–35]. However, the dynamics in these problems are often assumed to not change with time, i.e., they generally have constant coefficients, although the original SINDy algorithm is able to account explicitly for forcing and parameterized dynamics. More recently, SINDy has been extended to deal with parametric partial differential equations [32] by allowing the coefficients $\boldsymbol{\xi}$ of each term in the library to be time dependent. In the present study, we propose a time-varying SINDy to discover intrinsically and strongly time-varying dynamics:

$$\dot{\mathbf{x}} = \boldsymbol{f}_t(\mathbf{x}), \tag{12}$$

where $\boldsymbol{f}_t$ changes with time but is not assumed as an explicitly time-dependent function. Then, the coefficients of the terms identified with SINDy are time varying so that the active terms can vary dramatically with time:

$$\dot{x}_k = \boldsymbol{\Theta}(\mathbf{x})\boldsymbol{\xi}_k(t). \tag{13}$$

We assume $t \in [t - w, t]$ with window size w over which the coefficient vector $\boldsymbol{\xi}_k(t)$ is determined. The basic idea is shown in Fig. 6. We introduce a time sampling window w which moves across the time-series data collected from a time-varying dynamical system [see Fig. 6(a)] and conduct

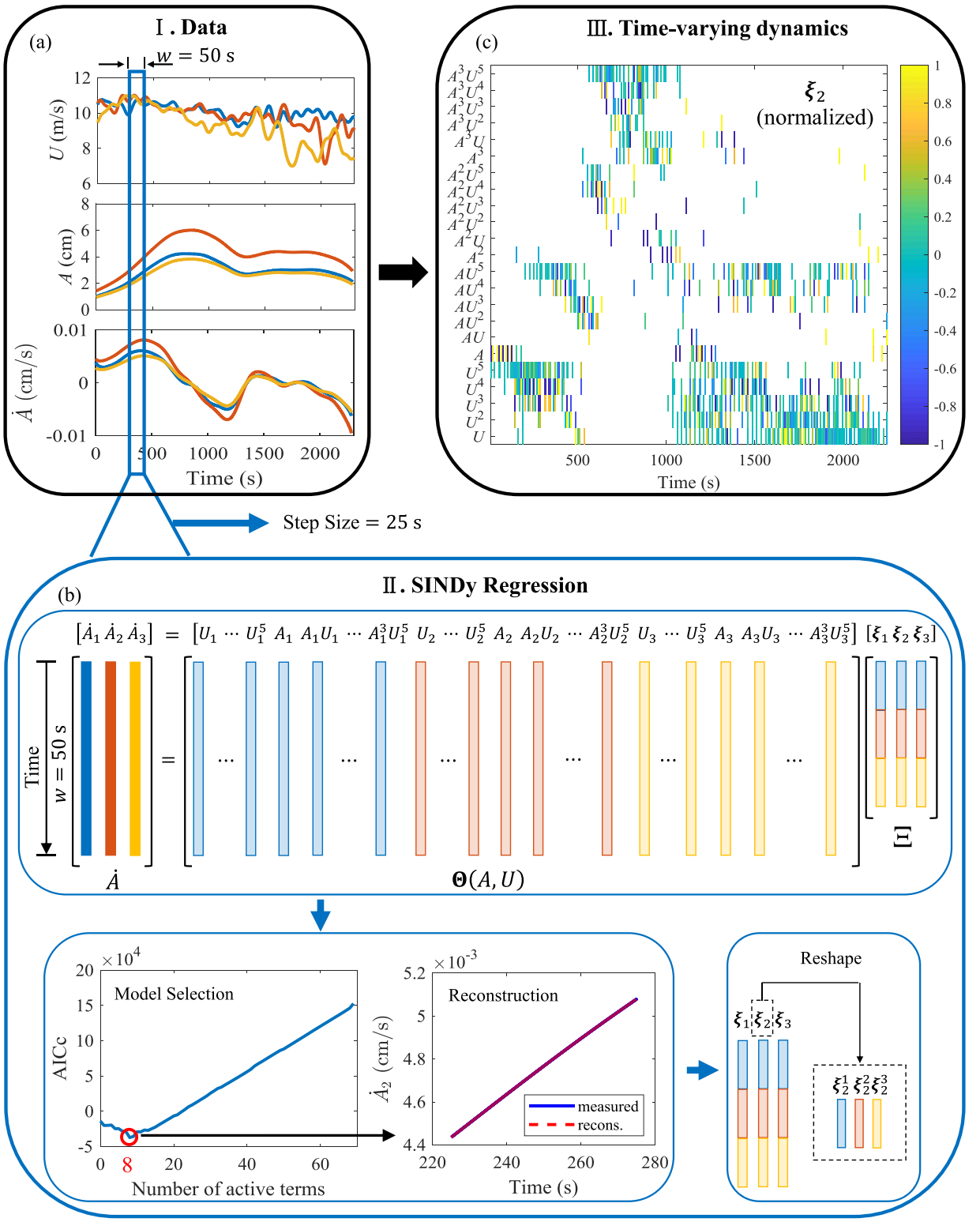

FIG. 6. Schematic of the time-varying SINDy framework, demonstrated on the aerodynamics of a VIV event on a bridge. (a) Data are collected from the measurement system, including a history of time-varying mean wind speed U, amplitudes A, and time derivatives $\dot{A}$. (b) A typical SINDy is conducted in a moving time window at each time instant. The time window is swept across the entire VIV event with a size of 50 s and a moving step size of 25 s. Each component of the obtained model $\boldsymbol{\xi}$ is reshaped into a three-column matrix, where each column corresponds to sensor measurements at one bridge section, respectively, for a more interpretable representation of the obtained time-varying aerodynamics. (c) A time series of the model in terms of $\boldsymbol{\xi}$ is obtained that captures the time-varying aerodynamics of an entire VIV event.

a basic SINDy regression on the data in the window at each time step [see Fig. 6(b)]. We can then sort the obtained active terms and corresponding coefficients in order to reveal the intrinsically time-varying dynamics.

