Data-enabled Koopman-based Load Shedding for Power System Frequency Safety

Qianni Cao and Chen Shen

Abstract—**Under-frequency load shedding (UFLS) serves as the very last resort for preventing total blackouts and cascading events. Fluctuating operating conditions and weak resilience of the future grid require UFLS adapt to various operating conditions and non-envisioned faults. This paper develops a novel data-enabled Koopman-based load shedding (KLS) to achieve the optimal one-shot load shedding for power system frequency safety. The KLS yields a network that facilitates a coordinate transformation from the delay-embedded space to a new space, wherein the dynamics can be expressed in a linear manner. The network is specifically tailored to effectively track parameter variations in the dynamic model of the power system. Linear dynamics support the development of a real-time decided load shedding strategy, while parameter tracking enables the adaptability of the KLS to non-envisioned operating conditions and faults. To address approximation inaccuracies and the discrete nature of load shedding, a safety margin tuning scheme is integrated into the KLS framework, ensuring that the system frequency trajectory remains within the safety range. Simulation results show the adaptability, prediction capability, and control effect of the proposed KLS.**

Index Terms—**Koopman-based load shedding, under-frequency load shedding, time delay, optimal emergency frequency control, parameter uncertainty.**

I. INTRODUCTION

CONTINUOUS penetration of new energy generation has posed a threat to frequency safety. In modern power systems, the declining system inertia leads to a low frequency nadir under sudden large active power deficits, which causes serious consequences.

Under-frequency load shedding (UFLS) serves as the very last resort for preventing total blackouts and cascading events. Traditional UFLS typically falls into one of two categories: offline decision-making real-time matching method [1], [2], or online decision-making real-time matching method [2], [3]. Both methods rely on the anticipation of faults, and the former considers anticipated operating conditions [4]. For the sake of clarity, we will refer to both methods as prescheduled UFLS hereafter. However, in modern power systems, the fluctuation of renewable energy generation results in less predictable operating conditions and strong uncertainties during under-frequency events. As a consequence, the traditional prescheduled UFLS becomes less effective.

Therefore, it is necessary to develop UFLS that adapts to non-envisioned operating conditions and events, and decide the one-shot optimal load shedding amount to promise the hard limits on the frequency trajectories. Specifically, the one-shot strategy is designed to prevent delayed control initiation, avoiding lower frequency nadirs and potential cascading outages from line overloads due to shifted power flows [5].

In this paper, the UFLS is designed for emergency frequency control, which is a real-time decided and data-enabled strategy. This strategy does not depend on anticipated operating conditions and faults, thus preventing the possibility of under- or over-shedding that may occur with prescheduled UFLS. With the frequency prediction capability under non-envisioned operating conditions and power imbalances, this strategy leverages the online input/output measurements to achieve safe and optimal online control with minimal one-shot load shedding, instead of shedding loads at multi-stages, thereby speeding up the recovery of system frequency.

To realize real-time decided UFLS, a challenge is to obtain load shedding amount, which ensures power system frequency safety for the current operating conditions and faults. In recent years, with the application of wide-area measurement systems in the power grid, data-driven control methods have rapidly developed, to some extent, addressing the modeling challenges and poor timeliness associated with traditional prescheduled UFLS. Data-driven control aims to extract valuable information from system responses, potentially enabling control measures to adapt to various operating conditions and disturbances.

Currently, there are no data-driven emergency control methods implemented in power systems. The emergency control methods implemented in power systems are primarily contingency-based control measures, which are event-based methods. However, some research has been conducted on data-driven emergency control for actual power system, such as the research work presented in [6]. The significance of the proposed data-driven method lies in its applicability to future scenarios with large-scale integration of renewable en-

---

Manuscript received: March 18, 2024; revised: June 4, 2024; accepted: September 10, 2024. Date of CrossCheck: September 10, 2024. Date of online publication: October 4, 2024.

This work was supported by National Key R&D Program of China (No. 2021YFB2400800).

Q. Cao and C. Shen (corresponding author) are with the State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: cqn20@mails. tsinghua. edu. cn; shenchen@mail.tsinghua.edu.cn).

