

# Estimating Transient Stability Regions of Large-scale Power Systems Part I: Koopman Operator and Reduced-order Model

Yuqing Lin, Tianhao Wen, Lei Chen, Q. H. Wu, *Life Fellow, IEEE, Fellow, CSEE*,  
and Yang Liu, *Member, IEEE, Member, CSEE*

**Abstract**—This paper presents an estimation of transient stability regions for large-scale power systems. In Part I, a Koopman operator based model reduction (KOMR) method is proposed to derive a low-order dynamical model with reasonable accuracy for transient stability analysis of large-scale power systems. Unlike traditional reduction methods based on linearized models, the proposed method does not require linearization, but captures dominant modes of the original nonlinear dynamics by employing a Koopman operator defined in an infinite-dimensional observable space. Combined with the Galerkin projection, the obtained dominant Koopman eigenvalues and modes produce a reduced-order nonlinear model. To approximate the Koopman operator with sufficient accuracy, we introduce a Polynomial-based Multi-trajectory Kernel Dynamic Mode Decomposition (PMK-DMD) algorithm, which outperforms traditional DMD in various scenarios. In the end, the proposed method is applied to the IEEE 10-machine-39-bus power system and IEEE 16-machine-68-bus power system, which demonstrates that our method is significantly superior to the modal analysis method in both qualitative and quantitative aspects.

**Index Terms**—Data driven method, dynamic mode decomposition, Koopman operator, model reduction, power systems.

## I. INTRODUCTION

**T**RANSIENT stability analysis investigates the dynamic behavior of power systems after large disturbances, which is of great significance in power system planning, operation, and control [1], [2]. Stability region (SR), mathematically known as *domain of attraction*, provides a quantitative description of transient stability [3]. It is defined as an invariant set such that all trajectories from points in the set converge to a corresponding asymptotically stable equilibrium point [4]. However, calculating an entire SR is a challenging task. Especially for a large-scale power system, no one has yet been able to realize the derivation of an entire SR, because the dynamical model of the power system is high-dimensional and complex [2]. Model reduction is currently the most promising

solution [5]. In this paper, we will present a work on estimation of SRs for large-scale power systems based on reduced-order models. Our work is presented in two parts. Part I introduces how to capture the nonlinear dynamics of a power system using the Koopman operator and derive a low-order dynamical model with reasonable accuracy. Part II describes how to use the reduced-order model combined with sum-of-squares programming to approximate a reduced-order stability region and map it to the full-order state space. We will focus on the first part of our work below.

A wide variety of model reduction tools for large-scale power system are developed by the following concepts. A well-known concept is *coherency*, which describes the similarity of generator responses following given disturbances to derive a reduced nonlinear external system that retains the relevant dynamics [5]. For identifying coherent groups of machines, several approaches have been proposed, including time simulation [6], modal coherency [7], slow coherency [8] and weak-link methods [9]. However, the quality of the obtained equivalent reduced-order model may be unsatisfactory, relying on the disturbances [7] and the system conditions [10] chosen to determine coherency. Another important concept is *modal analysis*, which computes and analyses the eigenvalues of the Jacobi matrix defined on an equilibrium point of the linearized dynamical system [5], [11], [12]. Modal analysis provides a good insight into various oscillatory modes present in the system and does not depend on any disturbances. However, dominant modals may be difficult to extract for constructing a reduced-order model and may not reflect the oscillatory behavior under large perturbations [5]. Additionally, there are concepts used to derive reduced-order models approximating the input-output behavior of external systems, such as balanced realization [13], selective modal analysis [14], and singular perturbations [15], among others. Since SRs are generally defined on autonomous models rather than input-output models, these concepts are not applicable to derive a model for studying SRs.

The modal analysis only discusses modes derived from linearized models that do not reflect nonlinear dynamics away from the equilibrium point under large perturbations [5]. Such a drawback can be overcome by introducing the so-called Koopman operator and studying its spectrum [16]–[18]. The Koopman operator is an infinite-dimensional linear operator that describes how measurements (observable functions or

Manuscript received February 18, 2024; revised May 23, 2024; accepted June 22, 2024. Date of online publication January 10, 2025; date of current version January 15, 2025.

Y. Q. Lin, T. H. Wen, L. Chen, Q. H. Wu, and Y. Liu (corresponding author, email: epyangliu@scut.edu.cn) are with the School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China.

DOI: 10.17775/CSEEJPES.2024.01170

observables) of a finite-dimensional dynamical system evolve through its nonlinear dynamics [17]. Indeed, it does not rely on linearization and captures the full information of the nonlinear dynamical system. Thus, accurately approximating the Koopman operator and employing spectral decomposition technique allow us to capture the dominant modes of the power system following large perturbations and realize model reduction. Currently, it is common to estimate Koopman operators using data-driven approaches such as Dynamic Mode Decomposition (DMD) [18], Extended DMD (EDMD) [19], [20], kernel DMD [21], etc. Among of them, DMD selects linear functions of the state as the observable functions, forming a subspace that is not “rich” enough for effective approximation [18]. EDMD and kernel DMD construct a “sufficiently rich” subspace spanned by a great deal of nonlinear observables and produce a more accurate approximation. However, finding the right set of observables is nontrivial and an unsolved problem, both for EDMD and kernel DMD [22]. Moreover, EDMD may face heavy computational burdens, because the number of the needed observables grows rapidly as the dimension of the state space increases [17]. More details and applications of the Koopman operator can be found in [23]–[26].

The contribution of Part I in this paper is threefold.

a) We propose a Polynomialization based Multi-trajectory Kernel Dynamic Mode Decomposition (PMK-DMD) algorithm to obtain a reliable approximation of the Koopman operator. PMK-DMD analytically constructs a significantly large space of observables and is computationally tractable, by applying polynomialization for the original system and utilizing the so called kernel trick respectively. Thus, PMK-DMD can address the challenges faced by EDMD and kernel DMD. Furthermore, PMK-DMD employs multiple trajectories to capture the dynamics within the global state space, minimizing the reliance on individual trajectories.

b) The estimated Koopman operator combined with the Galerkin projection method [27] is utilized to derive a low-dimensional and fairly accurate model reflecting nonlinear dynamics of the original high-dimensional system. In contrast, traditional methods either obtain linear models (Modal analysis) or have low accuracy (Coherency analysis). Moreover, we introduce a quantitative metric for assessing the quality of the reduced-order model.

c) The proposed Koopman operator based model reduction (KOMR) method is applied to the IEEE 10-machine-39-bus power system case and IEEE 16-machine-68-bus power system case. In the former case, we compare the effectiveness of the proposed method with that of the traditional modal analysis in terms of order reduction. Moreover, we also compare the performance of PMK-DMD and DMD in various scenarios, and discuss the computational savings due to model reduction.

In Part I of this paper, for the first time, the Koopman operator theory is successfully applied to the system-level, not device-level, model reduction of a power system. The obtained reduced-order model captures the dominant nonlinear dynamics of the system following large disturbances, which is conducive to transient stability analysis.

**Notation:** The set of  $n \times m$  real and complex matrices are represented by  $\mathbb{R}^{n \times m}$  and  $\mathbb{C}^{n \times m}$ , respectively.  $[\mu_1, \mu_2,$

$\dots, \mu_c]_{\text{diag}}$  denotes a diagonal matrix constructed by vector  $[\mu_1, \mu_2, \dots, \mu_c]$ .  $\mathbf{T}^+$  is the Moore-Penrose pseudoinverse of matrix  $\mathbf{T}$ .

## II. PRELIMINARIES

### A. Koopman Operator

The *Koopman operator*, proposed by B.O. Koopman in 1931 [16], is an infinite-dimensional linear operator that describes how measurements (observable functions) of a finite-dimensional dynamical system evolve through its nonlinear dynamics. Consider a continuous-time dynamical system

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) \quad (1)$$

where  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\mathbf{x} = [x_1, x_2, \dots, x_n]^\top$ . The continuous-time Koopman operator  $\mathcal{K}_c$  is an infinite-dimensional linear operator that acts on all observable function vectors  $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^l$  [17]. Define a finite time- $g$  flow  $\mathbf{S}_g : \mathbf{x}(t) \rightarrow \mathbf{x}(t+g) = \mathbf{x}(t) + \int_t^{t+g} \mathbf{f}(\mathbf{x}(\tau))d\tau$ . Then we have

$$\mathcal{K}_c \mathbf{h}(\mathbf{x}(t)) = (\mathbf{h} \circ \mathbf{S}_g)(\mathbf{x}(t)) = \mathbf{h}(\mathbf{S}_g(\mathbf{x}(t))) \quad (2)$$

where  $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_l(\mathbf{x})]^\top$  and  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  ( $i = 1, 2, \dots, l$ ) is a scalar-valued observable function. All observables (observable functions) form a Hilbert space.

System (1) will induce a discrete-time dynamical system given by the flow map  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  which maps the state  $\mathbf{x}_k \triangleq \mathbf{x}(k\Delta t)$  to a future time  $\mathbf{x}_{k+1}$ , where  $k = 1, 2, \dots$

$$\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k) = \mathbf{x}_k + \int_{k\Delta t}^{(k+1)\Delta t} \mathbf{f}(\mathbf{x}(\tau))d\tau \quad (3)$$

The discrete-time Koopman operator is given by  $\mathcal{K}_d$  such that

$$\mathcal{K}_d \mathbf{h}(\mathbf{x}_k) = (\mathbf{h} \circ \mathbf{F})(\mathbf{x}_k) = \mathbf{h}(\mathbf{F}(\mathbf{x}_k)) = \mathbf{h}(\mathbf{x}_{k+1}) \quad (4)$$

We consider the spectral decomposition [18] of Koopman operator  $\mathcal{K}_d$ :

$$\mathcal{K}_d \phi_i(\mathbf{x}_k) = \mu_i \phi_i(\mathbf{x}_k) \quad (5)$$

where  $i = 1, 2, \dots$ ,  $\mu_i \in \mathbb{C}$  are discrete-time *Koopman eigenvalues* and  $\phi_i : \mathbb{R}^n \rightarrow \mathbb{C}$  are *Koopman eigenfunctions* that define a set of intrinsic measurement coordinates [18]. Then one can represent the evolution of the observables using an eigenfunction expansion solution of the Koopman operator

$$\mathbf{h}(\mathbf{x}_k) = \sum_{i=1}^{\infty} \phi_i(\mathbf{x}_k) \zeta_i \quad (6)$$

where  $\zeta_i \in \mathbb{C}^l$  is the  $i$ th *Koopman mode* associated with the  $i$ th Koopman eigenfunction  $\phi_i$ . And the observables  $\mathbf{h}(\mathbf{x}_k)$  can be evolved using (4), (5) and (6) so that

$$\begin{aligned} \mathcal{K}_d \mathbf{h}(\mathbf{x}_k) &= \mathcal{K}_d \sum_{i=1}^{\infty} \phi_i(\mathbf{x}_k) \zeta_i = \sum_{i=1}^{\infty} \mathcal{K}_d \phi_i(\mathbf{x}_k) \zeta_i \\ &= \sum_{i=1}^{\infty} \mu_i \phi_i(\mathbf{x}_k) \zeta_i = \mathbf{h}(\mathbf{x}_{k+1}) \end{aligned} \quad (7)$$

Furthermore, we obtain

$$\mathbf{h}(\mathbf{x}_k) = \sum_{i=1}^{\infty} \mu_i^{k-1} \phi_i(\mathbf{x}_1) \zeta_i, \quad k = 1, 2, \dots \quad (8)$$

As for  $\mathcal{K}_c$ , copman eigenvalues  $\lambda_i \in \mathbb{C}$  are defined as [17]

$$\begin{cases} \mathcal{K}_c \phi_i(\mathbf{x}_k) = e^{\lambda_i \Delta t} \phi_i(\mathbf{x}_k) \\ \mu_i = e^{\lambda_i \Delta t} \end{cases} \quad (9)$$

Specially, if  $\mathbf{h}(\mathbf{x}_k) = \mathbf{x}_k$  with  $l = n$ , according to (6) and (8), we have

$$\mathbf{x}_k = \sum_{i=1}^{\infty} \phi_i(\mathbf{x}_k) \zeta_i = \sum_{i=1}^{\infty} \mu_i^{k-1} \phi_i(\mathbf{x}_1) \zeta_i \quad (10)$$

The Koopman operator is infinite-dimensional and we can approximate it by searching for a finite number of Koopman eigenvalues, modes and eigenfunctions. Since data-driven approaches, based on sampled data, are generally popular for the approximation, we will focus on discrete-time dynamical system (3) and its related Koopman operator  $\mathcal{K}_d$ .