C. SINDy to model time-varying bridge aerodynamics

The time-independent model of VIV described by Eq. (3) only accounts for a simple laboratory experiment where the wind speed is stationary or even constant. This would give time-independent constants for the SINDy parameters. However, the real VIV of a prototype bridge in the field is typically a time-varying, nonlinear dynamical system characterized by the time-varying aerodynamic regime which results from the time variability of natural wind. Equation (3) thus fails in simulating real VIV events. In the present study, we propose a time-varying SINDy model of vibrational displacement amplitude to discover the time-varying bridge aerodynamics from measured VIV events of a long-span bridge.

The input to the time-varying SINDy algorithm consists of time-series data of time-varying mean wind speeds $\mathbf{U}$, vibration displacement amplitudes of the bridge deck $\mathbf{A}$, and the time derivatives $\dot{\mathbf{A}}$ obtained by numerical differentiation for a measured VIV event. Here the wind speed U_k and displacement amplitude A_k, $k = 1, 2, 3$, denote the respective measurement at the kth sensor location along the bridge. In particular, the subscripts 1, 2, and 3 indicate the sensor locations at the bridge sections S1, S2, and S3, respectively. We learn the time-parametrized model over a short-term window with a duration of 50 s, which moves across the VIV event timeline with a step size of 25 s, as shown in Fig. 6(a). A SINDy regression is then performed for data in each 50-s time window, as shown in Fig. 6(b). Although the analytic model in Eq. (5) is unable to describe the time-varying aerodynamics during an entire VIV event, it guides our construction of candidate functions for the library $\boldsymbol{\Theta}$. Specifically, we expand the terms in Eq. (5) and propose a set of polynomial products of the time-varying mean wind speed $\mathbf{U}$ and the vibration displacement amplitude $\mathbf{A}$:

$$\mathbf{A}^i \odot \mathbf{U}^j, \tag{14}$$

where $i = 0, 1, 2, 3$, $j = 0, 1, 2, 3, 4, 5$ are the element-wise power and i and j do not both equal zero. We believe that the higher order of vibration amplitude A in the candidate function implies a higher level of self-excited effects. Note that the Scanlan's model [see Eq. (5)] has well described the highest level of self-excited effects induced by constant wind speed in the VIV wind speed range. The time-varying wind speed in the field which may get out of the VIV wind speed range is only possible to decrease the level of self-excited effects. We thus believe that the highest order of vibration amplitude in the constructed candidate functions should not be more than that in the Scanlan's model [see Eq. (5)], i.e., three.

In addition to the choice of polynomial terms, the single-section model characterizing wind tunnel tests with Eq. (5) is generalized to a higher-dimension variant by incorporating the sensors placed at the S1, S2, and S3 along the bridge span. After computing time derivative data $\dot{\mathbf{A}}$, the proposed SINDy architecture takes the form

$$\dot{\mathbf{A}}(t) = \boldsymbol{\Theta}(\mathbf{A}, \mathbf{U})\, \boldsymbol{\Xi}(t), \tag{15}$$

where the library of candidate functions is defined by

$$\boldsymbol{\Theta}^T(\mathbf{A}, \mathbf{U}) = \begin{bmatrix} \mathbf{A} \\ \mathbf{U} \\ \mathbf{A}^2 \\ \mathbf{A} \odot \mathbf{U} \\ \mathbf{U}^2 \\ \mathbf{A}^2 \odot \mathbf{U} \\ \vdots \\ \mathbf{A}^3 \odot \mathbf{U}^5 \end{bmatrix} \tag{16}$$

and "$\odot$" denotes the element-wise multiplication of $\mathbf{A}$ and $\mathbf{U}$, e.g., $\mathbf{A} \odot \mathbf{U} = [A_1U_1,\ A_2U_2,\ A_3U_3]^T$. Note that the dynamics of A_k at the kth location depend on sensor information at all three locations, i.e., they depend on A_l and U_l with $l = 1, 2, 3$.

The *Akaike information criterion* with a correction for small sample sizes (AICc) [57] is proposed to aid the SINDy regression for model selection. Alternatives, such as BIC and/or the description length approach used by Small *et al.* [41], could also be potentially used in evaluating models. Trajectory reconstructions are then used to ensure the accuracy of the model. For interpretability and visualization, we reshape the obtained models $\boldsymbol{\xi}_k$ into three sets of models $\boldsymbol{\xi}_k^l$, $l = 1, 2, 3$ corresponding to sensors at locations S1, S2, and S3. For example, $\boldsymbol{\xi}_k^1$ is a vector of coefficients of terms $[U_1, \ldots, A_1^3U_1^5]$.

SINDy results in a set of models for a VIV event after the 50-s time window moves through the entire event, as shown in Fig. 6(c). It is found that the active terms and coefficients vary significantly with time. It should be noted that the candidate terms are sorted in ascending polynomial order of vibration amplitude A and wind speed U from bottom to top in Fig. 6(c) and that a higher polynomial order of A implies a stronger wind-structure interaction with a higher level of self-excited (motion-induced) effects. In the same way, we have conducted the proposed time-varying SINDy on 31 measured VIV events in total and report the results for three VIVs in Fig. 7. From the obtained time-varying dynamics for all the VIV events, We can intuitively find four dynamical regimes which are distinguished by the polynomial order of vibration displacement amplitudes $\mathbf{A}$. Accordingly, we rewrite the time-varying SINDy model [see Eq. (15)] for these discovered different dynamical regimes specifically and respectively in Table I. The dynamics at any moment during a VIV event must be from one of or the mix of the discovered regimes.