DOI: 10.35833/MPCE.2024.000296

ergy sources and power electronic devices. In such scenarios, the fluctuations in renewable energy generation will result in less predictable operating conditions and strong uncertainties during under frequency events, necessitating the adoption of data-driven methods.

To ensure that control strategies effectively maintain system frequency within a safety range while optimizing cost-efficiency, a general solution is the formulation of an optimal control problem. This problem aims to minimize control costs and encompasses a variety of constraints such as operational constraints, control quantity limitations, and restrictions related to power system stability. Among these, the most complex and crucial aspect is the constraint pertaining to the dynamic characteristics of the system.

A standard solution is utilizing classical representations of system dynamics, such as the swing equation and the first-order primary frequency response (PFR) dynamics. These methods require prior knowledge of system parameters [3], [7], [8], or need to know the power deficit in the system when faults occur [6]. However, it is difficult to accurately obtain parameters such as system inertia due to the high penetration of power electronic converter-interfaced devices [9]. Moreover, the power deficit in the system cannot be measured directly in practical power system. Although data-driven methods are used to estimate these parameters [10], the accuracy of the model itself may be limited. For example, the traditional PFR model does not explicitly consider the impact of load-frequency dependence [11].

As power systems are becoming more complex and data are becoming more readily available, it is in favor to develop data-driven methods that use only input/output data measured from the unknown system [12]. Data-driven methods are suitable for applications when models are too complex for control design and thorough modelling and parameter identification are too costly. In recent years, many research works have focused on the design of data-driven load shedding strategies for power systems. Reference [13] proposes a data-driven distributional soft actor critic (DSAC) method to solve the emergency frequency control (EFC) problem. Reference [14] investigates the optimal frequency control problem using reinforcement learning with stability guarantees.

Fluctuating operating conditions and weak resilience of the future power grid require frequency control strategies to adapt to various operating conditions and non-envisioned faults. Consequently, data-driven control should possess the capability to accommodate various operating conditions and non-envisioned faults. However, for state-of-the-art data-driven UFLS, system responses are collected offline before the online control operation begins. These responses are used to estimate a model that matches the observed data in an appropriate sense. Nevertheless, offline observations are unable to cover all potential operating conditions and faults. Moreover, there is no guarantee that a system model trained for specific predetermined scenarios will generalize to the data outside of the distribution of the training set [13]-[15]. In [13], when addressing unforeseen operating conditions, the trained DSAC agent requires additional episodes to converge for strategy implementation. In [14], the reinforcement learning-based controller relies on dataset diversity to adapt to various system operating conditions and faults. In [10], a set of predefined scenarios including anticipated operating conditions and faults should be selected first, and the proposed control strategy can only be adopted in scenarios close to the predefined scenarios. Consequently, the control effect, evaluated by the frequency nadir and amount of control, degrades as the number of training trajectories decreases. Nevertheless, it is important to ensure that data-driven control strategies adapt to unanticipated operating conditions and faults. When incorporating an offline-trained system representation for online control, parameter variations should be tracked in the system model using measurements.

Given that nonlinear system dynamics render the optimal control problem intractable, it is desirable that the learned system model is linear. The Koopman-based control framework [16] has emerged as a dominant perspective in data-driven control due to its ability to provide linear representations of system dynamics in the context of black-box systems. Currently, Koopman-based methods are being widely studied in power systems [17], [18]. Reference [19] proposes fitting a global linear dynamic model of wind turbines to provide frequency regulation services. Reference [20] proposes an algorithm incorporating Koopman mode analysis for power system partitioning. Reference [21] develops a robust generalized maximum likelihood Koopman-based Kalman filter to realize state estimation of power systems.

Compared with the purely physics-informed learning method, which may utilize classical representations of power systems like the system frequency response (SFR) model and rely on the data for parameter identification, the data-driven Koopman-based method increases flexibility within a predefined model structure [16]. For instance, it allows the inclusion of terms such as polynomials or neural networks [22], [23]. Nonetheless, most Koopman-based methods rely on historical data to obtain fixed Koopman linear representations. When the operating conditions of the system change or different faults occur, the parameters in the dynamical model of the system change [24]. Although some recent research works include online algorithm of Koopman-based methods for extracting inter-area modes [25], adapting Koopman linear representations to parameter variations in the system model remains an open challenge.