It is worth noting that  $\mathcal{K}_d$  captures all dynamic features about nonlinear dynamical system (3) [17], and its eigenfunctions specify a nonlinear change of coordinates in which system (3) becomes a linear infinite-dimensional system (7). Then by the spectral decomposition, one can easily analyze which modes dominate and excite the dynamic behavior of the system.

### B. Extended Dynamic Mode Decomposition

Under certain conditions, the infinite-dimensional Koopman operator  $\mathcal{K}_d$  can be approximated in finite dimensions using a data-driven method, Extended Dynamic Mode Decomposition (EDMD) [19], [20]. Consider a data set of  $m$  snapshot pairs:

$$\{(\mathbf{x}_j, \mathbf{y}_j)\}_{j=1}^m, \quad \mathbf{y}_j = \mathbf{F}(\mathbf{x}_j) \quad (11)$$

where  $\mathbf{x}_j, \mathbf{y}_j \in \mathbb{R}^n$ . The data matrices are defined as  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$  and  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m]$ , where  $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times m}$ . And the matrices of observables are defined as:

$$\begin{aligned} \mathbf{H}_x &= [\mathbf{h}(\mathbf{x}_1), \mathbf{h}(\mathbf{x}_2), \dots, \mathbf{h}(\mathbf{x}_m)]^\top \\ \mathbf{H}_y &= [\mathbf{h}(\mathbf{y}_1), \mathbf{h}(\mathbf{y}_2), \dots, \mathbf{h}(\mathbf{y}_m)]^\top \end{aligned} \quad (12)$$

where  $\mathbf{H}_x, \mathbf{H}_y \in \mathbb{R}^{m \times l}$ . From (4), we have that  $\mathbf{H}_y = \mathbf{H}_x \mathcal{K}_d$ . Then EDMD gives a finite-dimensional approximation to the Koopman operator:

$$\mathbf{K} \triangleq \mathbf{H}_x^+ \mathbf{H}_y \quad (13)$$

where  $\mathbf{K} \in \mathbb{R}^{l \times l}$ . Note that EDMD requires that  $l \ll m$  holds [28]. The properties of  $\mathbf{K}$ , whose derivation can be found in [20], are introduced as follows:

a) The  $i$ th eigenvalue of  $\mathbf{K}$ ,  $\mu_i$ , is an estimate of the eigenvalue of  $\mathcal{K}_d$ .

b) The corresponding right eigenvector,  $\mathbf{v}_i$ , approximates the Koopman eigenfunction  $\phi_i(\mathbf{x}_k)$  by

$$\phi_i \triangleq \mathbf{h}^\top \mathbf{v}_i \quad (14)$$

c) The left eigenvector,  $\mathbf{w}_i$ , is involved in the approximation of the Koopman mode,  $\zeta_i$ .

Specially, when  $\mathbf{H}_x = \mathbf{X}^\top \in \mathbb{R}^{m \times n}$  and  $\mathbf{H}_y = \mathbf{Y}^\top \in \mathbb{R}^{m \times n}$ , EDMD degenerates into the traditional DMD, giving the result  $\mathbf{K} \triangleq (\mathbf{X}^\top)^+ \mathbf{Y}^\top \in \mathbb{R}^{n \times n}$ .

### C. Model Reduction

For discrete-time dynamical system (3), assume that a transformation matrix  $\mathbf{T}$  can be found to make

$$\mathbf{x}_k \approx \mathbf{T} \mathbf{z}_k \quad (15)$$

holds, where  $\mathbf{T} \in \mathbb{R}^{n \times r}$ ,  $\mathbf{z}_k \in \mathbb{R}^r$  and  $r < n$ . Then, by the Galerkin projection method [27], one can derive a discrete-time reduced-order model

$$\mathbf{z}_{k+1} = \mathbf{T}^+ \mathbf{F}(\mathbf{T} \mathbf{z}_k) \quad (16)$$

And the analogous continuous-time reduced-order model is given by

$$\frac{d\mathbf{z}}{dt} = \mathbf{T}^+ \mathbf{f}(\mathbf{T} \mathbf{z}) \quad (17)$$

Note that  $\mathbb{R}^r$  is a linear subspace of  $\mathbb{R}^n$  since  $r < n$ . Reduced-order model (16) evolves in  $\mathbb{R}^r$  and approximates full-order system (3) in some sense. Generally, searching for a matrix  $\mathbf{T}$  to minimize  $\|\mathbf{x}_k - \mathbf{T} \mathbf{z}_k\|$  is extremely challenging. Fortunately, the Koopman theory provides an efficient way to calculate  $\mathbf{T}$ , which is presented in Section IV.

## III. A DATA-DRIVEN APPROXIMATION OF THE KOOPMAN OPERATOR

Approximating the Koopman operator as accurately as possible is a significant task. From (14), the set of observables is required to be sufficiently *large* so that  $\phi_i(\mathbf{x}_k) \in \text{span}\{h_1, \dots, h_j\}$ , which implies that more observables will produce a more accurate approximation of the Koopman operator. However, EDMD requires an  $l \times l$  matrix  $\mathbf{K}$  to be formed, taking  $\mathcal{O}(l^2 m)$  calculation time. As the dimension of state space  $n$  increases, the number of observables required to be considered,  $l$ , grows rapidly, resulting in a huge amount of computation and the possibility of violating the condition  $l \ll m$ . Moreover, EDMD starts from a set of pre-configured dictionary functions to select observables, lacking analytical construction. Consequently, EDMD fails to *systematically* construct a *large* observable space to accurately estimate the Koopman operator of a high-dimensional dynamical system.

In order to consider the case  $l \gg m$  and obtain a more precise approximation, we propose a Polynomialization based Multi-trajectory Kernel Dynamic Mode Decomposition (PMK-DMD) algorithm, within the architecture of EDMD. Firstly, PMK-DMD renders dynamical system (1) into a polynomial form by Lie derivatives, producing a new state vector  $\bar{\mathbf{x}}$ . Then, PMK-DMD uses  $\bar{\mathbf{x}}$  and kernel DMD to systematically construct an adequate set of observable function and estimate the Koopman operators, which can significantly reduce the amount of computation. Finally, a multi-trajectory approach is introduced in PMK-DMD to reduce the dependence of the results on sampled data.

### A. Polynomialization of the Dynamical System

For system (1), assume that it can be rewritten as a linear combination of elementary functions  $e(\mathbf{x})$ , that is, the  $i$ th equation becomes:

$$\dot{x}_i = \eta_0 + \boldsymbol{\eta}_x^\top \mathbf{x} + \eta_1 e_1(\mathbf{x}) + \dots + \eta_g e_g(\mathbf{x}) \quad (18)$$

where  $i \in \{1, 2, \dots, n\}$ ,  $\eta_{0,1,\dots,g}$  are constant coefficients and  $\boldsymbol{\eta}_x$  is a constant vector. The elementary functions include  $\sin \boldsymbol{x}$ ,  $\cos \boldsymbol{x}$ ,  $\boldsymbol{x}^\alpha$ ,  $e^{\boldsymbol{x}}$ ,  $\boldsymbol{x}/(b + \boldsymbol{x})$  and so on, as well as compositions of these elementary functions, which cover a wide range of models encountered in engineering.

The polynomialization procedure, as shown in Algorithm 1, converts system (18) into a polynomial system [22], [29]:

$$\dot{\bar{\boldsymbol{x}}} = \bar{\boldsymbol{f}}(\bar{\boldsymbol{x}}) \quad (19)$$

where  $\bar{\boldsymbol{x}} \in \mathbb{R}^q$  ( $q = n + g + k$ ) is a new state vector,  $\bar{\boldsymbol{f}} : \mathbb{R}^q \rightarrow \mathbb{R}^q$  is a polynomial vector field with degree  $d$ ,  $k$  is the number of new elementary functions introduced by Step 0~4 in Algorithm 1. The specific form is as follows:

$$\bar{\boldsymbol{x}} = [x_1, x_2, \dots, x_n, \tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{g+k}]^\top \quad (20a)$$

$$\bar{f}_i = \eta_0 + \boldsymbol{\eta}_x^\top \boldsymbol{x} + \eta_1 \tilde{x}_1 + \dots + \eta_g \tilde{x}_g, \quad i = 1, 2, \dots, n \quad (20b)$$

$$\bar{f}_{n+j} = \mathcal{L}_f e_j(\boldsymbol{x}), \quad j = 1, 2, \dots, g + k \quad (20c)$$

---

**Algorithm 1:** Polynomialization procedure

---

**Input:** System (18).

**Output:** Polynomial system  $\dot{\bar{\boldsymbol{x}}} = \bar{\boldsymbol{f}}(\bar{\boldsymbol{x}})$ .

**Step 0:** Set  $j = 1$ ,  $\bar{j} = g$ .

**Step 1:** Introduce a new variable  $\tilde{x}_j = e_j(\boldsymbol{x})$ .

**Step 2:** Replace  $e_j(\boldsymbol{x})$  by  $\tilde{x}_j$  in the original equations (18).

**Step 3:** Add the Lie derivative of  $\tilde{x}_j$ , i.e.,

$\dot{\tilde{x}}_j = \mathcal{L}_f e_j(\boldsymbol{x}) = \frac{\partial e_j(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{f}$ , in the set of original equations.

**Step 4:** If  $j \leq \bar{j}$ , set  $j = j + 1$  and go to Step 1. Otherwise, go to Step 5.

**Step 5:** If the above procedure introduces  $k$  new elementary functions, define them as

$\{e_{g+1}(\boldsymbol{x}), e_{g+2}(\boldsymbol{x}), \dots, e_{g+k}(\boldsymbol{x})\}$ . Set  $j = j + 1$ ,  $\bar{j} = \bar{j} + k$ , and go to Step 1. Otherwise, go to Step 6.

**Step 6:** Let  $\bar{\boldsymbol{x}} = [x_1, x_2, \dots, x_n, \tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{g+k}]^\top$ . The new vector field constructed by the above procedure is denoted by  $\bar{\boldsymbol{f}}$ . Return  $\dot{\bar{\boldsymbol{x}}} = \bar{\boldsymbol{f}}(\bar{\boldsymbol{x}})$ .

---

Furthermore, for compositions of elementary functions  $e(\boldsymbol{x}) = (e_2 \circ e_1)(\boldsymbol{x}) = e_2(e_1(\boldsymbol{x}))$ , we have:

a) Introduce new variables  $\tilde{x}_1 = e_1(\boldsymbol{x})$  and  $\tilde{x}_2 = e_2(\tilde{x}_1)$ .

b) Replace  $e_2(e_1(\boldsymbol{x}))$  by  $\tilde{x}_2$  in the original equations.

c) Add  $\dot{\tilde{x}}_1 = \frac{\partial e_1(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{f}$  and  $\dot{\tilde{x}}_2 = \frac{\partial e_2(\tilde{x}_1)}{\partial \tilde{x}_1} \dot{\tilde{x}}_1$  in the set of original equations.

**Example 1.** Consider a dynamical system:

$$\begin{aligned} \dot{x}_1 &= 1 + x_1 + 1/(1 + e^{-x_2}) \\ \dot{x}_2 &= \cos x_2 \end{aligned}$$

Then by the polynomialization procedure, we obtain:

$$\begin{aligned} \bar{\boldsymbol{x}} &= [x_1, x_2, 1/(1 + e^{-x_2}), e^{-x_2}, \cos x_2, \sin x_2]^\top \\ \dot{\tilde{x}}_1 &= 1 + \tilde{x}_1 + \tilde{x}_3, \quad \dot{\tilde{x}}_2 = \tilde{x}_5 \\ \dot{\tilde{x}}_3 &= \tilde{x}_3^2 \tilde{x}_4 \tilde{x}_5, \quad \dot{\tilde{x}}_4 = -\tilde{x}_4 \tilde{x}_5 \\ \dot{\tilde{x}}_5 &= -\tilde{x}_6 \tilde{x}_5, \quad \dot{\tilde{x}}_6 = \tilde{x}_5^2 \end{aligned}$$

**Example 2.** Consider a single-machine infinite-bus power system [22], [30]:

$$\dot{\delta} = \omega, \quad \dot{\omega} = 21.36 - \omega - 21.36 \cos \delta - 78.58 \sin \delta$$

By Algorithm 1, a polynomial system is produced:

$$\begin{aligned} \bar{\boldsymbol{x}} &= [\delta, \omega, \cos \delta, \sin \delta]^\top \\ \dot{\tilde{x}}_1 &= \tilde{x}_2, \quad \dot{\tilde{x}}_2 = 21.36 - \tilde{x}_2 - 21.36 \tilde{x}_3 - 78.58 \tilde{x}_4 \\ \dot{\tilde{x}}_3 &= -\tilde{x}_4 \tilde{x}_2, \quad \dot{\tilde{x}}_4 = \tilde{x}_3 \tilde{x}_2 \end{aligned}$$

## B. Kernel DMD

### 1) Construct Observable Functions

For system (19), the set of all monomials defined on  $\bar{\boldsymbol{x}}$  with degree up to  $d$  is represented by

$$\mathcal{M} = \{1, \bar{x}_1, \dots, \bar{x}_q, \bar{x}_1^2, \bar{x}_1 \bar{x}_2, \dots, \bar{x}_2^2, \bar{x}_2 \bar{x}_3, \dots\}$$

And the number of elements of set  $\mathcal{M}$  can be calculated by  $N_M = (q + d)! / (q! d!)$ .