The effects of window size and moving step size are further studied, as shown in Fig. 8. The time-varying SINDy results with the three different wind sizes and moving step sizes have shown almost identical evolution of active terms and only different time scales. The time-varying SINDy result in this paper is robust to the window size and moving step size because the dynamics of the wind-bridge system changes slowly over time. However, if the dynamics of the studied system changes fast over time, then a self-adaptive window size and moving step size as a function of changing rate of the time-varying dynamics could be a better alternative.

In the VIV wind speed range for the bridge considered in Ref. [58], we find a strong correlation between the time variation of aerodynamics and wind speed (see Fig. 9). Specifically,

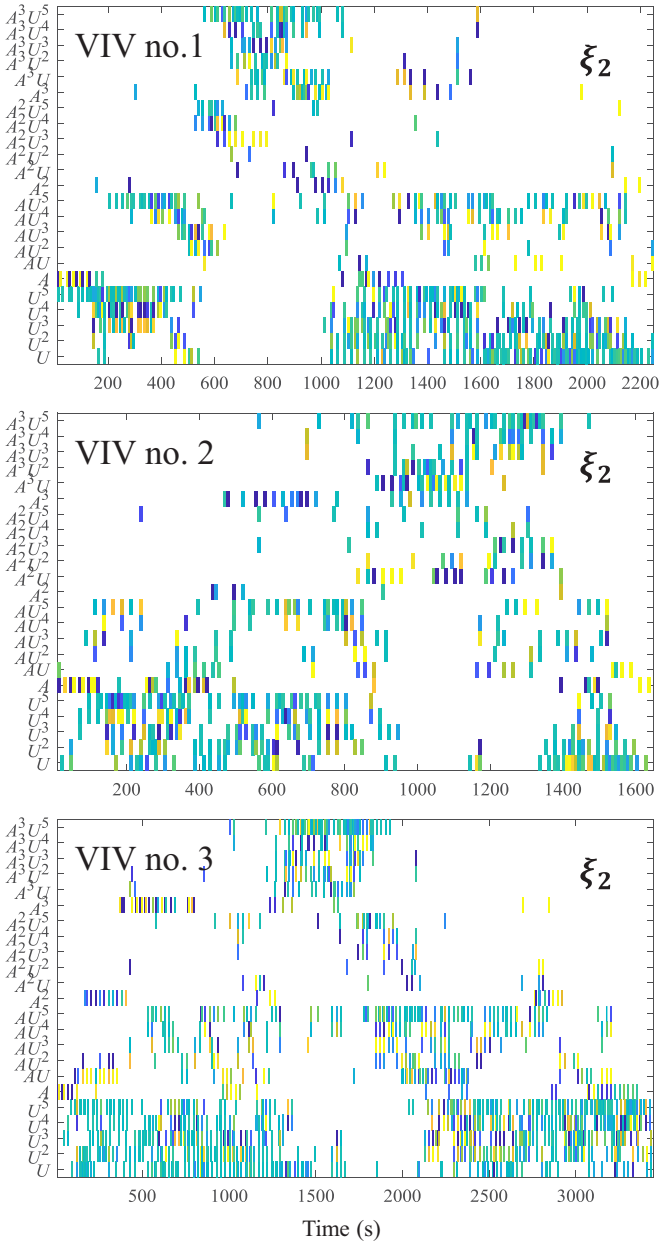

FIG. 7. Time-varying dynamics of three exemplary VIV events discovered by time-varying SINDy.

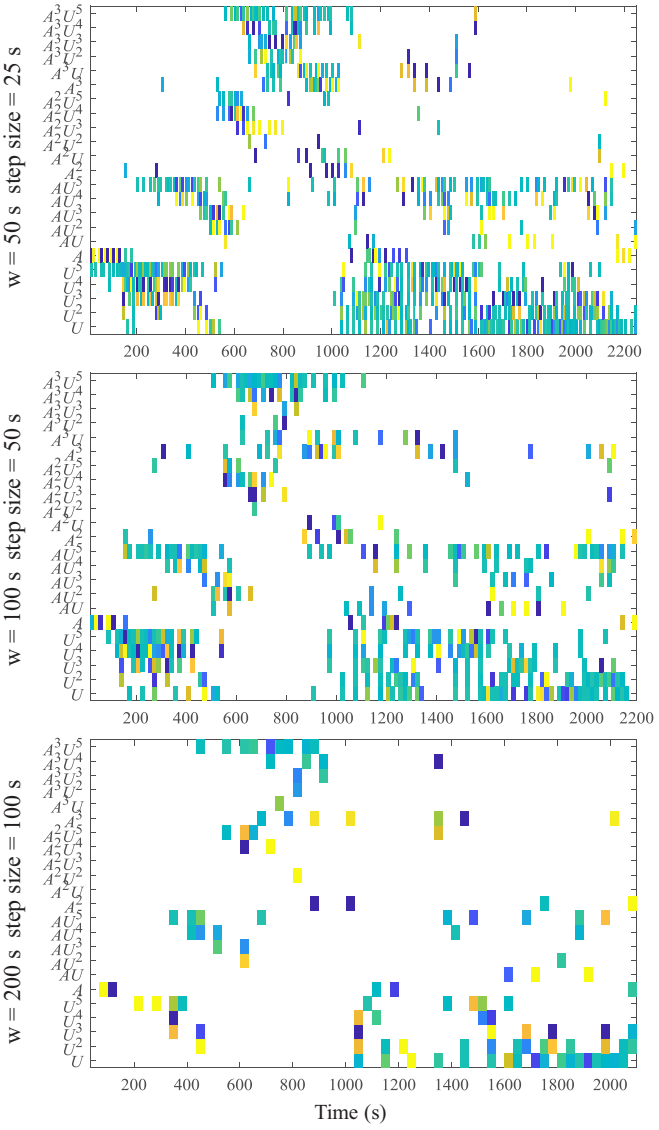

FIG. 8. The effects of window size and moving step size on the time-varying SINDy results for a VIV event.