Moreover, for the state-of-the-art Koopman-based methods, representation errors of Koopman operators are inevitable [16], [26]. Inaccurate estimations and predictions may degrade the quality of the obtained optimal control. In [27], the impact of representation errors of Koopman eigenpairs on the control performance of linear quadratic regulator (LQR) controllers is evaluated. The adaptive nature of closed-loop control theoretically allows one to compensate for modeling discrepancies and to account for disturbances [16]. Nevertheless, in practice, load shedding measures are rarely implemented as closed-loop strategies [28]. Hence, for Koopman-based UFLS, it is necessary to address the potential impact of representation errors on open-loop control.

In most load shedding strategy, it is often assumed that the load shedding amount at each bus can be a continuous value [13], [29]. This assumption overlooks the fact that load shedding is accomplished by shedding candidate feeders. In practical engineering, the feasible load shedding amount is restricted to discrete values, with each discrete interval representing the load associated with a feeder.

Although the optimal control problem can be formulated as an mixed-integer linear programming (MILP) problem to decide whether to shed a feeder or not [30], solving MILP is an NP-hard problem, making it difficult for online control. Therefore, [2], [31], and [32] decide the control value at a stage beforehand and then round it to the nearest discrete value. However, this leads to a discrepancy between the actual and optimal load shedding amounts, which affects the dynamic behavior of the system after control. Consequently, when designing the control strategy, it is essential to explicitly consider the potential effects of the discrepancy and ensure that the dynamic system remains within the safety range.

According to the literature review, to address the frequency stability issue in modern power systems with high penetration of renewable energy sources, we face three main challenges: ① identifying the system for optimal control under unanticipated operating conditions and power imbalance; ② analyzing the system performance when applying control policies computed from an inaccurate Koopman linear representations; and ③ designing a control strategy that works with feasible discrete load shedding values.

In response to the aforementioned research gaps, the main contributions of this paper are outlined as follows.

1) We introduce a novel data-enabled predictive control, referred to as Koopman-based load shedding (KLS), that achieves optimal one-shot load shedding amount for power system frequency safety. The proposed KLS demonstrates the adaptability to non-envisioned operating conditions under frequency events, allowing for precise load shedding strategies.

2) We investigate how approximation inaccuracies in the Koopman linear representations influence the control strategies and controlled frequency trajectories.

3) By formulating a safety margin tuning scheme within the KLS framework, we ensure that the system frequency trajectory remains within the prescribed hard limits when approximation inaccuracies exist and the feasible amount of load shedding is restricted to discrete values.

The rest of this paper is organized as follows. Section II presents the design of EFC. Section III provides error estimation and safety margin design. In Section IV, a case study is presented. Section V provides the conclusion.

II. DESIGN OF EFC

In this section, we design a deep neural network to learn a coordinate transformation from the delay-embedded measurement space into a new space where it is possible to represent the frequency dynamics linearly. An optimal control problem is then formulated to obtain the one-shot load shedding amount.

A. Identifying Koopman Linear Representation for Frequency Dynamics

Let an autonomous nonlinear dynamic system be governed by:

$$\boldsymbol{x}_{t+1}=\boldsymbol{f}(\boldsymbol{x}_t,\boldsymbol{y}_t) \tag{1}$$

where $t=1,2,...T$, and T is the prediction horizon; $\boldsymbol{x}\in\mathbb{R}^{n_x}$ is the state; $\boldsymbol{y}\in\mathbb{R}^{n_y}$ is the algebraic variable; and $\boldsymbol{f}(\cdot)$ is a nonlinear function. Considering the computational inefficiency of calculating optimal control for high-dimensional nonlinear dynamical functions, Koopman theory [33] provides a perspective that nonlinear dynamics can be represented in terms of an infinite-dimensional linear operator acting on the space of all possible measurement functions of the system. Even if $\boldsymbol{f}(\cdot)$ is unknown, it is still possible to estimate the Koopman operator using the system measurements.