Then we regard  $\mathcal{M}$  as a set of observable functions. Such an approach is derived from paper [22]. In [22], the author takes all the monomials that appear in  $\bar{\boldsymbol{f}}(\bar{\boldsymbol{x}})$  as observable functions, denoted by  $\mathcal{M}_c$ . Obviously,  $\mathcal{M}_c \subset \mathcal{M}$ . The approach implies that

$$\bar{f}_i \in \text{span}\{\mathcal{M}_c\} \Rightarrow \bar{f}_i \in \text{span}\{\mathcal{M}\} \quad (21)$$

which makes the following condition more likely to be true, with sufficient large  $N_M$ :

$$\phi_i \in \text{span}\{\mathcal{M}_c\} \quad \text{or} \quad \phi_i \in \text{span}\{\mathcal{M}\} \quad (22)$$

Condition (22) guarantees the accuracy of Koopman eigenfunction estimation. Selecting  $\mathcal{M}_c$  as observables in EDMD provides a more satisfactory result compared with other traditional methods, which has been verified in [22]. Since  $\mathcal{M}_c \subset \mathcal{M}$ , it is advisable to choose  $\mathcal{M}$  as observable functions.

Now we recall (2), (4) and let  $\boldsymbol{h}(\bar{\boldsymbol{x}}) = \mathcal{M}$ ,  $l = N_M$ . Rather than explicitly representing  $\boldsymbol{h}$ , the kernel trick is a common technique for implicitly computing inner products of  $\boldsymbol{h}$  [31], [32]. We define a kernel function  $p : \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}$  that computes inner products in feature space. A brief example with polynomial kernel is shown as follows:

$$p(\boldsymbol{x}, \boldsymbol{y}) = (1 + \boldsymbol{y}^\top \boldsymbol{x})^2 \quad (23)$$

with  $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ , whose expanded form is:

$$\begin{aligned} p(\boldsymbol{x}, \boldsymbol{y}) &= (1 + x_1 y_1 + x_2 y_2)^2 \\ &= 1 + 2x_1 y_1 + 2x_2 y_2 + 2x_1 x_2 y_1 y_2 + x_1^2 y_1^2 + x_2^2 y_2^2 \\ &= \boldsymbol{h}(\boldsymbol{y})^\top \boldsymbol{h}(\boldsymbol{x}) \\ \boldsymbol{h}(\boldsymbol{x}) &= [1, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1 x_2, x_1^2, x_2^2]^\top \\ \boldsymbol{h}(\boldsymbol{y}) &= [1, \sqrt{2}y_1, \sqrt{2}y_2, \sqrt{2}y_1 y_2, y_1^2, y_2^2]^\top \end{aligned} \quad (24)$$

where  $\boldsymbol{h}(\boldsymbol{x})$  contains all monomials defined on  $\boldsymbol{x}$  with degree up to 2 and can be regarded as observables from (22).

Therefore, we use  $\bar{\boldsymbol{x}}$  and a polynomial kernel function  $p$  to implicitly represent observables  $\boldsymbol{h}(\bar{\boldsymbol{x}})$ , i.e.  $\mathcal{M}$ , with degree  $d$ .

$$p(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \triangleq (1 + \bar{\boldsymbol{y}}^\top \bar{\boldsymbol{x}})^d = \boldsymbol{h}(\bar{\boldsymbol{y}})^\top \boldsymbol{h}(\bar{\boldsymbol{x}}) \quad (25)$$

## 2) Approximate the Koopman Operator

We recall (12) and modify it from the above analysis, that is:

$$\begin{aligned} \mathbf{H}_x &= [\mathbf{h}(\bar{\mathbf{x}}_1), \mathbf{h}(\bar{\mathbf{x}}_2), \dots, \mathbf{h}(\bar{\mathbf{x}}_m)]^\top \\ \mathbf{H}_y &= [\mathbf{h}(\bar{\mathbf{y}}_1), \mathbf{h}(\bar{\mathbf{y}}_2), \dots, \mathbf{h}(\bar{\mathbf{y}}_m)]^\top \end{aligned} \quad (26)$$

where  $\mathbf{H}_x, \mathbf{H}_y \in \mathbb{R}^{m \times l}$  and  $l = N_M$ . Each pair  $\{\mathbf{h}(\bar{\mathbf{x}}_i), \mathbf{h}(\bar{\mathbf{y}}_i)\}_{i=1}^m$  represents the values of a very large array of observable functions  $\mathbf{h}$ , associated with  $\mathcal{M}$ , at the points  $\{\bar{\mathbf{x}}_i, \bar{\mathbf{y}}_i\}_{i=1}^m$  defined on the snapshot pair  $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^m$ . Note that we consider the case  $l \gg m$ .

Then take the singular value decomposition (SVD) of  $\mathbf{H}_x$ :

$$\mathbf{H}_x \approx \mathbf{Q}\mathbf{\Sigma}\mathbf{Z}^\top \quad (27)$$

where  $\mathbf{Q} \in \mathbb{R}^{m \times c}$ ,  $\mathbf{\Sigma} \in \mathbb{R}^{c \times c}$  and  $\mathbf{Z} \in \mathbb{R}^{l \times c}$ . Here  $c$  is the rank of proper orthogonal decomposition (POD) truncation [33]. The left singular vectors  $\mathbf{Q}$  are POD modes. The columns of  $\mathbf{Q}$  are orthonormal, i.e.  $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$ . Similarly,  $\mathbf{Z}^\top \mathbf{Z} = \mathbf{I}$ .

An eigenvector of  $\mathbf{K}$  with  $\mu \neq 0$  could be written as:

$$\mathbf{v} = \mathbf{Z}\hat{\mathbf{v}} \quad (28)$$

for some  $\hat{\mathbf{v}} \in \mathbb{C}^c$ . The eigenvalue problem of  $\mathbf{K}$  can be written as:

$$\mu \mathbf{v} = \mathbf{K} \mathbf{v} \quad (29)$$

From (13) and (27)–(29), we have

$$\begin{aligned} \mu \mathbf{Z}\hat{\mathbf{v}} &= \mathbf{H}_x^+ \mathbf{H}_y \mathbf{Z}\hat{\mathbf{v}} \\ &= \mathbf{Z}\mathbf{\Sigma}^{-1} \mathbf{Q}^\top \mathbf{H}_y \mathbf{Z}\hat{\mathbf{v}} \\ &= \mathbf{Z}[(\mathbf{\Sigma}^{-1} \mathbf{Q}^\top)(\mathbf{H}_y \mathbf{H}_x^\top)(\mathbf{Q}\mathbf{\Sigma}^{-1})]\hat{\mathbf{v}} \\ \Rightarrow \mu \hat{\mathbf{v}} &= [(\mathbf{\Sigma}^{-1} \mathbf{Q}^\top)(\mathbf{H}_y \mathbf{H}_x^\top)(\mathbf{Q}\mathbf{\Sigma}^{-1})]\hat{\mathbf{v}} \end{aligned} \quad (30)$$

Thus, (30) is the eigenvalue problem of matrix  $\hat{\mathbf{K}}$ , whose form is:

$$\hat{\mathbf{K}} \triangleq (\mathbf{\Sigma}^{-1} \mathbf{Q}^\top) \hat{\mathbf{H}} (\mathbf{Q}\mathbf{\Sigma}^{-1}) \quad (31)$$

where  $\hat{\mathbf{H}} = \mathbf{H}_y \mathbf{H}_x^\top \in \mathbb{R}^{m \times m}$ . The eigenvalue of  $\hat{\mathbf{K}}$  is equal to that of  $\mathbf{K}$ , and the eigenvector of  $\hat{\mathbf{K}}$ , say  $\hat{\mathbf{v}}$ , can be used to compute  $\mathbf{v}$  from (28). Note that  $\hat{\mathbf{K}} \in \mathbb{R}^{c \times c}$ , so the computational cost of the eigenvalue decomposition is determined by the rank of POD truncation  $c$  rather than the dimension of observable space  $l$  that may be fairly large.

Since  $l$  is large, it may be difficult to take SVD for  $\mathbf{H}_x \in \mathbb{R}^{m \times l}$  in (27). To address this problem, we define a matrix  $\hat{\mathbf{S}}$ :

$$\begin{aligned} \hat{\mathbf{S}} &\triangleq \mathbf{H}_x \mathbf{H}_x^\top = \mathbf{Q}\mathbf{\Sigma}^2 \mathbf{Q}^\top \\ \Rightarrow \hat{\mathbf{S}} \mathbf{Q} &= \mathbf{Q}\mathbf{\Sigma}^2 \end{aligned} \quad (32)$$

Obviously, given  $\hat{\mathbf{S}} \in \mathbb{R}^{m \times m}$ , we can obtain  $\mathbf{Q}$  and  $\mathbf{\Sigma}$  via its eigenvalue decomposition, whose computational cost is determined by the number of snapshots  $m$  rather than  $l$ .

The  $ij$ -th elements of  $\hat{\mathbf{H}}$  and  $\hat{\mathbf{S}}$  are

$$\hat{\mathbf{H}}_{ij} = \mathbf{h}(\bar{\mathbf{x}}_j)^\top \mathbf{h}(\bar{\mathbf{y}}_i), \quad \hat{\mathbf{S}}_{ij} = \mathbf{h}(\bar{\mathbf{x}}_j)^\top \mathbf{h}(\bar{\mathbf{x}}_i) \quad (33)$$

We find that the calculation of  $\hat{\mathbf{H}}_{ij}$  and  $\hat{\mathbf{S}}_{ij}$  involves only the inner product of  $\mathbf{h}$ . Then, from (25),  $\hat{\mathbf{H}}_{ij}$  and  $\hat{\mathbf{S}}_{ij}$  can be calculated by

$$\hat{\mathbf{H}}_{ij} = p(\bar{\mathbf{y}}_i, \bar{\mathbf{x}}_j), \quad \hat{\mathbf{S}}_{ij} = p(\bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) \quad (34)$$

Note that (34) takes  $\mathcal{O}(m^2 q)$  calculation time each for matrix  $\hat{\mathbf{H}}$  and  $\hat{\mathbf{S}}$  to construct  $\hat{\mathbf{K}}$ , which distinctly outperforms EDMD that consumes  $\mathcal{O}(l^2 m)$  calculation time to construct  $\mathbf{K}$ , because of  $l \gg m$  and  $l \gg q$ .

After  $\hat{\mathbf{K}}$  is constructed, we show how to approximate the Koopman eigenvalues, modes, and eigenfunctions.

According to (29) and (30), the eigenvalues of  $\hat{\mathbf{K}}$ , i.e.  $\{\mu_i\}_{i=1}^c$ , can be regarded as the Koopman eigenvalues.

Let  $\hat{\mathbf{V}}$  be the matrix whose columns are the right eigenvectors of  $\hat{\mathbf{K}}$ . Using (14) and (28), we define the matrix of eigenfunction values as:

$$\begin{aligned} \Phi_x &\triangleq \mathbf{H}_x \mathbf{Z} \hat{\mathbf{V}} = (\mathbf{H}_x \mathbf{H}_x^\top) (\mathbf{Q}\mathbf{\Sigma}^{-1}) \hat{\mathbf{V}} \\ &= \hat{\mathbf{S}} (\mathbf{Q}\mathbf{\Sigma}^{-1}) \hat{\mathbf{V}} = (\mathbf{Q}\mathbf{\Sigma}^2 \mathbf{Q}^\top) (\mathbf{Q}\mathbf{\Sigma}^{-1}) \hat{\mathbf{V}} \\ &= \mathbf{Q}\mathbf{\Sigma} \hat{\mathbf{V}} \end{aligned} \quad (35)$$

Obviously,  $\Phi_x \in \mathbb{C}^{m \times c}$ . The  $ij$ -th element of  $\Phi_x$  represents the  $j$ -th eigenfunction  $\phi_j(\mathbf{x})$  evaluated at the  $i$ -th snapshot point  $\mathbf{x}_i$ .