during the first stage (0–600 s), the wind speeds at S1, S2, and S3 all stay within the VIV win speed range, resulting in full development of wind-structure interaction with an increasing motion-induced (self-excited) effect. This is indicated by the increasing polynomial order of A in active terms with time. During the second stage (600–1000 s), the aerodynamic system reaches the steady state of high wind-structure interaction with the strong motion-induced (self-excited) effects. Here only the terms with the highest polynomial order in A are active. During the third stage (1000 s to the end), wind speeds at S2 and S3 fall out of the VIV wind speed range, resulting in

TABLE I. The obtained SINDy models for the discovered different dynamical regimes which are distinguished by the polynomial order of vibration displacement amplitude $\mathbf{A}$ in the active terms accounting for the level of self-excited effect in the bridge-wind interaction. $\boldsymbol{\Xi}_{R1}(t)$, $\boldsymbol{\Xi}_{R2}(t)$, $\boldsymbol{\Xi}_{R3}(t)$, and $\boldsymbol{\Xi}_{R4}(t)$ are the corresponding subsets of $\boldsymbol{\Xi}(t)$, respectively.

Dynamical regime	SINDy model	Characteristics
Regime 1	$\dot{\mathbf{A}}(t) = [\mathbf{U}, \mathbf{U}^2, \mathbf{U}^3, \mathbf{U}^4, \mathbf{U}^5, \mathbf{A}]\boldsymbol{\Xi}_{R1}(t)$	No self-excited effect
Regime 2	$\dot{\mathbf{A}}(t) = [\mathbf{A}\odot\mathbf{U}, \mathbf{A}\odot\mathbf{U}^2, \mathbf{A}\odot\mathbf{U}^3, \mathbf{A}\odot\mathbf{U}^4, \mathbf{A}\odot\mathbf{U}^5, \mathbf{A}^2]\boldsymbol{\Xi}_{R2}(t)$	Slight self-excited effect
Regime 3	$\dot{\mathbf{A}}(t) = [\mathbf{A}^2\odot\mathbf{U}, \mathbf{A}^2\odot\mathbf{U}^2, \mathbf{A}^2\odot\mathbf{U}^3, \mathbf{A}^2\odot\mathbf{U}^4, \mathbf{A}^2\odot\mathbf{U}^5, \mathbf{A}^3]\boldsymbol{\Xi}_{R3}(t)$	Medium self-excited effect
Regime 4	$\dot{\mathbf{A}}(t) = [\mathbf{A}^3\odot\mathbf{U}, \mathbf{A}^3\odot\mathbf{U}^2, \mathbf{A}^3\odot\mathbf{U}^3, \mathbf{A}^3\odot\mathbf{U}^4, \mathbf{A}^3\odot\mathbf{U}^5]\boldsymbol{\Xi}_{R4}(t)$	Strong self-excited effect

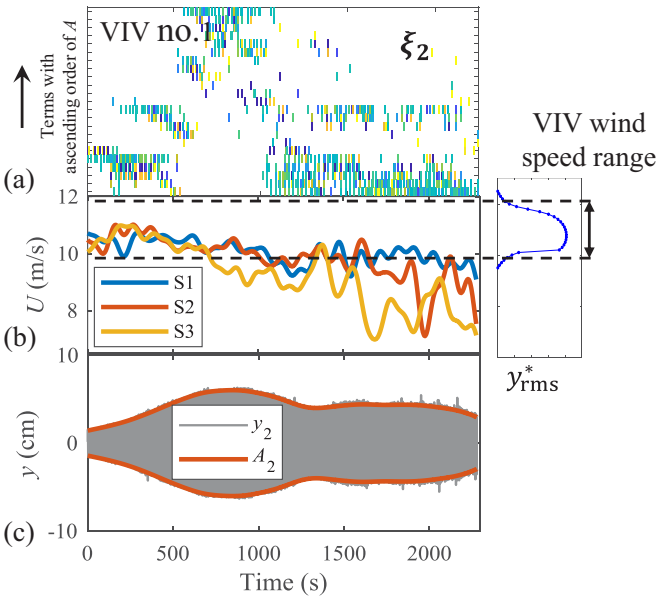

FIG. 9. Interpretation of time-varying aerodynamics for a VIV event and the VIV wind speed range obtained in Ref. [52]. (a) The time series of model sets $\boldsymbol{\xi}_2$. (b) The history of time-varying mean wind speeds U compared with the VIV wind speed. (c) The history of displacement y_2 with the amplitude A_2.

a significant decrease of motion-induced (self-excited) effects. This is indicated by the decreasing polynomial order of A in the active terms, i.e., the system becomes weak coupled. The obtained time-dependent, nonlinear dynamics is capable of producing a parsimonious model of the aerodynamics of a real bridge VIV event.

D. Simulation of measured VIVs by the obtained time-dependent SINDy models

We have obtained a specific parametric model for each measured VIV event by the proposed time-varying SINDy. Each VIV is thus represented by an ODE with the corresponding time-dependent parameter $\boldsymbol{\Xi}(t)$ [see Eq. (15)]. To validate the obtained models, we simulate all entire VIV events by numerically solving the parametric models with the corresponding time-dependent parameters $\boldsymbol{\Xi}(t)$ given the measured initial states $\mathbf{A}(t=0)$ and the measured wind histories $\mathbf{U}(t)$. The normalized mean-square error (NMSE) is calculated to evaluate the prediction performance. The comparisons between the simulated and measured states for three VIV events as examples show a near perfect agreement with an averaged NMSE of 0.0023 (see Fig. 10), indicating the high accuracies of the obtained models.