However, the estimation of the Koopman operator relies exclusively on data, either numerical or experimental. In the context of power systems, it is a common practice to rely on numerical data acquired from simulations. When the dataset is collected, it is necessary to preset the operating conditions and emergency events to trigger the system dynamics. Diverse operational conditions lead to variations in the parameters of the state space model of the grid given in (1), as well as the Koopman linear representations [16]. Thus, a significant challenge in linear predictive system modeling is to adapt to different operating conditions and faults in the training set.

Although it is feasible to incorporate a wide range of operating conditions and events within the training set, it is impractical to exhaustively account for every scenario. The linear representation should possess the ability to generalize. This ensures that the linear dynamic model of the system, trained for specific predetermined scenarios, can extend its prediction capability. Such generalization allows the model to perform effectively for systems operating under non-predefined conditions and faults that are not included in the sample set.

In order to explicitly represent the variations of operating conditions and the complexity of emergency events in the state space model of (1), we utilize a vector of variables $\boldsymbol{m}$ to represent a subset of uncertain model parameters that are challenging to obtain online. To incorporate the uncertainty of these parameters into the system model, a modified model is evaluated as:

$$\begin{cases}\boldsymbol{x}_{t+1}=\boldsymbol{f}(\boldsymbol{x}_t,\boldsymbol{y}_t,\boldsymbol{m}_t)\\ \boldsymbol{m}_{t+1}=\boldsymbol{h}(\boldsymbol{x}_t,\boldsymbol{y}_t,\boldsymbol{m}_t)\end{cases} \tag{2}$$

Compared with (1), (2) provides a more general form of a deterministic power system model, which accounts for uncertainties. $\boldsymbol{x}_a^{\mathrm{T}}=[\boldsymbol{x}^{\mathrm{T}},\ \boldsymbol{m}^{\mathrm{T}}]$ can be defined as pseudo-state variables. The augmented model was first introduced in [34].

Since $\boldsymbol{m}$ is hard to measure, which constitutes hidden or latent variables that are not directly measured but are dynamically important. Thus, the challenge of adapting the linear representations to accommodate the parameter variations transforms into that of accounting for the hidden variables in the model.

Time-delay embedding provides an method to augment these hidden variables. Under certain conditions, given by Takens' embedding theorem [35], the delay-augmented state yields an attractor that is diffeomorphic to the underlying, though unmeasured, full-state attractor [36]. Here, we design a deep neural network to learn a coordinate transformation from the delay-embedded space into a new space. In the new space, it is possible to represent the dynamics in a linear form. Additionally, the network also tracks the parameter variations in the system with input/output data. The introduction of time-delay embedding enables the identification of hidden variables corresponding to unexpected operating conditions and faults. Owing to the capability of time-delay embedding to discern these hidden variables, prediction precision of system frequency is enhanced when unexpected faults occur under unexpected operating conditions. These are evidenced in the case study.

The deep neural network is illustrated in Supplementary Material A. The dataset collected for training the network is described as follows. For a given power system, a specific anticipated operating condition $\alpha \in \mathfrak{A}$ is considered, and a representative fault $\beta \in \mathfrak{B}$ is introduced, where $\mathfrak{A}$ denotes a predefined set of typical operating conditions; and $\mathfrak{B}$ denotes a predefined set of typical faults. Additionally, load shedding amount $\boldsymbol{u}=[u_1,...,u_i,...,u_I]\in \mathcal{U}$ is defined at each load node i, where $u_i\,(i=1,2,...,I)$ (in per unit, and the load level at bus i is the base value for u_i) is a uniformly distributed random number between 0 and 1; and $\mathcal{U}$ is a predefined set of load shedding trajectories in training sets. Subsequently, time-series data of the inertia center frequency of the power system (referred to as the system frequency hereafter) in (3) are collected at time points $t=1,2,...,T$, resulting in a sequence of data $\{\omega_1^{\alpha,\beta},\omega_2^{\alpha,\beta},...,\omega_T^{\alpha,\beta}\}$ and $[\boldsymbol{y}_1^{\alpha,\beta},\boldsymbol{y}_2^{\alpha,\beta},...,\boldsymbol{y}_T^{\alpha,\beta}]$. The obtained $\varOmega$ and $\boldsymbol{u}\in\mathcal{U}$ are utilized as training data for the linear prediction model.