To compute the Koopman modes, we use (10) and derive a regression problem:

$$\mathbf{X} = \Theta \Phi_x^\top \quad (36)$$

where  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m] \in \mathbb{R}^{n \times m}$  is the data matrix and  $\Theta = [\zeta_1, \zeta_2, \dots, \zeta_c] \in \mathbb{C}^{n \times c}$  is the Koopman mode matrix. Thus, we have

$$\begin{aligned} \Theta^\top &= \Phi_x^+ \mathbf{X}^\top = \hat{\mathbf{V}}^{-1} \mathbf{\Sigma}^{-1} \mathbf{Q}^\top \mathbf{X}^\top \\ &= \hat{\mathbf{W}}^\top \mathbf{\Sigma}^{-1} \mathbf{Q}^\top \mathbf{X}^\top \end{aligned} \quad (37)$$

where  $\hat{\mathbf{W}}$  is the matrix whose rows are the left eigenvectors of  $\hat{\mathbf{K}}$ , due to  $\hat{\mathbf{W}}^\top \hat{\mathbf{V}} = \mathbf{I}$ .

Ultimately, according to (10) and (36), a prediction function can be obtained:

$$\mathcal{P}(k) = \hat{\mathbf{x}}_k = \Theta \mathbf{D}^{k-1} \phi(\mathbf{x}_1), \quad k = 1, 2, \dots \quad (38)$$

where  $\hat{\mathbf{x}}_k$  is an estimate of  $\mathbf{x}_k$ ,  $\mathbf{D} = [\mu_1, \mu_2, \dots, \mu_c]_{\text{diag}} \in \mathbb{C}^{c \times c}$  is the Koopman eigenvalue matrix,  $\phi(\mathbf{x}_1) = [\phi_1(\mathbf{x}_1), \phi_2(\mathbf{x}_1), \dots, \phi_c(\mathbf{x}_1)]^\top \in \mathbb{C}^{c \times 1}$  is a vector whose element is the eigenfunction evaluated at  $\mathbf{x}_1$ .

## C. Multi-trajectory Approach

The above analysis estimates the Koopman operator based on a single trajectory  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$ . Because the estimated Koopman operator will be utilized for model reduction, we require it to capture the inherent characteristics of transient dynamics and not depend on specific trajectories. Hence, a multi-trajectory approach is proposed to reduce the effect of trajectories on the Koopman operator.

Consider  $s$  trajectories, each containing  $m$  snapshots. The data matrix is defined as:

$$\begin{aligned} \underline{\mathbf{X}} &= [\mathbf{x}_{11}, \mathbf{x}_{12}, \dots, \mathbf{x}_{1m}, \mathbf{x}_{21}, \mathbf{x}_{22}, \dots, \mathbf{x}_{2m} \\ &\quad \dots, \mathbf{x}_{s1}, \mathbf{x}_{s2}, \dots, \mathbf{x}_{sm}] \end{aligned} \quad (39)$$

Then  $\underline{\mathbf{Y}}_{ij} = \mathbf{F}(\underline{\mathbf{X}}_{ij})$  and  $\underline{\mathbf{X}}, \underline{\mathbf{Y}} \in \mathbb{R}^{n \times \underline{m}}$ ,  $\underline{m} = s \times m$ . Similar to (26), the matrices of observables are defined as:

$$\underline{\mathbf{H}}_x = [\mathbf{h}(\bar{\mathbf{x}}_{11}), \dots, \mathbf{h}(\bar{\mathbf{x}}_{1m}), \mathbf{h}(\bar{\mathbf{x}}_{21}), \dots, \mathbf{h}(\bar{\mathbf{x}}_{2m}),$$

$$\begin{aligned} & \cdots, \mathbf{h}(\bar{\mathbf{x}}_{s1}), \cdots, \mathbf{h}(\bar{\mathbf{x}}_{sm})]^\top \\ \underline{\mathbf{H}}_y &= [\mathbf{h}(\bar{\mathbf{y}}_{11}), \cdots, \mathbf{h}(\bar{\mathbf{y}}_{1m}), \mathbf{h}(\bar{\mathbf{y}}_{21}), \cdots, \mathbf{h}(\bar{\mathbf{y}}_{2m}), \\ & \cdots, \mathbf{h}(\bar{\mathbf{y}}_{s1}), \cdots, \mathbf{h}(\bar{\mathbf{y}}_{sm})]^\top \end{aligned} \quad (40)$$

where  $\underline{\mathbf{H}}_x, \underline{\mathbf{H}}_y \in \mathbb{R}^{m \times l}$  and  $\underline{\mathbf{H}}_y = \underline{\mathbf{H}}_x \mathcal{K}_d$  holds. Then the approximation matrix of  $\mathcal{K}_d$  is:

$$\underline{\mathbf{K}} \triangleq \underline{\mathbf{H}}_x^+ \underline{\mathbf{H}}_y \quad (41)$$

where  $\underline{\mathbf{K}} \in \mathbb{R}^{l \times l}$ , similar to (13). Here  $\underline{\mathbf{K}}$  drives the evolution of observable functions for each trajectory.

Similar to the procedure in Section III.B.2), we define two matrices  $\hat{\underline{\mathbf{H}}} \in \mathbb{R}^{m \times m}$  and  $\hat{\underline{\mathbf{S}}} \in \mathbb{R}^{m \times m}$ :

$$\hat{\underline{\mathbf{H}}} = \begin{bmatrix} \hat{\mathbf{H}}^{(11)} & \cdots & \hat{\mathbf{H}}^{(1s)} \\ \vdots & \vdots & \vdots \\ \hat{\mathbf{H}}^{(s1)} & \cdots & \hat{\mathbf{H}}^{(ss)} \end{bmatrix}, \quad \hat{\underline{\mathbf{S}}} = \begin{bmatrix} \hat{\mathbf{S}}^{(11)} & \cdots & \hat{\mathbf{S}}^{(1s)} \\ \vdots & \vdots & \vdots \\ \hat{\mathbf{S}}^{(s1)} & \cdots & \hat{\mathbf{S}}^{(ss)} \end{bmatrix} \quad (42)$$

where  $\hat{\mathbf{H}}^{(\alpha\beta)}, \hat{\mathbf{S}}^{(\alpha\beta)} \in \mathbb{R}^{m \times m}$  with  $\alpha, \beta \in \{1, 2, \dots, s\}$ , and

$$\begin{aligned} \hat{\mathbf{H}}_{ij}^{(\alpha\beta)} &= \mathbf{h}(\bar{\mathbf{x}}_{\beta j})^\top \mathbf{h}(\bar{\mathbf{y}}_{\alpha i}) = p(\bar{\mathbf{y}}_{\alpha i}, \bar{\mathbf{x}}_{\beta j}) \\ \hat{\mathbf{S}}_{ij}^{(\alpha\beta)} &= \mathbf{h}(\bar{\mathbf{x}}_{\beta j})^\top \mathbf{h}(\bar{\mathbf{x}}_{\alpha i}) = p(\bar{\mathbf{x}}_{\alpha i}, \bar{\mathbf{x}}_{\beta j}) \end{aligned} \quad (43)$$

with  $i, j \in \{1, 2, \dots, m\}$ . Then we compute the eigendecomposition of  $\hat{\underline{\mathbf{S}}}$  from (32):

$$\hat{\underline{\mathbf{S}}}\underline{\mathbf{Q}} = \underline{\mathbf{Q}}\underline{\Sigma}^2 \quad (44)$$

where  $\underline{\mathbf{Q}} \in \mathbb{R}^{m \times c}$  and  $\underline{\Sigma} \in \mathbb{R}^{c \times c}$ . From (31), the matrix  $\hat{\underline{\mathbf{K}}}$  can be constructed:

$$\hat{\underline{\mathbf{K}}} \triangleq (\underline{\Sigma}^{-1} \underline{\mathbf{Q}}^\top) \hat{\underline{\mathbf{H}}} (\underline{\mathbf{Q}} \underline{\Sigma}^{-1}) \quad (45)$$

and  $\hat{\underline{\mathbf{K}}} \in \mathbb{R}^{c \times c}$ . Similarly, take the eigendecomposition of  $\hat{\underline{\mathbf{K}}}$  to calculate its right eigenmatrix  $\hat{\underline{\mathbf{V}}}$  and eigenvalue matrix  $\underline{\mathbf{D}}$ .

Finally, according to (35)–(38), the Koopman eigenfunction matrix  $\underline{\Phi}_x$ , Koopman mode matrix  $\underline{\Theta}$  and prediction function  $\mathcal{P}(\alpha, k)$  can be easily obtained:

$$\underline{\Phi}_x = \underline{\mathbf{Q}} \underline{\Sigma} \hat{\underline{\mathbf{V}}} \quad (46a)$$

$$\underline{\Theta}^\top = \underline{\Phi}_x^+ \underline{\mathbf{X}}^\top \quad (46b)$$

$$\mathcal{P}(\alpha, k) = \hat{\mathbf{x}}_{\alpha k} = \underline{\Theta} \underline{\mathbf{D}}^{k-1} \underline{\phi}(\mathbf{x}_{\alpha 1}) \quad (46c)$$

$$\alpha = 1, 2, \dots, s. \quad k = 1, 2, \dots$$

where  $\underline{\Phi}_x \in \mathbb{C}^{m \times c}$ ,  $\underline{\Theta} = [\zeta_1, \zeta_2, \dots, \zeta_c] \in \mathbb{C}^{n \times c}$ , and  $\hat{\mathbf{x}}_{\alpha k}$  is an estimate of  $\mathbf{x}_{\alpha k}$ ,  $\underline{\mathbf{D}} = [\mu_1, \mu_2, \dots, \mu_c]_{\text{diag}} \in \mathbb{C}^{c \times c}$  is the Koopman eigenvalue matrix,  $\underline{\phi}(\mathbf{x}_{\alpha 1}) = [\phi_1(\mathbf{x}_{\alpha 1}), \phi_2(\mathbf{x}_{\alpha 1}), \dots, \phi_c(\mathbf{x}_{\alpha 1})]^\top \in \mathbb{C}^{c \times 1}$  is a vector whose element is the eigenfunction evaluated at  $\mathbf{x}_{\alpha 1}$ .

#### D. Framework of PMK-DMD Algorithm

Based on the above analysis, we summarize the procedure of PMK-DMD algorithm, as shown in Algorithm 2, and its framework is given by Fig. 1. PMK-DMD has the following several outstanding advantages:

a) The observable space is sufficiently large and can be systematically constructed. Since Koopman operator  $\mathcal{K}_d$  is infinite-dimensional, its approximation  $\underline{\mathbf{K}} \in \mathbb{R}^{l \times l}$  needs to be as large as possible to ensure a high degree of fit, i.e.

#### Algorithm 2: PMK-DMD algorithm

**Input:** Data matrices  $\underline{\mathbf{X}}$  and  $\underline{\mathbf{Y}}$ , rank of POD truncation  $c$ .

**Output:** Koopman eigenvalue matrix  $\underline{\mathbf{D}}$ , eigenfunction matrix  $\underline{\Phi}_x$ , mode matrix  $\underline{\Theta}$  and prediction function  $\mathcal{P}(\alpha, k)$ .

**Step 1:** Implement the polynomialization of dynamical system (18) by Algorithm 1, to obtain new state vector  $\bar{\mathbf{x}}$ . Since  $\bar{\mathbf{x}}$  is a function vector of original states  $\mathbf{x}$ , we have  $\bar{\mathbf{x}}_{\alpha k} = \bar{\mathbf{x}}(\mathbf{x}_{\alpha k})$  and  $\bar{\mathbf{y}}_{\alpha k} = \bar{\mathbf{y}}(\mathbf{y}_{\alpha k})$ , with  $k \in \{1, 2, \dots, m\}$ ,  $\alpha \in \{1, 2, \dots, s\}$ .

**Step 2:** Construct matrices  $\hat{\underline{\mathbf{H}}}$  and  $\hat{\underline{\mathbf{S}}}$  from (43), by a polynomial kernel function  $p$  defined in (25).

**Step 3:** Compute the eigendecomposition of  $\hat{\underline{\mathbf{S}}}$  using (44) and  $c$  to obtain  $\underline{\mathbf{Q}}$  and  $\underline{\Sigma}$ .

**Step 4:** Construct  $\hat{\underline{\mathbf{K}}}$  by (45). Take the eigendecomposition of  $\hat{\underline{\mathbf{K}}}$  to produce its right eigenmatrix  $\hat{\underline{\mathbf{V}}}$  and eigenvalue matrix  $\underline{\mathbf{D}}$ .

**Step 5:** Calculate the Koopman eigenfunction matrix  $\underline{\Phi}_x$ , Koopman mode matrix  $\underline{\Theta}$  and prediction function  $\mathcal{P}(\alpha, k)$ , from (46a)–(46c).

**Step 6:** Return  $\underline{\mathbf{D}}$ ,  $\underline{\Phi}_x$ ,  $\underline{\Theta}$  and  $\mathcal{P}(\alpha, k)$ .