V. DISTINGUISHED DYNAMICAL REGIMES: CLUSTERING OF DYNAMIC MODELS

The analysis of the identified time-varying aerodynamic response from the VIV events (see Fig. 9 for a single VIV event) indicates the existence of several distinct dynamical regimes, all of which contribute to revealing the underlying, time-varying aerodynamic physics. The patterns associated with different SINDy model structures, e.g., as shown in Table I, indicate distinct dynamical regimes. This motivates the application of cluster analysis on the model sets to automatically discover the potential modes of aerodynamic behavior in the VIVs.

A. Clustering algorithm

In the clustering algorithm [59] applied in this study, two quantities are calculated for each data point i: the local density

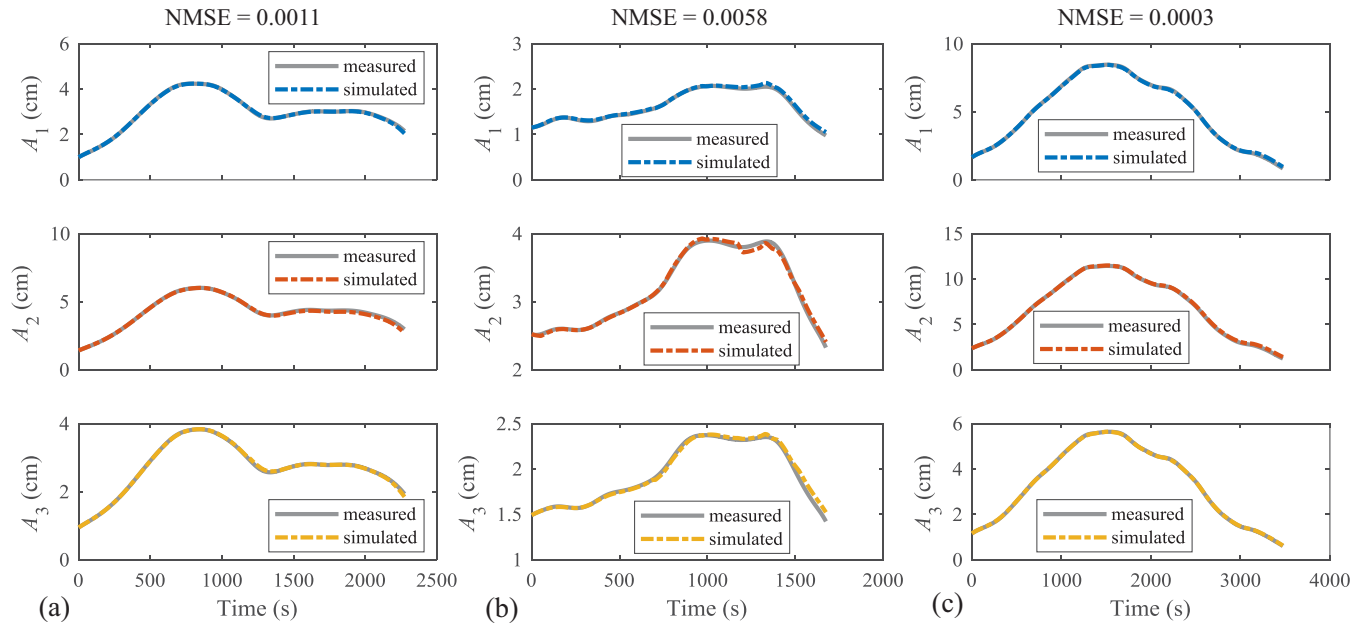

FIG. 10. Simulations for three entire VIV events as examples by solving the ODE with the corresponding obtained time-dependent parameter $\boldsymbol{\Xi}(t)$ [see Eq. (15)] with only the measured initial state $\mathbf{A}(t=0)$ and the measured wind history $\mathbf{U}(t)$ given. (a) VIV event No. 1 with NMSE of 0.0011. (b) VIV event No. 2 with NMSE of 0.0058. (3) VIV event No. 3 with NMSE of 0.0003.

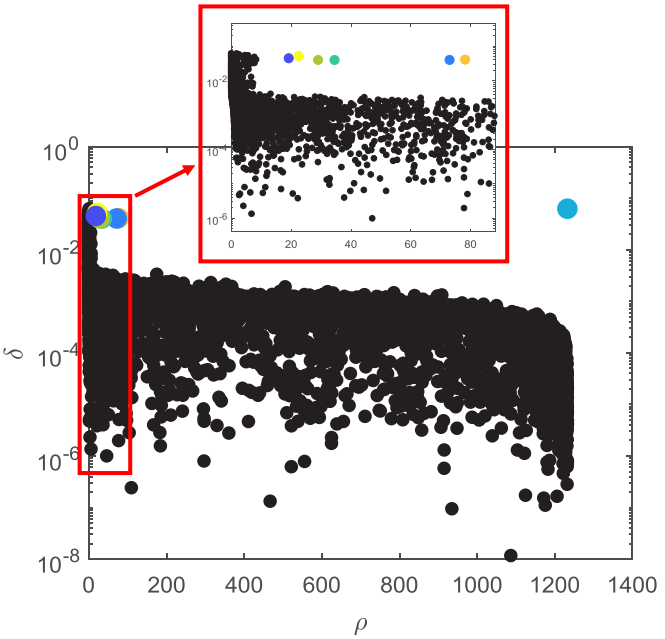

FIG. 11. Decision graph: Seven cluster centers (colored) are determined by points for which the value of δ is anomalously large.

ρ_i and the distance δ_i. The local density of data point i is defined as

$$\rho_i = \sum_j e^{-\frac{d_{ij}^2}{d_c^2}}, \tag{17}$$

where d_{ij} is the Euclidean distance between data point i and j, d_c is a cutoff distance, and $e^{-\frac{d_{ij}^2}{d_c^2}}$ is actually the Gaussian radial distance. The quantity ρ_i thus measures the local density of data point i within the radial radius d_c. The distance δ_i is defined as the minimum distance between the point i and any other point with a higher density:

$$\delta_i = \min_{j:\rho_j>\rho_i} (d_{ij}). \tag{18}$$

But for the point with the highest global density, the distance δ_i is defined as the maximum distance between data point i and any other point as there is no data point with a higher density.