$$\varOmega=\{\omega^{\alpha,\beta},\boldsymbol{y}^{\alpha,\beta}|\alpha\in\mathfrak{A},\beta\in\mathfrak{B},\boldsymbol{u}\in\mathcal{U}\} \tag{3}$$

Details for the network architecture and the loss function are given in Supplementary Material A. The latent extraction layers in the network are specifically designed to monitor variations in the parameters of the state-space model.

Based on Koopman theory, we assume that the frequency dynamics of the system are governed by the linear dynamic system equation represented in (4).

$$\boldsymbol{g}_{t+1}=\boldsymbol{A}\boldsymbol{g}_t+\boldsymbol{B}\boldsymbol{u}_t \tag{4}$$

$$\boldsymbol{g}_t=\begin{bmatrix}\omega_t\\ \boldsymbol{\varphi}(\omega_{t-\tau:t},\boldsymbol{y}_{t-\tau:t})\end{bmatrix} \tag{5}$$

where ω_t is the deviation of frequency from its nominal value (in per unit) at time t; $\omega_{t-\tau:t}=\{\omega_{t-\tau},\omega_{t-\tau+\Delta t},...,\omega_t\}$ and $\boldsymbol{y}_{t-\tau:t}=[y_{t-\tau},y_{t-\tau+\Delta t},...,y_t]$ are the time series of system frequency (state variable) and voltage (algebraic variable), respectively; $\boldsymbol{\varphi}$ denotes a neural network with a prescribed activation function and connectivity structure; $\boldsymbol{g}$ is a set of finite Koopman observables, which forms a subspace of the infinite dimensional Koopman observables; and $\boldsymbol{A}$ and $\boldsymbol{B}$ are the matrices in the Koopman linear representation. With the loss function and the algorithm provided in Supplementary Material A, it is feasible to approximate the parameters of $\boldsymbol{\varphi}$, along with the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

After the parameters of $\boldsymbol{\varphi}$, $\boldsymbol{A}$, and $\boldsymbol{B}$ are approximated from data, we can utilize (4) to predict the future trajectory of the system frequency variation, providing the control sequence $\boldsymbol{u}_t$, which is a time series of $\omega_{1-\tau:1}^{\alpha,\beta}$ and $\boldsymbol{y}_{1-\tau:1}^{\alpha,\beta}$. Herein, we refer to the dynamic system described by (4) as a Koopman linear system.

Remark 1 The time intervals among the time points $t=1,2,...,T$ may not be consistent with the time intervals among $t-\tau,t-\tau+\Delta t,...,t$.

The selection of time intervals is based on the observation that the system in the case study takes about 60 s to stabilize its frequency following a power deficit. As a result, the dynamics that change within 1 s are relatively slow. To reduce the number of steps in neural network prediction and thus lower training complexity, a 1 s time interval is chosen. The Δt setting of 1 ms is based on practical engineering considerations, where the sampling frequency for transient data in power grids is generally above 1000 Hz.

B. Koopman-based EFC Strategy

Combined with the Koopman model predictive control proposed in [37], the optimal load shedding amount is obtained by solving the following optimal control problem.