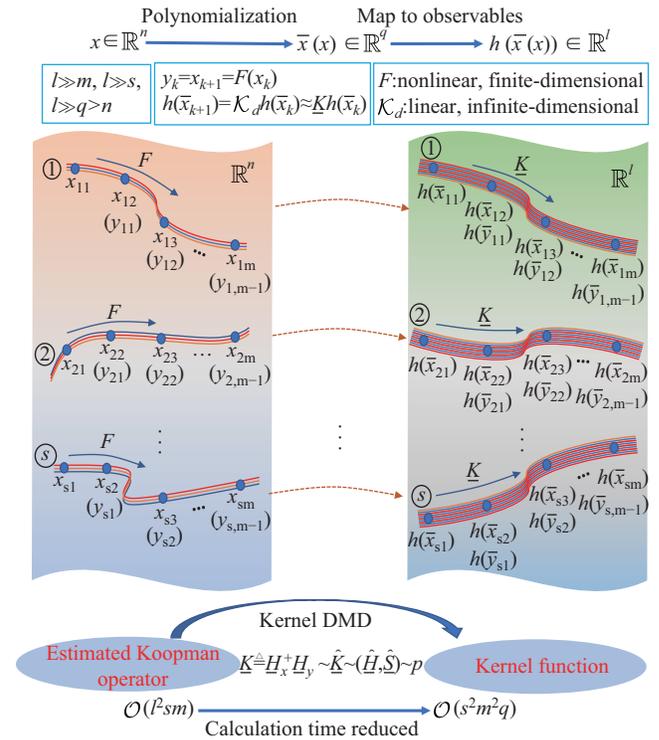


Fig. 1. Framework of PMK-DMD method. The thickness of these trajectories represents the number of components they contain, i.e. the dimension of the related space.

the number of observables  $l \rightarrow +\infty$ . PMK-DMD provides a method for systematically constructing numerous observables, by polynomializing the original system. All monomials defined on the new state vector  $\bar{\mathbf{x}}$  span an observable space, with a large dimension  $l = N_M$ .

b) The calculation time is highly reduced. Although the observable space is vast, PMK-DMD does not explicitly represent observable matrices  $\underline{\mathbf{H}}_x, \underline{\mathbf{H}}_y$ , but implicitly represents them through their inner products  $\hat{\underline{\mathbf{H}}}, \hat{\underline{\mathbf{S}}}$ . Then,  $\hat{\underline{\mathbf{H}}}, \hat{\underline{\mathbf{S}}}$  can be easily computed by a kernel function  $p$ , which takes only



$$\alpha = 1, 2, \dots, s. \quad k = 1, 2, \dots \quad (52)$$

where  $\tilde{x}_{\alpha k}$  is an estimate of  $x_{\alpha k}$ ,  $\{\underline{\Theta}_r, \underline{D}_r, \underline{\phi}_r\}$  are matrices formed by keeping the relevant elements of  $L_e$  in matrices  $\{\underline{\Theta}_p, \underline{D}_p, \underline{\phi}_p\}$ . Thus,  $\underline{\Theta}_r \in \mathbb{R}^{n \times r}$ ,  $\underline{D}_r \in \mathbb{R}^{r \times r}$  and  $\underline{\phi}_r \in \mathbb{R}^{r \times 1}$ . Obviously,

$$\mathbf{T} = \underline{\Theta}_r \quad (53)$$

with  $\mathbf{T} \in \mathbb{R}^{n \times r}$ . In (52),  $z_{\alpha, k} \in \mathbb{R}^{r \times 1}$  is the state vector for the reduced-order model.

Finally, according to Section II.C, the reduced-order model can be represented by

$$z_{\alpha, k+1} = \mathbf{T}^+ \mathbf{F}(\mathbf{T} z_{\alpha, k}), \quad (54a)$$

$$\text{or } \frac{dz}{dt} = \mathbf{T}^+ \mathbf{f}(\mathbf{T} z). \quad (54b)$$

$$\alpha = 1, 2, \dots, s. \quad k = 1, 2, \dots$$

Assume that  $x = \mathbf{0}$  is the equilibrium point of system (1), i.e.  $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ , and it is asymptotically stable. Then from (54b),  $z = \mathbf{0}$  is also the equilibrium point of the reduced-order model. And we require the transformation matrix  $\mathbf{T}$  not to change the stability of  $z = \mathbf{0}$ . The Jacobian matrix at  $z = \mathbf{0}$  for system (54b) is denoted by

$$\mathbf{J}_r = \mathbf{T}^+ \mathbf{J} \mathbf{T} \quad (55)$$

where  $\mathbf{J}$  is the Jacobian matrix at  $x = \mathbf{0}$  for system (1). If the real part of all eigenvalues of  $\mathbf{J}_r$  are negative,  $z = \mathbf{0}$  is also asymptotically stable. Otherwise, we add additional eigenvalues to the set  $L_e$  and reconstruct  $\mathbf{T}$  such that  $\mathbf{J}_r$  satisfies the above condition.

Furthermore, since the modulus of each column of  $\mathbf{T}$  may be affected by the trajectories, we scale it such that  $\mathbf{T}_i = \mathbf{T}_i / \|\mathbf{T}_i\|$ , with  $i = 1, 2, \dots, r$ . Note that the above treatment does not change the eigenvalues of  $\mathbf{J}_r$ .

Finally, we call the whole process of using PMK-DMD to estimate a Koopman operator from sampling data and then deducing a reduced-order model through the dominant Koopman modes as *Koopman operator based model reduction* (KOMR) method.

## V. VERIFICATION AND COMPARISON

This section applies the above KOMR method to the IEEE 10-machine-39-bus power system case and the IEEE 16-machine-68-bus power system case. The resulting reduced-order model is nonlinear and can be used for transient stability analysis. In the IEEE 10-machine-39-bus power system case, we compare the effectiveness of the proposed method with that of the traditional modal analysis in terms of order reduction. Furthermore, we also compare the performance of PMK-DMD and DMD in various scenarios. At last, when we attempt to estimate stability regions, the computational savings due to model reduction are quantitatively shown. As for the IEEE 16-machine-68-bus power system case, the effectiveness of our method on larger scale systems is presented.

### A. Case I: IEEE 10-machine-39-bus Power System

#### 1) Modeling of Power Systems

Here, we introduce the internal node model [2] of a  $g$ -machine power system (Generator 1 is regarded as the reference machine):

$$\begin{aligned} \frac{d\delta_{i1}}{dt} &= \omega_i - \omega_1, \\ \frac{d\omega_i}{dt} &= \frac{\omega_b}{2H_i} (P_{mi} - P_{ei} - D_i \omega_i) \end{aligned} \quad (56)$$

where  $i = 1, 2, \dots, g$ ,  $\delta_{i1}$  represents the relative rotor angle [rad] and  $\omega_i$  is the rotor angular velocity [rad/s].  $P_{ei} = \sum_{j=1}^g E_i E_j \{G_{ij} \cos(\delta_{i1} - \delta_{j1}) + B_{ij} \sin(\delta_{i1} - \delta_{j1})\}$  is the electrical power [p.u.].  $E_i$ ,  $P_{mi}$  represent the internal voltage and mechanical input respectively [p.u.].  $B_{ij}$ ,  $G_{ij}$  are the susceptance and conductance between generators  $i$  and  $j$  respectively [p.u.].  $H_i$  is the time inertia constant [s] and  $D_i$  is the damping coefficient [p.u.].  $\omega_b = 120\pi$  is the reference value of  $\omega_i$ . Electrical loads are modeled as constant impedances. For simplicity, we shift the equilibrium point of system (56) to the origin. Assuming that  $(\delta_{21}^*, \delta_{31}^*, \dots, \delta_{g1}^*, 0, \dots, 0)$  is the stable equilibrium point, we define a new state vector  $x \in \mathbb{R}^n$  ( $n = 2g - 1$ ) as:

$$\begin{aligned} x &= [x_1, \dots, x_{g-1}, x_g, \dots, x_n] \\ &= [\Delta\delta_{21}, \dots, \Delta\delta_{g1}, \omega_1, \omega_2, \dots, \omega_g] \\ &= [\delta_{21} - \delta_{21}^*, \dots, \delta_{g1} - \delta_{g1}^*, \omega_1, \omega_2, \dots, \omega_g] \end{aligned} \quad (57)$$

The IEEE 10–39 power system is a well-known benchmark model for research, as shown in Fig. 2, and see [2] for more details.

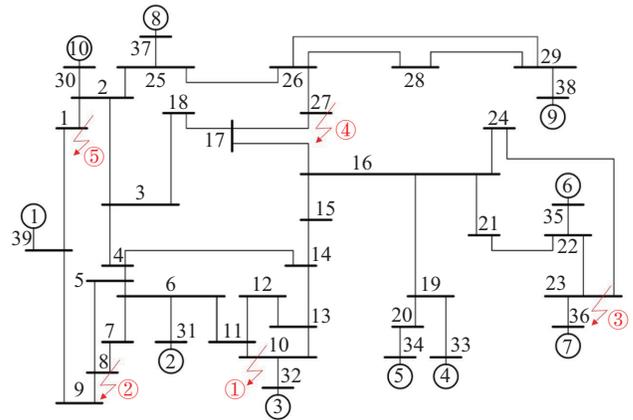


Fig. 2. One-line diagram of the IEEE 10–39 system.

In this case,  $g = 10$  and  $n = 19$ . More parameters are presented in Appendix A. As shown in Fig. 2, we set five separate faults on Bus 10, 8, 23, 27 and 1 respectively, locating throughout the whole area. By the time-domain simulation [35], these faults provide five post-fault trajectories, i.e.  $s = 5$ . Since low-frequency oscillations (often less than 8 Hz) are dominant in the electromechanical dynamics, we choose a sampling frequency of 20 Hz, i.e. the uniform sampling period  $\Delta t = 0.05$ s, according to the Nyquist–Shannon sampling theorem [36]. Set the number of sampling snapshots  $m = 120$ . Then the data matrices  $\underline{X}$  and  $\underline{Y}$  can be obtained.

When estimating Koopman operators, the sampled data should preferably come from the state trajectories inside the domain of attraction of the underlying system [17]. This means that unstable trajectories are not recommended, and the stability of the resulting reduced-order model is a sufficient condition for that of the full-order model. Thus, the trajectories chosen in this case are asymptotically stable.

## 2) The KOMR Method and the Traditional Modal Analysis in Terms of Model Reduction

We apply the KOMR method to the power system using PMK-DMD (Algorithm 2) with different values of  $c$ . Then compute the  $R^2 = R(\hat{x}_{\alpha k})$  of the results, as shown in Fig. 3. Generally, a small  $c$  tends to produce a small  $r$ , the dimension of the reduced-order model, but the performance about the related  $R^2$  may be disappointing. Hence, from Fig. 3, we choose a case with relatively small  $c$  and large  $R^2$ , i.e. the case of  $c = 19$ .

When  $c = 19$ , the continuous-time Koopman eigenvalues are presented in Fig. 4, and all of them are with negative real part because the post-fault trajectories sampled are convergent. Then we calculate  $C_i$ , the contribution of each eigenvalue, as shown in Table I, where  $\lambda_i^{(*)}$  denotes a pair of complex-conjugate eigenvalues  $\{\lambda_i, \lambda_i^*\}$ . It can be found that the contributions of  $\lambda_1^{(*)}$  and  $\lambda_2$  are outstanding. We choose

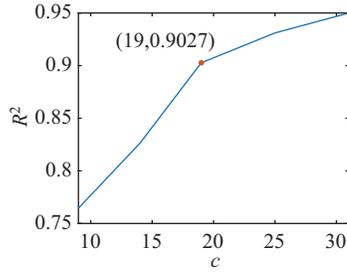


Fig. 3. Relationship of  $R^2$  and  $c$ .

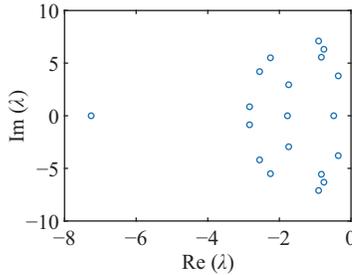


Fig. 4. Koopman eigenvalues.