By plotting all the data points with the two quantities defined by Eqs. (17) and (18), the cluster centers are recognized fast and easily as points for which the value of δ_i is anomalously large without a definite pre-specified number of clusters.

After the identification of cluster centers, each remaining point is assigned to the same cluster as its nearest neighbor of higher density. It is noted that this algorithm is sensitive only to the relative magnitude of ρ for different points and the clustering results are robust against the parameter d_c for large data sets [59].

B. Cluster analysis of the obtained dynamic models

We consider the time series of the model coefficient vector $\boldsymbol{\xi}_2$ obtained by the time-varying SINDy algorithm for each of the measured 31 VIV events. Each model set (consisting of ξ_2^1, ξ_2^2, and ξ_2^3) is considered as a data point in the 23-dimensional model space for this cluster analysis, where each dimension corresponds to a term in the candidate function library. With ploting all the model sets with the two quantities defined by Eqs. (17) and (18), seven cluster centers are identified as points for which the value of δ_i is anomalously large (see Fig. 11). And the corresponding clusters are obtained after the assignment of each remaining model set to the same cluster as its nearest neighbor of higher density.

The obtained clusters along with their members are shown in Fig. 12. It can be found that the model sets in the same cluster have common dominant terms. Specifically, the common dominant terms in C1 are U, U^2, U^3, U^4, and U^5, indicating purely forced vibrations by wind. C2 and C3 are dominated by the same term A, however with different signs. Thus, these clusters represent linear dynamics with respect to A. The most dominant term in C4 is A^2, followed by A and A^3. C5 and C6 have the same dominant term A^3, which corresponds to the parameter α in Eq. (5) related to the aerodynamic parameter Y_1

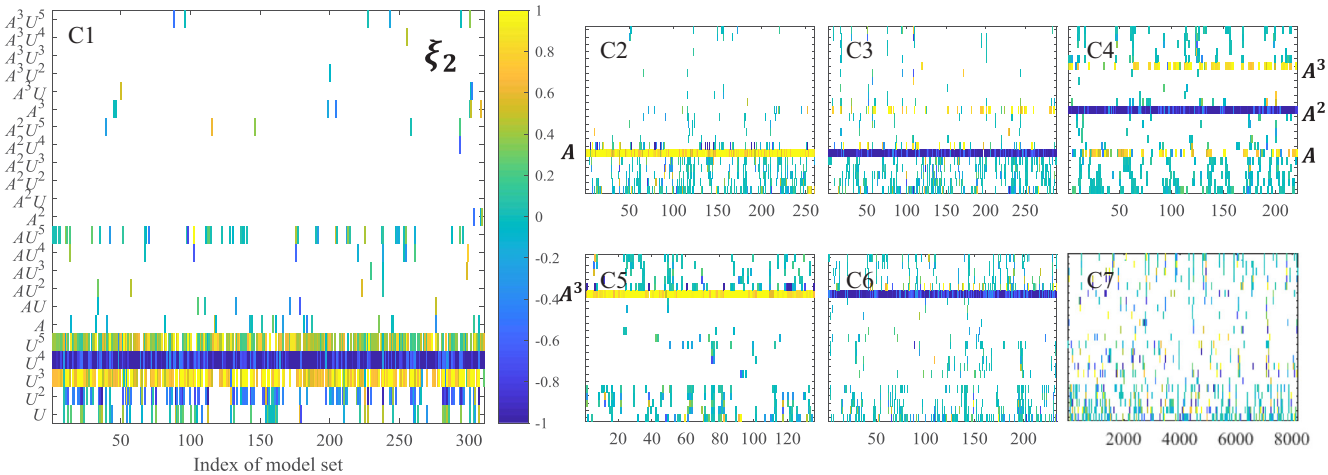

FIG. 12. The obtained seven clusters in the model sets. The model sets in the same cluster have common donimant terms except Cluster 7 (C7). These clusters are distinguished by the polynomial order of the vibration amplitude A in the dominant terms.

in Eq. (2). We can thus know that the discovered terms with A^3 actually correspond to the aerodynamic damping component of the motion induced force in the Simiu and Scanlan's model [see Eq. (2)]. In C7 no term is dominant, but instead the dynamics are mixed where both wind-induced force and self-excited force account for the vibration of the bridge. It can be found that these clusters are distinguished by the polynomial order of the vibration amplitude A in the dominant terms, which is just consistent with the intuitively discovered four dynamical regimes shown in Table I, indicating that different dynamical regimes in VIV aerodynamics of this bridge are distinguished by the level of self-excited effects in the wind-structure interaction. As analyzed with Fig. 9, the temporal dynamical regime of the bridge-wind system is intrinsically determined by the temporal wind condition and bridge state.

VI. CONCLUSIONS

In the present work, we have developed a data-driven method to discover time-varying aerodynamics of a long-span bridge during VIV events based on field measurements. Using the SINDy algorithm, we are able to identify parsimonious, time-varying dynamical systems which result from VIV events of the bridge subjected to nonstationary wind characterized by time-varying mean wind speed. Thus we are able to posit new, data-driven models highlighting the time-varying aerodynamics of the wind-bridge system during VIV events subjected to time-varying wind, which may get out of VIV wind speed range for a while.

The wind-bridge aerodynamical system is shown to have distinct, time-dependent modes of behavior, thus requiring parametric models to account for the diversity of dynamics. The obtained time-varying SINDy models have visualized and revealed the evolution of bridge aerodynamics over time during VIV events. The variation of aerodynamics is mainly reflected in the level of self-excited effects, which is intrinsically determined by temporal wind condition and bridge motion state. Clustering of obtained models has discovered potential modes of bridge aerodynamics during VIV events and clearly show distinct dynamical regimes of the wind-bridge system that are distinguished by the level of self-excited effects. Simulation of VIV displacement amplitude history by the obtained time-varying SINDy model which is actually a time-dependent ODE has shown high accuracies of the model. All the above indicate that time-varying SINDy architecture and clustering analysis are effective in identifying parsimonious, time-varying aerodynamical systems which result from VIV events of the bridge.