$$\begin{cases}\min\ \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}\\ \text{s.t.}\ \ \bar{\omega}_t\geq\omega_{\min}\quad t=1,2,...,T\\ \qquad\bar{\omega}_T\geq\omega_{\infty\min}\quad t=1,2,...,T\\ \qquad\boldsymbol{g}_{t+1}=\boldsymbol{A}\boldsymbol{g}_t+\boldsymbol{B}\boldsymbol{u}_t\quad t=1,2,...,T\end{cases} \tag{6}$$

where $\boldsymbol{R}$ is a diagonal weight matrix representing load criticality, with larger diagonal entries for more critical loads; $\bar{\omega}_t$ is the system frequency at time t predicted by Koopman linear system; $\omega_{\min}$ is the minimal allowed system frequency; and $\omega_{\infty\min}$ is the minimum allowed steady-state frequency. Here, we assume that the prediction horizon T is sufficiently long for the system frequency to reach a steady state by T. Note that the symbol $\bar{\ }$ indicates the variables with discrimination or prediction errors. The optimal control problem (6) is a quadratic programming problem with $\boldsymbol{R}$ being a positive definite matrix. This problem can be solved in polynomial time.

Remark 2 The values of $\omega_{\min}$ and $\omega_{\infty\min}$ can be adjusted based on the interplay between data-driven load shedding and the traditional UFLS. The traditional UFLS, as it initiates load shedding after a certain deviation in system frequency occurs (e.g., when the system frequency drops to 49 Hz), may lead to unexpected severe consequences due to the delayed timing of load shedding, such as greater power deficits and consequently higher load losses. If the objective of data-driven load shedding is to avoid triggering the traditional UFLS, $\omega_{\min}$ can be set to be 49 Hz. On the other hand, if the data-driven load shedding aims to fully replace the traditional UFLS and ensure that the minimum frequency of the system remains above the minimum operating frequency of synchronous generators (e.g., 47 Hz), $\omega_{\min}$ can be set to be 47 Hz. In this paper, we choose $\omega_{\min}$ as 49 Hz as an illustrative example to show the effectiveness of the proposed KLS.

In practice, continuous adjustment of load shedding amount is difficult to achieve, and it is often necessary to choose whether or not to shed a load on a particular feeder line, resulting in a series of discrete values for the actual load shedding. Let the optimal load shedding amount obtained by solving the optimal control problem (6) be denoted as $\bar{\boldsymbol{u}}_* = [\bar{u}_{1*}, ..., \bar{u}_{i*}, ..., \bar{u}_{I*}]$, the actual load shedding amount would be given as:

$$Q_d(\bar{u}_{i*}) = \begin{cases} nd & nd \le \bar{u}_{i*} < n(d+0.5) \\ n(d+1) & n(d+0.5) \le \bar{u}_{i*} < n(d+1) \end{cases} \tag{7}$$

where d is the quantization interval, which physically refers to the load shedding amount on a single feeder line; n is a positive integer; Q_d is an operator to round the optimal load shedding amount to the nearest feasible value; and $Q_d(\bar{u}_{i*})$ is the actual load shedding amount at each load node i when the discrete interval is d. Solving the optimal control problem in (6) and rounding the resulting solution as described in (7) are referred to as KLS.

Remark 3 An alternative method is to round the computed control quantity to the nearest larger value, as outlined below.

$$Q_d(\bar{u}_{i*}) = n(d+1) \quad nd < \bar{u}_{i*} \le n(d+1) \tag{8}$$

However, the strategy in (8) results in over-shedding. With the same system linear representation, the strategy presented in (8) with a larger load shedding amount compared with (7) is more likely to ensure that the system frequency does not violate safety constraints. Nevertheless, when implementing the strategy described in (7), it is also possible to ensure the safety of the system frequency by tuning a safety margin in the constraints of (6). The design of the safety margin in KLS will be presented in this paper. By employing the load shedding amount given in (7) along with the safety margin, we achieve a reduced level of load shedding amount compared with the strategy in (8).

III. ERROR ESTIMATION AND SAFETY MARGIN DESIGN

In Section II, we employ constraints in the optimal control problem to ensure that the frequency predicted by the Koopman prediction model remains above the acceptable minimum values. However, in actual power systems, the optimal load shedding amount obtained from (6) may cause the system frequency to violate the prescribed hard limits. The main reasons are as follows.