TABLE I  
CONTRIBUTION OF EACH EIGENVALUE IN CASE I

	$\lambda_1^{(*)}$	$\lambda_2$	$\lambda_3^{(*)}$	$\lambda_4^{(*)}$	$\lambda_5^{(*)}$	$\lambda_6^{(*)}$
$\lambda_i^{(*)}$	-0.36 $\pm 3.79i$	-0.48	-0.76 $\pm 6.31i$	-0.83 $\pm 5.56i$	-0.90 $\pm 7.10i$	-1.74 $\pm 2.95i$
$C_i$	0.6171	0.3025	0.1008	0.0214	0.0360	0.0166
	$\lambda_7$	$\lambda_8^{(*)}$	$\lambda_9^{(*)}$	$\lambda_{10}^{(*)}$	$\lambda_{11}$	
$\lambda_i^{(*)}$	-1.77	-2.25 $\pm 5.51i$	-2.55 $\pm 4.20i$	-2.83 $\pm 0.86i$	-7.26	
$C_i$	0.0084	0.0144	0.0164	0.0628	0.0033	

different  $L_e$  to search their effect on the reduced-order model, as shown in Table II, where  $R^2 = R(\tilde{x}_{\alpha k})$ . Notice that the  $L_e$  selected by ordering the real parts of eigenvalues may not produce a satisfying  $R^2$ , like  $L_e = \{\lambda_1^{(*)}, \lambda_2, \lambda_3^{(*)}, \lambda_4^{(*)}\}$  with  $R^2 = 0.7879$ , compared with  $L_e = \{\lambda_1^{(*)}, \lambda_2, \lambda_3^{(*)}, \lambda_5^{(*)}\}$  with  $R^2 = 0.8475$ .

TABLE II  
THE EFFECT OF DIFFERENT  $l_e$  ON THE REDUCED-ORDER MODEL  
IN CASE I

$L_e$	$R^2$	$r$	$z = \mathbf{0}$
$\{\lambda_1^{(*)}, \lambda_2\}$	0.7768	3	stable
$\{\lambda_1^{(*)}, \lambda_2, \lambda_3^{(*)}\}$	0.8097	5	stable
$\{\lambda_1^{(*)}, \lambda_2, \lambda_3^{(*)}, \lambda_4^{(*)}\}$	0.7879	7	stable
$\{\lambda_1^{(*)}, \lambda_2, \lambda_3^{(*)}, \lambda_5^{(*)}\}$	0.8475	7	stable

Finally, we choose  $L_e = \{\lambda_1^{(*)}, \lambda_2, \lambda_3^{(*)}\}$  as the leading Koopman eigenvalues, because they have considerable contributions, and the resulting reduced-order model is stable and with a large  $R^2$ . Obviously,  $r = 5$ , that is the 19-dimensional model (56) can be reduced to a 5-dimensional model (54b). The resulting  $\mathbf{T}$  is given in Appendix A.

To verify the validity of the reduced-order model, we compare the proposed method with the traditional modal analysis [5], [11], [12]. First, we consider the linearized system around the stable equilibrium  $\mathbf{x} = \mathbf{0}$  of (56), represented by  $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$  with  $\mathbf{A} \in \mathbb{R}^{19 \times 19}$ . Let  $\mathcal{V}_i \in \mathbb{C}^{19}$  be right eigenvectors and  $\epsilon_i \in \mathbb{C}$  be eigenvalues of  $\mathbf{A}$ :

$$\mathbf{A}\mathcal{V}_i = \epsilon_i \mathcal{V}_i \quad (58)$$

Table III shows the eigenvalues  $\epsilon_i$ , and we find that  $\epsilon_1^{(*)}$ - $\epsilon_9^{(*)}$  are with the same real part, because system (56) uses the uniform damping ratio, i.e.  $D_1/H_1 = D_2/H_2 = \dots = D_g/H_g$ . In this case, the traditional modal analysis fails to select dominant modals by sorting the real part of the eigenvalues. Reluctantly, we select the low-frequency eigenvalues  $\mathbf{E} = [\epsilon_1^{(*)}, \epsilon_5^{(*)}, \epsilon_6^{(*)}]_{\text{diag}}$  as dominant modals and obtain the associated eigenvectors  $\mathcal{V} = [\mathcal{V}_1, \mathcal{V}_5, \mathcal{V}_6]$ . After taking a realification procedure for  $\mathbf{E}$  and  $\mathcal{V}$ ,  $\mathcal{V} \in \mathbb{R}^{19 \times 6}$  can be regarded as the transformation matrix and  $\mathbf{x} \approx \mathcal{V}\mathbf{z}$  holds, where  $\mathbf{z} \in \mathbb{R}^6$ . Then from (58), we obtain the reduced-order model by the modal analysis:

$$\dot{\mathbf{z}} = \mathcal{V}^+ \mathbf{A}\mathcal{V}\mathbf{z} = \mathbf{E}\mathbf{z} \quad (59)$$

where  $\mathbf{E} \in \mathbb{R}^{6 \times 6}$ .

TABLE III  
EIGENVALUES OF  $\mathbf{A}$  FOR THE MODAL ANALYSIS IN CASE I

$\epsilon_1^{(*)}$	$\epsilon_2^{(*)}$	$\epsilon_3^{(*)}$	$\epsilon_4^{(*)}$	$\epsilon_5^{(*)}$
-0.377	-0.377	-0.377	-0.377	-0.377
$\pm 3.784i$	$\pm 9.787i$	$\pm 9.65i$	$\pm 9.161i$	$\pm 5.846i$
$\epsilon_6^{(*)}$	$\epsilon_7^{(*)}$	$\epsilon_8^{(*)}$	$\epsilon_9^{(*)}$	$\epsilon_{10}$
-0.377	-0.377	-0.377	-0.377	
$\pm 6.389i$	$\pm 7.072i$	$\pm 7.898i$	$\pm 8.03i$	-0.754

The trajectories governed by the full-order model (blue), reduced-order model based on KOMR (red) and reduced-order model based on the traditional modal analysis (green) are

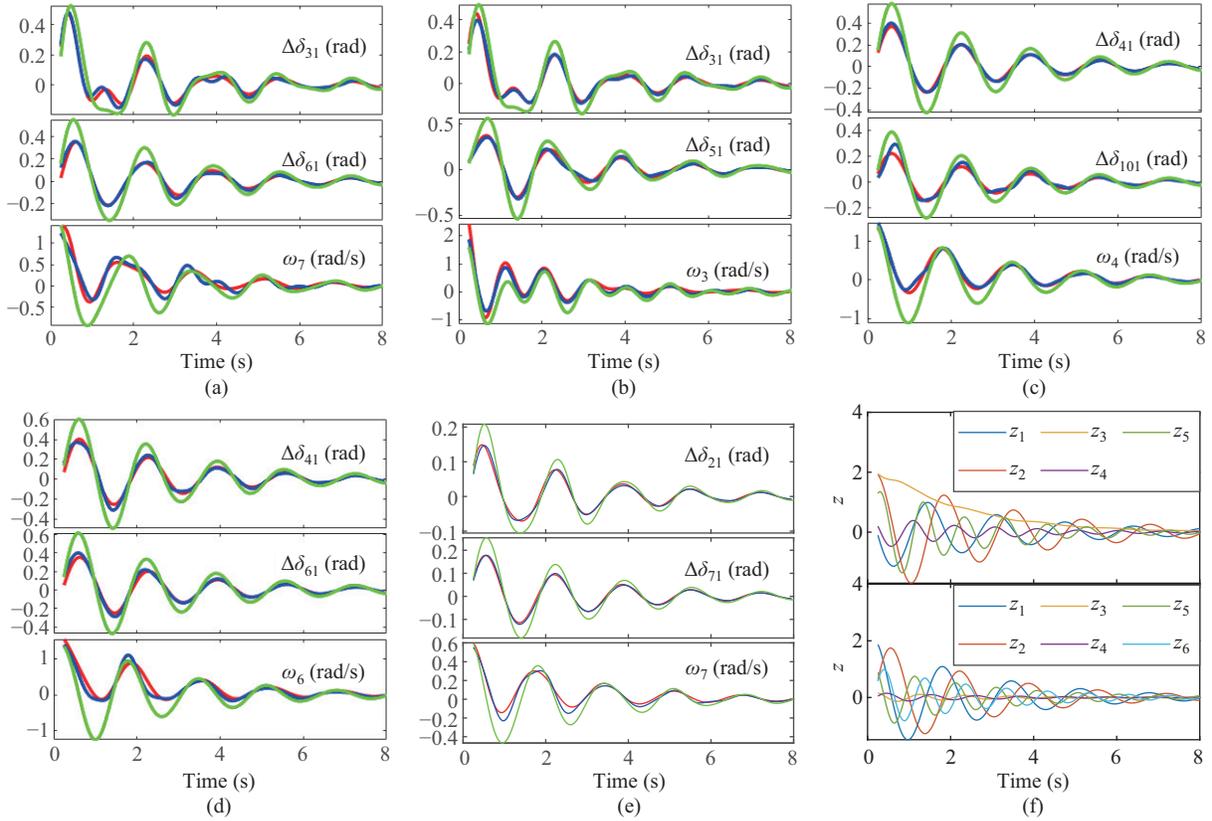


Fig. 5. Comparison of the trajectories governed by the full-order model (blue), reduced-order model based on KOMR (red) and reduced-order model based on the traditional modal analysis (green) in Case I. In (a)–(e), we select 3 components of each trajectory for presentation. The trajectories driven by the reduced-order systems, such as trajectory 1 shown in (f), are projected into the full-order state space for comparison. (a) Trajectory 1. (b) Trajectory 2. (c) Trajectory 3. (d) Trajectory 4. (e) Trajectory 5. (f) Reduced-order trajectory 1.

shown in Fig. 5(a)–(e). The trajectories driven by the reduced-order systems, such as trajectory 1 in Fig. 5(f) (above is system (54b), below is system (59)), are projected into the full-order state space for comparison. And the projection are  $x \approx Tz$  and  $x \approx Vz$  respectively. From Fig. 5, we notice that the 5-order model based on Koopman operator commendably fit the trajectories of the full-order system. The degree of fit is quantified as  $R^2 = R(\tilde{x}_{\alpha k}) = 0.8096$ . However, the 6-order model based on the modal analysis provides a disappointing result, with  $R^2 = 0.4091$ . Therefore, the model reduction based on Koopman operator is able to produce a lower-order model while guaranteeing a higher degree of fit. This suggests that the Koopman eigenvalues can perfectly characterize the dynamics under large perturbations, whereas the eigenvalues of the linearized model do not. Also, note that  $\lambda_1^{(*)} \approx \epsilon_1^{(*)}$ , indicating that the Koopman modes accurately contain the dominant eigenvalue of the linearized model.

### 3) Performance of PMK-DMD and Traditional DMD

There are several data-driven methods to approximate the Koopman operator, including DMD, EDMD, etc. Among them, EDMD and the proposed PMK-DMD return the same estimate if using the same observable functions, except that EDMD is slower to compute, see analysis in Section III.D(b). In this section, we compare the performance of the proposed PMK-DMD and traditional DMD [18], as shown in Fig. 6. Since DMD can only discuss a single trajectory, here it con-

siders the data from Trajectory 2, while PMK-DMD considers the data from all the above-mentioned trajectories. Moreover, both of the methods are set to be without POD truncation. In addition, we define the Root Mean Squared Error (RMSE) and Sum of Absolute Error (SAE):

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{m} \sum_{k=1}^m (x_k - \hat{x}_k)^2}, \\ \text{SAE} &= \sum_{k=1}^m |x_k - \hat{x}_k| \end{aligned} \quad (60)$$

In normal cases, both PMK-DMD and DMD fit the full-order trajectory well, with the former having lower RMSE and SAE, as shown in Fig. 6(a). When a part of the measurements are missing, e.g. the measurements are without all  $\delta$ , DMD is significantly worse than PMK-DMD, as presented in Fig. 6(b). Additionally, when the original data carries white noise with an amplitude of 0.2, PMK-DMD can resist the effect of the noise, capture the mode characteristics of the original data, and have a smaller fitting error, from Fig. 6(c). Moreover, when the modes obtained from the selected state trajectories are utilized to estimate other state trajectories, PMK-DMD demonstrates even better performance, shown in Fig. 6(d). The reasons why PMK-DMD performs better in the above scenarios are twofold. First, the number of Koopman eigenvalues PMK-DMD estimates is consistent with the number of observables

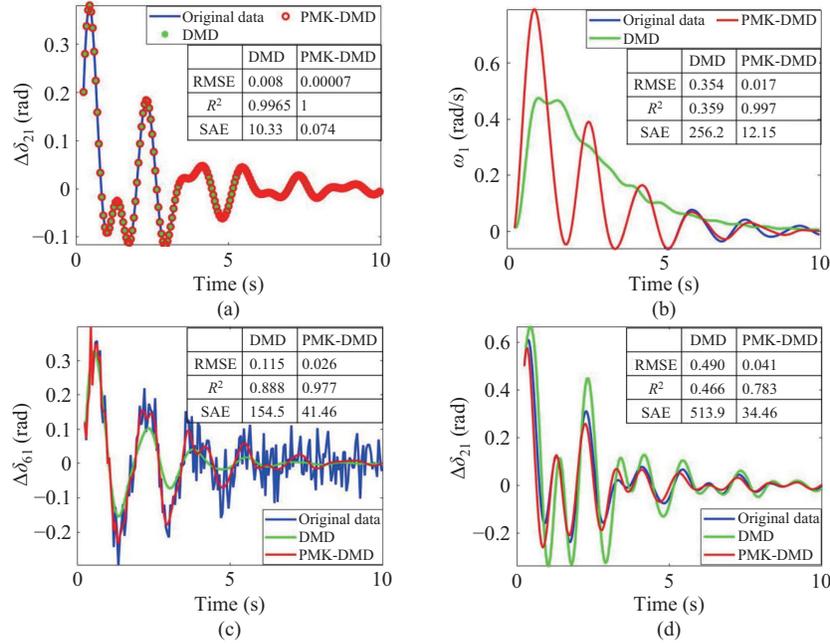


Fig. 6. Comparison of the performance of PMK-DMD and DMD. (a) Normal case. (b) Lose part of measurements. (c) Noise interference. (d) Estimate new trajectories.