ACKNOWLEDGMENTS

S. Li acknowledges funding support from the China Scholarship Council (Grant No. 201706120256). S. Laima acknowledges funding support from the National Natural Science Foundation of China (Grant No. 51503138). H.L. acknowledges funding support from the National Natural Science Foundation of China (Grant No. 51638007). S.L.B. acknowledges funding support from the Army Research Office (ARO W911NF-17-1-0422) and Air Force Office of Scientific Research (AFOSR FA9550-18-1-0200). E.K. also gratefully acknowledges support by the Washington Research Foundation, the Gordon and Betty Moore Foundation (Award No. 2013-10-29), and the Alfred P. Sloan Foundation (Award No. 3835). J.N.K. acknowledges support from the Air Force Office of Scientific Research (AFOSR) Grant No. FA9550-17-1-0329.

[1] S. L. Brunton, J. L. Proctor, and J. Nathan Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016).

[2] A. G. Davenport, The response of slender, line-like structures to a gusty wind, Proc. Inst. Civil Eng. **23**, 389 (1962).

[3] R. H. Scanlan, The action of flexible bridges under wind, ii: Buffeting theory, J. Sound Vib. **60**, 201 (1978).

[4] Y. K. Lin, Motion of suspension bridges in turbulent winds, J. Eng. Mech. ASCE, **105**, 921 (1979).

[5] Y. K. Lin and J. N. Yang, Multimode bridge response to wind excitations, J. Eng. Mech. **109**, 586 (1983).

[6] R. E. D. Bishop and A. Y. Hassan, The lift and drag forces on a circular cylinder oscillating in a flowing fluid, Proc. R. Soc. Lond. A **277**, 51 (1964).

[7] R. T. Hartlen and I. G. Currie, Lift-oscillator model of vortex-induced vibration, J. Eng. Mech. Div. **96**, 577 (1970).

[8] W. D. Iwan and R. D. Blevins, A model for vortex induced oscillation of structures, J. Appl. Mech. **41**, 581 (1974).

[9] E. Simiu and R. H. Scanlan, *Wind Effects on Structures: An Introduction to Wind Engineering* (Wiley, New York, 1978).

[10] T. Theodorsen, General theory of aerodynamic instability and the mechanism of flutter, NACA Report 496 (1935).

[11] F. Bleich and Advisory Board on the Investigation of Suspension Bridges, *The mathematical Theory of Vibration in Suspension Bridges: A Contribution to the Work of the Advisory Board on the Investigation of Suspension Bridges* (Department of Commerce, Bureau of public roads, Washington, 1950).

[12] Y. Nakamura and T. Mizota, Unsteady lifts and wakes of oscillating rectangular prisms, J. Eng. Mech. Div. **101**, 855 (1975).

[13] S. Komatsu and H. Kobayashi, Vortex-induced oscillation of bluff cylinders, J. Wind Eng. Industr. Aerodynam. **6**, 335 (1980).

[14] M. Matsumoto, N. Shiraishi, H. Shirato, S. Stoyanoff, and T. Yagi, Mechanism of, and turbulence effect on vortex-induced oscillations for bridge box girders, J. Wind Eng. Industr. Aerodynam. **49**, 467 (1993).

[15] S. Laima, H. Li, W. Chen, and F. Li, Investigation and control of vortex-induced vibration of twin box girders, J. Fluids Struct. **39**, 205 (2013).

[16] H. Li, S. Laima, and H. Jing, Reynolds number effects on aerodynamic characteristics and vortex-induced vibration of a twin-box girder, J. Fluids Struct. **50**, 358 (2014).