First, the training of $\boldsymbol{g}$, $\boldsymbol{A}$, and $\boldsymbol{B}$ terminates when the loss function is less than a specified tolerance. Therefore, potential inadequate training leads to representation errors in $\boldsymbol{g}$, $\boldsymbol{A}$, and $\boldsymbol{B}$. Denote the finite Koopman observables and Koopman system matrix identified from data as $\bar{\boldsymbol{A}}$, $\bar{\boldsymbol{B}}$ and $\bar{\boldsymbol{g}}$, respectively. The representation errors manifest as minor prediction errors $\bar{\boldsymbol{g}}_{t+1} - \bar{\boldsymbol{A}}\bar{\boldsymbol{g}}_t - \bar{\boldsymbol{B}}\boldsymbol{u}_t$. Second, the actual load shedding amount may deviate from the optimal load shedding amount obtained since it should be rounded to a feasible value.

Hence, to ensure that the frequency remains above the acceptable minimum values, we propose adding a safety margin to the frequency limits in the constraints of (6). Firstly, we present in Section III-A the theoretical basis for ensuring power system frequency safety with a finite safety margin. Secondly, an explicit method is introduced for calculating the safety margin in Section III-B.

A. Impact of Koopman Linear Representation Errors on Control Effects

This subsection aims to determine whether even slight deviations between the learnt and the accurate Koopman linear dynamics can compromise the desired system properties, potentially violating the imposed hard limits.

In Section II-B, the formulation of optimal control problem (6) is based on the assumption that the identification of $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{g}$ is accurate. However, in real applications, due to the training error, only $\bar{\boldsymbol{A}}$, $\bar{\boldsymbol{B}}$, and $\bar{\boldsymbol{g}}$ can be identified from data.

When solving the optimal control problem in (6), we can only employ $\bar{\boldsymbol{A}}$, $\bar{\boldsymbol{B}}$, and $\bar{\boldsymbol{g}}$ identified from data, assuming that the equality in the following equation holds.

$$\bar{\boldsymbol{g}}_{t+1} = \bar{\boldsymbol{A}}\bar{\boldsymbol{g}}_t + \bar{\boldsymbol{B}}\boldsymbol{u}_t \tag{9}$$

With $\bar{\boldsymbol{A}}$, $\bar{\boldsymbol{B}}$, and $\bar{\boldsymbol{g}}$, we utilize (9) to predict the future trajectory of Koopman observables with $\boldsymbol{x}_1$ and $\boldsymbol{u}$. We denote the trajectory of Koopman observables predicted by (9) as $[\hat{\boldsymbol{g}}_2, \hat{\boldsymbol{g}}_3, \cdots, \hat{\boldsymbol{g}}_T]$.

In order to analyze how Koopman linear representation errors influence the control strategy, we assume that $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{B}}$ can be expressed as (10) and (11), and $\boldsymbol{\Delta_A}$ and $\boldsymbol{\Delta_B}$ are bounded in terms of the induced norm as shown in (12).

$$\bar{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta_A} \tag{10}$$

$$\bar{\boldsymbol{B}} = \boldsymbol{B} + \boldsymbol{\Delta_B} \tag{11}$$

$$\begin{cases} \|\boldsymbol{\Delta_A}\| \le \varepsilon_{\boldsymbol{A}} \\ \|\boldsymbol{\Delta_B}\| \le \varepsilon_{\boldsymbol{B}} \end{cases} \tag{12}$$

where $\boldsymbol{A} \in \mathbb{R}^{p \times p}$; $\bar{\boldsymbol{A}} \in \mathbb{R}^{p \times p}$; $\boldsymbol{B} \in \mathbb{R}^{p \times q}$; $\bar{\boldsymbol{B}} \in \mathbb{R}^{p \times q}$; and $\varepsilon_{\boldsymbol{A}}$ and $\varepsilon_{\boldsymbol{B}}$ are the preset upper bounds of the representation error of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

$\bar{\boldsymbol{g}}$ can be expressed as:

$$\bar{\boldsymbol{g}}_t = \boldsymbol{g}_t + \boldsymbol{\Delta}_{\boldsymbol{g}_t} \tag{13}$$

where $\boldsymbol{g}_t \in \mathbb{R}^p$ is the discrepancy to the accurate observables $\boldsymbol{g}_t \in \mathbb{R}^p$, and we assume that $\|\boldsymbol{\Delta}_{\boldsymbol{g}_t}\| \le \varepsilon_{\boldsymbol{g}}$ is very small compared with $\|\boldsymbol{g}_t\|$ for $t = 1, 2, ..., T$.