$l$ , whereas that DMD estimates coincides with the number of measured states  $n$ , with  $l \gg n$ . And it is demonstrated that more Koopman eigenvalues portray the nonlinear dynamic behavior of the system more accurately [18]. Second, PMK-DMD utilizes multiple trajectories at the same time, and seeks a  $\underline{K}$  to drive the evolution of observable functions related to these trajectories. The estimated Koopman operator will potentially reflect the inherent dynamic characteristics of system (3) and is weakly dependent on the trajectories used, which is well validated by Fig. 6(d).

#### 4) Computational Savings Due to Model Reduction

When estimating the stability region for power systems, we previously proposed an Expanding Annular Domain (EAD) algorithm combined with Sum of Squares (SOS) programming [37]. However, the computational effort of this method grows dramatically with the order of the model, mainly related to the number of monomials employed [38]. In this section, we discuss the relationship between the order of the model  $n$  and the number of monomials  $M$  (reflecting the amount of computation). In the Part II of this paper, we will transform the model into a polynomial system by 5-order Taylor truncation, then a 2-degree polynomial Lyapunov function will be used to estimate the stability region, according to [38]. Thus, 6-degree SOS polynomials will be involved and according to the theory of complete square-matrix representation (CSMR) of polynomials [38], the number of monomials that make up each SOS polynomial is:

$$M = \frac{(n+3)!}{n!3!} \quad (61)$$

The result is shown in Fig. 7, and we notice that the full-order model ( $n = 19$ ) uses 27.5 times as many monomials as the 5-order model, representing a 27.5-fold difference in the amount of computation involved between the two.

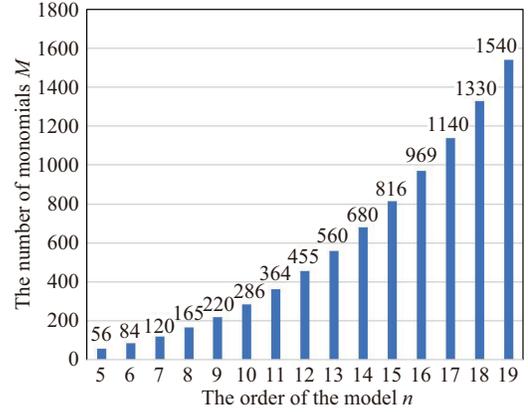


Fig. 7. Relationship between the order of the model and the number of monomials (reflecting the computational effort).

Therefore, the computational savings due to model reduction are enormous.

#### B. Case II: IEEE 16-machine-68-bus Power System

Consider a more complex and larger scale case, the IEEE 16-machine-68-bus power system, as shown in Fig. 8. See [39] for details. Here we modify model (56) by introducing the relative rotor angular velocity  $\omega_{i1} = \omega_i - \omega_1$  ( $i = 2, 3, \dots, g$ ) and the uniform damping ratio, i.e.,  $L \triangleq D_1/H_1 = D_2/H_2 = \dots = D_g/H_g$ . Then system (56) can be rewritten as:

$$\begin{aligned} \frac{d\delta_{i1}}{dt} &= \omega_{i1} \\ \frac{d\omega_{i1}}{dt} &= \frac{\omega_b}{2H_i}(P_{mi} - P_{ei}) - \frac{\omega_b}{2H_1}(P_{m1} - P_{e1}) - \lambda\omega_{i1} \end{aligned} \quad (62)$$

where  $i = 2, 3, \dots, g$ . Similarly, we define the state vector  $\mathbf{x} \in \mathbb{R}^n$  ( $n = 2g - 2$ ) as:

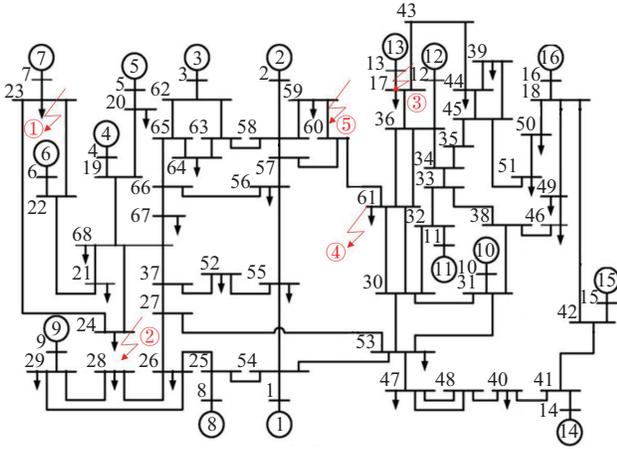


Fig. 8. One-line diagram of the IEEE 16-68 system.

$$\begin{aligned} \mathbf{x} &= [x_1, \dots, x_{g-1}, x_g, \dots, x_n] \\ &= [\Delta\delta_{21}, \dots, \Delta\delta_{g1}, \omega_{21}, \omega_{31}, \dots, \omega_{g1}] \end{aligned} \quad (63)$$

In this case,  $g = 16$ ,  $n = 30$  and  $L = 0.1257$ . As shown in Fig. 8, we set five separate faults on Bus 23, 24, 13, 61 and 59 respectively, providing five post-fault trajectories, i.e.  $s = 5$ . Similar to case I, we set  $\Delta t = 0.05s$  and  $m = 200$ . Then by the time-domain simulation, the data matrices  $\mathbf{X}$  and  $\mathbf{Y}$  can be obtained.

Here we set  $c = 21$ , and apply the PMK-DMD algorithm to the power system, which producing 21 Koopman eigenvalues. The contribution of each Koopman eigenvalue is shown in

Table IV. We find that the greatest contribution is made by  $\lambda_3^{(*)}$  rather than  $\lambda_1^{(*)}$  who has the smallest absolute real part. Then the effect of different  $L_e$  on the reduced-order model is presented in Table V. We choose several eigenvalues with relatively large contributions  $L_e = \{\lambda_1^{(*)}, \lambda_2^{(*)}, \lambda_3^{(*)}, \lambda_5^{(*)}, \lambda_7^{(*)}\}$  as the leading Koopman eigenvalues. Thus, the 30-dimensional model (62) can be reduced to a 10-dimensional model (54b). The resulting  $T$  is not convenient to show here due to its long length.

The trajectories governed by the full-order model (blue) and reduced-order model based on KOMR method (red) are shown in Fig. 9(a)–(e). The trajectories driven by the reduced-order system, such as trajectory 1 in Fig. 9(f), are projected into the full-order state space for comparison. From Fig. 9, we find that the fitting effect of the 10-order model on the trajectories of the full-order model is acceptable and  $R^2 = R(\tilde{\mathbf{x}}_{\alpha k}) = 0.6338$ . Therefore, the KOMR method can provide a satisfactory result for larger scale power systems.

 TABLE IV  
 CONTRIBUTION OF EACH KOOPMAN EIGENVALUE IN CASE II

	$\lambda_1^{(*)}$	$\lambda_2^{(*)}$	$\lambda_3^{(*)}$	$\lambda_4^{(*)}$	$\lambda_5^{(*)}$	$\lambda_6^{(*)}$	$\lambda_7^{(*)}$
$\lambda_i^{(*)}$	-0.47 $\pm 2.50i$	-0.58 $\pm 4.03i$	-0.77 $\pm 7.43i$	-1.31 $\pm 3.12i$	-1.53 $\pm 6.08i$	-1.86 $\pm 1.23i$	-2.03 $\pm 7.41i$
$C_i$	0.2477	0.1403	0.6031	0.0604	0.1467	0.0168	0.1244
	$\lambda_8^{(*)}$	$\lambda_9$	$\lambda_{10}^{(*)}$	$\lambda_{11}$	$\lambda_{12}$		
$\lambda_i^{(*)}$	-2.17 $\pm 10.97i$	-2.86	-3.85 $\pm 2.80i$	-4.19	-6.73		
$C_i$	0.0062	0.0016	0.0019	0.0036	0.0063		

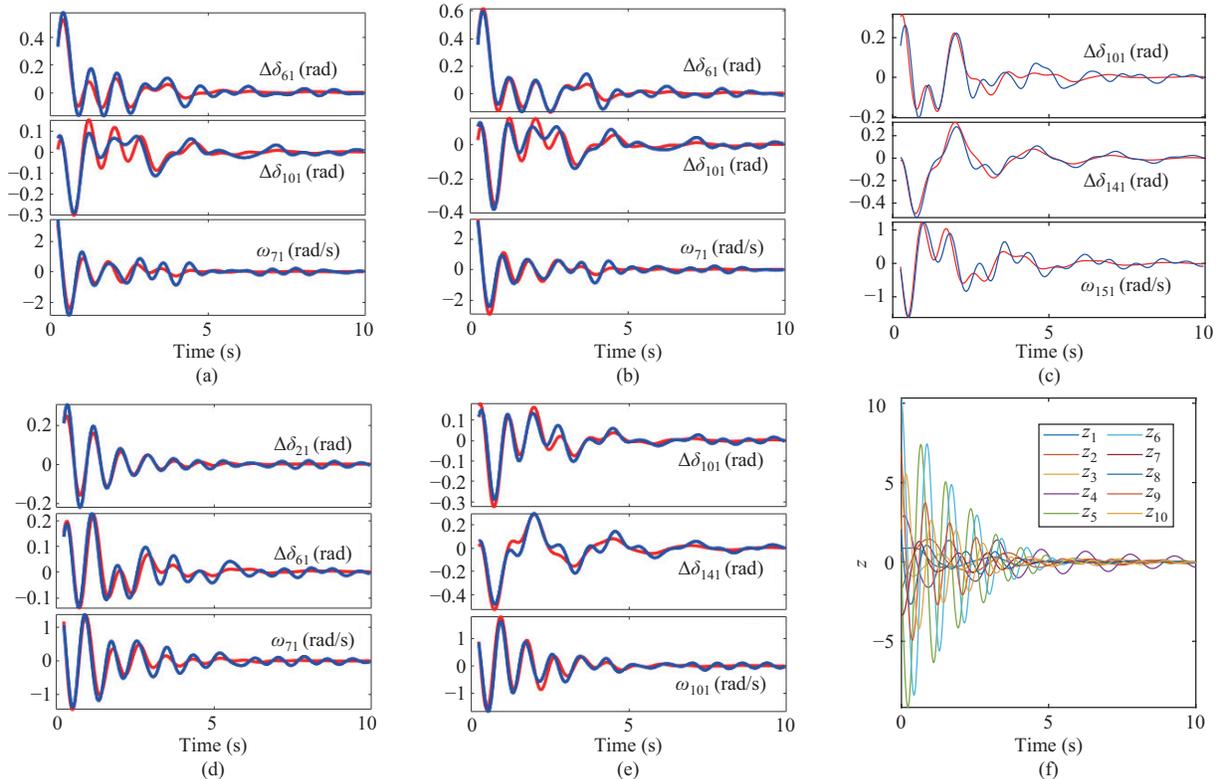


Fig. 9. Comparison of the trajectories governed by the full-order model (blue) and the reduced-order model (red) in Case II. In (a)–(e), we select 3 components of each trajectory for presentation. The trajectories driven by the reduced-order system, such as the reduced-order trajectory 1 shown in (f), are projected into the full-order state space for comparison. (a) Trajectory 1. (b) Trajectory 2. (c) Trajectory 3. (d) Trajectory 4. (e) Trajectory 5. (f) Reduced-order trajectory 1.