[17] R. H. Scanlan, State-of-the-art methods for calculating flutter, vortex-induced, and buffeting response of bridge structures, Technical report (1981).
[18] A. G. Davenport, J. P. C. King, and G. L. Larose, Taut strip model tests, in *Proceedings of the 1st International Symposium on Aerodynamics of Large Bridges* (Copenhagen, Denmark, 1992), pp. 113–124.
[19] R. H. Scanlan, Problematics in formulation of wind-force models for bridge decks, J. Eng. Mech. **119**, 1353 (1993).
[20] P. P. Sarkar, N. P. Jones, and R. H. Scanlan, Identification of aeroelastic parameters of flexible bridges, J. Eng. Mech. **120**, 1718 (1994).
[21] G. L. Larose and J. Mann, Gust loading on streamlined bridge decks, J. Fluids Struct. **12**, 511 (1998).
[22] X. Chen and A. Kareem, Advances in modeling of aerodynamic forces on bridge decks, J. Eng. Mech. **128**, 1193 (2002).
[23] X. Chen, M. Matsumoto, and A. Kareem, Time domain flutter and buffeting response analysis of bridges, J. Eng. Mech. **126**, 7 (2000).
[24] X. Chen and A. Kareem, Nonlinear response analysis of long-span bridges under turbulent winds, J. Wind Eng. Industr. Aerodynam. **89**, 1335 (2001).
[25] G. Diana, F. Resta, and D. Rocchi, A new numerical approach to reproduce bridge aerodynamic non-linearities in time domain, J. Wind Eng. Industr. Aerodynam. **96**, 1871 (2008).
[26] G. Diana, D. Rocchi, T. Argentini, and S. Muggiasca, Aerodynamic instability of a bridge deck section model: Linear and nonlinear approach to force modeling, J. Wind Eng. Industr. Aerodynam. **98**, 363 (2010).
[27] A. Larsen, Computer simulation of wind-structure interaction in bridge aerodynamics, Struct. Eng. Int. **8**, 105 (1998).
[28] Y. J. Ge, Z. X. Lin, F. C. Cao, J. B. Pang, and H. F. Xiang, Investigation and prevention of deck galloping oscillation with computational and experimental techniques, J. Wind Eng. Industr. Aerodynam. **90**, 2087 (2002).
[29] Y. Bai, D. Sun, and J. Lin, Three dimensional numerical simulations of long-span bridge aerodynamics using block-iterative coupling and DES, Comput. Fluids **39**, 1549 (2010).
[30] B. de Silva, D. M. Higdon, S. L. Brunton, and J. N. Kutz, Discovery of physics from data: Universal laws and discrepancy models, arXiv:1906.07906.
[31] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Data-driven discovery of partial differential equations, Sci. Adv. **3**, e1602614 (2017).
[32] S. Rudy, A. Alla, S. L. Brunton, and J. N. Kutz, Data-driven identification of parametric partial differential equations, SIAM J. Appl. Dyn. Syst. **18**, 643 (2019).
[33] N. M. Mangan, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Inferring biological networks by sparse identification of nonlinear dynamics, IEEE Trans. Molec. Biol. Multi-Scale Commun. **2**, 52 (2016).
[34] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Sparse identification of nonlinear dynamics with control (sindyc), IFAC-PapersOnLine **49**, 710 (2016).
[35] K. Champion, S. L. Brunton, and J. N. Kutz, Discovery of nonlinear multiscale systems: Sampling strategies and embeddings, SIAM J. Appl. Dyn. Syst. **18**, 312 (2019).
[36] K. Champion, P. Zheng, A. Y. Aravkin, S. L. Brunton, and J. N. Kutz, A unified sparse optimization framework to learn parsimonious physics-informed models from data, arXiv:1906.10612.
[37] J. Bongard and H. Lipson, Automated reverse engineering of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA **104**, 9943 (2007).
[38] M. Schmidt and H. Lipson, Distilling free-form natural laws from experimental data, Science **324**, 81 (2009).
[39] T. Cornforth and H. Lipson, Symbolic regression of multiple-time-scale dynamical systems, In *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation* (ACM, New York, 2012), pp 735–742.
[40] C. Yao and E. M. Bollt, Modeling and nonlinear parameter estimation with Kronecker product representation for coupled oscillators and spatiotemporal systems, Physica D **227**, 78 (2007).
[41] M. Small, K. Judd, and A. Mees, Modeling continuous processes from data, Phys. Rev. E **65**, 046704 (2002).
[42] Q. Li, F. Dietrich, E. M. Bollt, and I. G. Kevrekidis, Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the Koopman operator, Chaos **27**, 103111 (2017).
[43] P. R. Vlachas, W. Byeon, Z. Y. Wan, T. P. Sapsis, and P. Koumoutsakos, Data-driven forecasting of high-dimensional chaotic systems with long-short term memory networks, Proc. R. Soc. A **474**, 20170844 (2018).
[44] E. Yeung, S. Kundu, and N. Hodas, Learning deep neural network representations for Koopman operators of nonlinear dynamical systems, arXiv:1708.06850.
[45] N. Takeishi, Y. Kawahara, and T. Yairi, Learning Koopman invariant subspaces for dynamic mode decomposition, in *Advances in Neural Information Processing Systems* (Long Beach, California, 2017), pp. 1130–1140.
[46] C. Wehmeyer and F. Noé, Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics, J. Chem. Phys. **148**, 241703 (2018).
[47] A. Mardt, L. Pasquali, H. Wu, and F. Noé, VAMPnets for deep learning of molecular kinetics, Nat. Commun. **9**, 5 (2018).
[48] B. Lusch, J. N. Kutz, and S. L. Brunton, Deep learning for universal linear embeddings of nonlinear dynamics, Nat. Commun. **9**, 4950 (2018).
[49] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Multistep neural networks for data-driven discovery of nonlinear dynamical systems, arXiv:1801.01236.
[50] B. J. Vickery and R. I. Basu, Across-wind vibrations of structures of circular cross-section. part i. development of a mathematical model for two-dimensional conditions, J. Wind Eng. Industr. Aerodynam. **12**, 49 (1983).
[51] F. Ehsan and R. H. Scanlan, Vortex-induced vibrations of flexible bridges, J. Eng. Mech. **116**, 1392 (1990).
[52] S. Li, S. Laima, and H. Li, Cluster analysis of winds and wind-induced vibrations on a long-span bridge based on long-term field monitoring data, Eng. Struct. **138**, 245 (2017).
[53] E. Kaiser, J. N. Kutz, and S. L. Brunton, Sparse identification of nonlinear dynamics for model predictive control in the low-data limit, Proc. R. Soc. A **474**, 20180335 (2018).
[54] R. Chartrand, Numerical differentiation of noisy, nonsmooth data, ISRN Appl. Math. **2011**, 164564 (2011).
[55] P. Zheng, T. Askham, S. L. Brunton, J. N. Kutz, and A. Y. Aravkin, A Unified Framework for Sparse Relaxed Regularized Regression: SR3, IEEE Access **7**, 1404 (2018).

[56] L. Zhang and H. Schaeffer, On the convergence of the SINDy algorithm, SIAM Multiscale Model. Simul. **17**, 948 (2019).

[57] J. E. Cavanaugh, Unifying the derivations for the Akaike and corrected akaike information criteria, Stat. Probab. Lett. **33**, 201 (1997).

[58] S. Li, S. Laima, and H. Li, Data-driven modeling of vortex-induced vibration of a long-span suspension bridge using decision tree learning and support vector regression, J. Wind Eng. Industr. Aerodynam. **172**, 196 (2018).

[59] A. Rodriguez and A. Laio, Clustering by fast search and find of density peaks, Science **344**, 1492 (2014).