TABLE V  
THE EFFECT OF DIFFERENT  $l_e$  ON THE REDUCED-ORDER MODEL IN CASE II

$L_e$	$R^2$	$r$	$z = \mathbf{0}$
$\{\lambda_1^{(*)}, \lambda_3^{(*)}\}$	0.4797	4	stable
$\{\lambda_1^{(*)}, \lambda_2^{(*)}, \lambda_3^{(*)}\}$	0.5897	6	stable
$\{\lambda_1^{(*)}, \lambda_2^{(*)}, \lambda_3^{(*)}, \lambda_4^{(*)}\}$	0.5925	8	stable
$\{\lambda_1^{(*)}, \lambda_2^{(*)}, \lambda_3^{(*)}, \lambda_5^{(*)}\}$	0.6001	8	stable
$\{\lambda_1^{(*)}, \lambda_2^{(*)}, \lambda_3^{(*)}, \lambda_5^{(*)}, \lambda_7^{(*)}\}$	0.6338	10	stable
$\{\lambda_1^{(*)}, \lambda_2^{(*)}, \lambda_3^{(*)}, \lambda_4^{(*)}, \lambda_5^{(*)}\}$	0.5733	10	stable

## VI. CONCLUSION

Part I of this paper has proposed a KOMR method to produce a reduced-order model for a large-scale power system. The key to this method is to approximate the Koopman operator as accurately as possible, for which we propose PMK-DMD algorithm. PMK-DMD is computationally tractable as the size of the system expands, and it outperforms traditional DMD in various scenarios because a sufficiently large space of observable functions is systematically constructed and utilized. Then we derive the transformation matrix for a reduced-order model by selecting the leading Koopman eigenvalues and modes. On the IEEE 10–39 power system case, the obtained nonlinear reduced-order model is demonstrated to be valid and better than the traditional modal analysis method. This implies that the Koopman operator, defined in a Hilbert space spanned by observable functions, can reflect both linear and nonlinear dynamics of the underlying system in a large region. However, traditional methods like the modal analysis only focus on the linear dynamics around the equilibrium of the system.

In the future, we would like to refine our work in two directions. First, the proposed method will be applied to model reduction in large-scale power systems that include power electronic devices, such as wind farms, HVDC, and energy storage systems. The external characteristics of the grid-connected buses of these power electronic devices will be taken into account when performing system-level model reduction. As for device-level model reduction, full electromagnetic transient models will be introduced. Secondly, we will investigate the relationship between the order of the resulting model and the number of dominant oscillatory modes in the full-order system, to further reveal the intrinsic properties of transient dynamics. By analyzing several cases, we have initially found that the number of dominant oscillatory modes in most of power systems does not exceed 10 after a large perturbation, implying that the reduced-order model is at most 20-dimensional. This suggests that the dimensions of the reduced-order model are independent of the size of the power system.

## APPENDIX

### A. Results of PMK-DMD Algorithm in Case I

$n = 19, m = 120, s = 5, d = 2, q = 37, l = 741, c = 19,$   
 $r = 5. \mathbf{T} \in \mathbb{R}^{19 \times 5}$  and  $\mathbf{T} =$   
 $[-0.184, 0.035, 0.009, -0.280, 0.007; -0.198, 0.038, 0.008,$

$-0.300, 0.018; -0.263, 0.053, 0.016, -0.0006, 0.003;$   
 $-0.324, 0.074, 0.016, 0.101, 0.004; -0.249, 0.045,$   
 $0.010, 0.002, -0.017; -0.246, 0.048, 0.013, -0.021,$   
 $-0.015; -0.176, 0.035, 0.012, -0.025, 0.013; -0.227,$   
 $0.043, 0.009, 0.146, 0.023; -0.160, 0.03, 0.006, -0.021,$   
 $0.003; -0.136, -0.255, 0.302, -0.043, -0.044; 0.148, 0.228,$   
 $0.324, 0.050, 0.603; 0.162, 0.240, 0.373, 0.402, 0.656;$   
 $0.282, 0.398, 0.262, 0.040, -0.062; 0.546, 0.562, 0.246,$   
 $0.061, -0.254; 0.112, 0.373, 0.337, -0.509, -0.098;$   
 $0.164, 0.314, 0.278, -0.410, 0.014; 0.136, 0.138, 0.175,$   
 $0.343, 0.010; 0.141, 0.251, 0.434, 0.279, -0.351;$   
 $-0.012, 0.111, 0.354, -0.068, -0.016],$

where ‘;’ stands for row separator.

### B. Parameters of the Generators in Case I

$i$	1	2	3	4	5	6	7	8	9	10
$H_i$ (s)	500	30.3	35.8	28.6	26.0	34.8	26.4	24.3	34.5	42.0
$D_i$ (p.u.)	2.00	0.12	0.14	0.12	0.11	0.14	0.11	0.10	0.14	0.17
$x'_{di}$ (p.u.)	0.01	0.07	0.05	0.04	0.13	0.05	0.05	0.06	0.06	0.03

## REFERENCES

- [1] A. H. El-Abiad and K. Nagappan, “Transient stability regions of multimachine power systems,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-85, no. 2, pp. 169–179, Feb. 1966.
- [2] M. A. Pai, *Energy Function Analysis for Power System Stability*, New York: Springer Science & Business Media, 1989.
- [3] H. D. Chiang, *Direct Methods for Stability Analysis of Electric Power Systems: Theoretical Foundation, BCU Methodologies, and Applications*, Hoboken: John Wiley & Sons, 2011.
- [4] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed., Philadelphia: SIAM, 2002.
- [5] J. H. Chow, *Power System Coherency and Model Reduction*, New York: Springer, 2013.
- [6] R. Podmore, “Identification of coherent generators for dynamic equivalents,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 4, pp. 1344–1354, Jul. 1978.
- [7] S. Geeves, “A modal-coherency technique for deriving dynamic equivalents,” *IEEE Transactions on Power Systems*, vol. 3, no. 1, pp. 44–51, Feb. 1988.
- [8] J. H. Chow, J. Cullum, and R. A. Willoughby, “A sparsity-based technique for identifying slow-coherent areas in large power systems,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-103, no. 3, pp. 463–473, Mar. 1984.
- [9] R. Nath, S. S. Lamba, and K. S. Prakasa Rao, “Coherency based system decomposition into study and external areas using weak coupling,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 6, pp. 1443–1449, Jun. 1985.
- [10] F. Ma and V. Vittal, “Right-sized power system dynamic equivalents for power system operation,” *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1998–2005, Nov. 2011.
- [11] W. W. Price, E. M. Gulachenski, P. Kundur, F. J. Lange, G. C. Loehr, B. A. Roth, and R. F. Silva, “Testing of the modal dynamic equivalents technique,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 4, pp. 1366–1372, Jul. 1978.
- [12] F. Saccomanno, *Electric Power Systems: Analysis and Control*. Piscataway: Wiley-IEEE Press, 2003.
- [13] G. Troullinos, J. F. Dorsey, H. Wong, J. Myers, and S. Goodwin, “Estimating order reduction for dynamic equivalents,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 12, pp. 3475–3481, Dec. 1985.
- [14] I. J. Pérez-Arriaga, G. C. Verghese, and F. C. Schweppe, “Selective modal analysis with applications to electric power systems, part i: Heuristic introduction,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 9, pp. 3117–3125, Sep. 1982.

- [15] J. R. Winkelman, J. H. Chow, B. C. Bowler, B. Avramovic, and P. V. Kokotovic, "An analysis of interarea dynamics of multi-machine systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 2, pp. 754–763, Feb. 1981.
- [16] B. O. Koopman, "Hamiltonian systems and transformation in hilbert space," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 17, no. 5, pp. 315–318, Mar. 1931.
- [17] A. Mauroy, I. Mezić, and Y. Susuki, *The Koopman Operator in Systems and Control: Concepts, Methodologies, and Applications*, Cham: Springer, 2020.
- [18] J. Nathan Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, Philadelphia: SIAM, 2016.
- [19] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, Dec. 2009.
- [20] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the koopman operator: extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, Jun. 2015.
- [21] M. O. Williams, C. W. Rowley, and I. G. Kevrekidis, "A kernel-based method for data-driven koopman spectral analysis," *Journal of Computational Dynamics*, vol. 2, no. 247–265, Jun. 2015.
- [22] M. Netto, Y. Susuki, V. Krishnan, and Y. C. Zhang, "On analytical construction of observable functions in extended dynamic mode decomposition for nonlinear estimation and prediction," in *2021 American Control Conference (ACC)*, 2021, pp. 4190–4195.
- [23] E. Kaiser, J. Nathan Kutz, and S. L. Brunton, "Data-driven discovery of koopman eigenfunctions for control," *Machine Learning: Science and Technology*, vol. 2, no. 3, pp. 035023, Jun. 2021.
- [24] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. Nathan Kutz, "Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control," *PLoS One*, vol. 11, no. 2, pp. e0150171, Feb. 2016.
- [25] J. L. Proctor, S. L. Brunton, and J. Nathan Kutz, "Generalizing koopman theory to allow for inputs and control," *SIAM Journal on Applied Dynamical Systems*, vol. 17, no. 1, pp. 909–930, Jan. 2018.
- [26] X. Gong, X. Z. Wang, and G. Joos, "An online data-driven method for microgrid secondary voltage and frequency control with ensemble koopman modeling," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 68–81, Jan. 2023.
- [27] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge: Cambridge University Press, 2012.
- [28] Q. X. 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*, vol. 27, no. 10, pp. 103111, Jul. 2017.
- [29] C. J. Gu, "QLMOR: a projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 9, pp. 1307–1320, Sep. 2011.
- [30] P. M. Anderson and A. A. Fouad, *Power System Control and Stability*, John Wiley & Sons, 2008.
- [31] B. Schölkopf, "The kernel trick for distances," in *Proceedings of the 13th International Conference on Neural Information Processing Systems*, 2000, pp. 283–289.
- [32] G. Baudat and F. Anouar, "Kernel-based methods and function approximation," in *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222)*, 2001, pp. 1244–1249.
- [33] A. Chatterjee, "An introduction to the proper orthogonal decomposition," *Current Science*, vol. 78, no. 7, pp. 808–817, Apr. 2000.
- [34] A. Y. J. Akossou and R. Palm, "Impact of data structure on the estimators R-square and adjusted R-square in linear regression," *International Journal of Mathematics & Computation*, vol. 20, no. 3, pp. 84–93, Jan. 2013.
- [35] S. Q. Lin, W. Yao, Y. X. Xiong, Y. F. Zhao, Z. T. Shi, X. M. Ai, and J. Y. Wen, "Matpsst: a matlab/simulink-based power system simulation toolbox for research and education," *IET Generation, Transmission & Distribution*, vol. 17, no. 10, pp. 2272–2288, May 2023.
- [36] A. J. Jerri, "The shannon sampling theorem—its various extensions and applications: a tutorial review," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.
- [37] Y. Q. Lin, T. H. Wen, Y. Liu, and Q. H. Wu, "Expanding annular domain algorithm to estimate domains of attraction for power system stability analysis," *CSEE Journal of Power and Energy Systems*, vol. 10, no. 5, pp. 1925–1934, Sep. 2024.
- [38] G. Chesi, *Domain of Attraction: Analysis and Control Via SOS Programming*, London: Springer Science & Business Media, 2011.
- [39] G. Rogers, "The nature of power system oscillations," in *Power System Oscillations*, G. Rogers, Ed. Boston: Springer, 2000, pp. 7–30.



**Yuqing Lin** received a B.E. degree in Electrical Engineering from South China University of Technology, Guangzhou, China, in 2019. He is currently pursuing the Ph.D degree in South China University of Technology. His research interests include power system transient stability analysis, control of wind power generation systems, and model reduction.



**Tianhao Wen** received a B.E. degree in Electrical Engineering from Huazhong University of Science and Technology (HUST), Wuhan, China, in 2018. He obtained a Ph.D degree in South China University of Technology, Guangzhou, China in 2023. His research interests include nonlinear observers, power system transient stability analysis and control.



**Lei Chen** is currently pursuing a Ph.D. degree in Electrical Engineering at South China University of Technology, Guangzhou, China. His research interests include power system transient voltage stability analysis and power system planning and operation.



**Qing-Hua Wu** received an M.Sc. degree in Electrical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 1981, and a Ph.D. degree in Electrical Engineering from the Queen's University of Belfast (QUB), Belfast, U.K., in 1987. He worked as the Chair Professor at Liverpool University from 1995 to 2012. He is currently a Distinguished Professor and the Director of Energy Research Institute, South China University of Technology, Guangzhou, China. He is a Life Fellow of IEEE, Fellow of IET, CSEE Fellow, Fellow of AAIA, and Chartered Engineer. He has authored and coauthored more than 400 journal papers, 20 book chapters and 5 research monographs published by Springer Nature. His research interests include nonlinear adaptive control, mathematical morphology, artificial intelligence, and power system control and operation.



**Yang Liu** received a B.E. degree and a Ph.D degree in Electrical Engineering from South China University of Technology (SCUT), Guangzhou, China, in 2012 and 2017, respectively. He is currently a Lecturer in the School of Electric Power Engineering, SCUT. His research interests include the areas of power system stability analysis and control, control of wind power generation systems, and nonlinear control theory. He has authored or co-authored more than 30 peerreviewed SCI journal papers.