Combining (13), dynamics in (9) can be transformed into $\boldsymbol{g}$ coordinates as:

$$\boldsymbol{g}_{t+1} = \bar{\boldsymbol{A}}\boldsymbol{g}_t + \bar{\boldsymbol{B}}\boldsymbol{u}_t + \bar{\boldsymbol{A}}\Delta_{\boldsymbol{g}_t} - \Delta_{\boldsymbol{g}_t+1} \tag{14}$$

Define (15) and $\|\gamma_{t+1}\|$ satisfies (16).

$$\gamma_{t+1} = \bar{\boldsymbol{B}}\boldsymbol{u}_t + \bar{\boldsymbol{A}}\Delta_{\boldsymbol{g}_t} - \Delta_{\boldsymbol{g}_t+1} \tag{15}$$

$$\|\gamma_{t+1}\| \le \|\bar{\boldsymbol{A}}\boldsymbol{\Delta}_{\boldsymbol{g}_t}\| + \|\boldsymbol{\Delta}_{\boldsymbol{g}_t+1}\| \tag{16}$$

where γ denotes the disturbance.

Since $\|\bar{\boldsymbol{A}}\boldsymbol{\Delta}_{\boldsymbol{g}_t}\|$ and $\|\boldsymbol{\Delta}_{\boldsymbol{g}_t+1}\|$ are small, we assume the disturbance is norm-bounded by:

$$\gamma_t=\{\gamma\in\mathbb{R}^p, \|\gamma\|\le\varepsilon_\gamma\} \tag{17}$$

For dynamics in (2) and a fixed T, we define $\mathcal{G}(\boldsymbol{x})$, $\boldsymbol{u}$, and $\boldsymbol{\gamma}$ as the stacked states, inputs, and disturbances up to time T as (18)-(20), respectively. Note that we embed x_0 as the first component of the disturbance process.

$$\mathcal{G}^{\mathrm{T}}=[\boldsymbol{g}_0^{\mathrm{T}}\quad \boldsymbol{g}_1^{\mathrm{T}}\quad\cdots\quad \boldsymbol{g}_T^{\mathrm{T}}] \tag{18}$$

$$\boldsymbol{u}^{\mathrm{T}}=[\boldsymbol{u}_0^{\mathrm{T}}\quad \boldsymbol{u}_1^{\mathrm{T}}\quad\cdots\quad \boldsymbol{u}_T^{\mathrm{T}}] \tag{19}$$

$$\boldsymbol{\gamma}^{\mathrm{T}}=[\boldsymbol{x}_0^{\mathrm{T}}\quad \gamma_0^{\mathrm{T}}\quad\dots\quad \gamma_{T-1}^{\mathrm{T}}] \tag{20}$$

Based on system level synthesis (SLS) [38], [39], we have a direct optimization over system responses $\mathcal{T}_{\boldsymbol{g}}$ and $\mathcal{T}_{\boldsymbol{u}}$, which are defined as:

$$\begin{bmatrix}\mathcal{G}^{\mathrm{T}}\\ \boldsymbol{u}_*\end{bmatrix}=\begin{bmatrix}\mathcal{T}_{\boldsymbol{g}}\\ \mathcal{T}_{\boldsymbol{u}}\end{bmatrix}\boldsymbol{\gamma} \tag{21}$$

where $\mathcal{T}_{\boldsymbol{g}}\in\mathbb{R}^{pT\times pT}$ and $\mathcal{T}_{\boldsymbol{u}}\in\mathbb{R}^{qT\times pT}$ are two block-lower triangular matrices representing system responses.

It has been proved in [38], [39] that for any $\mathcal{T}_{\boldsymbol{g}}$ and $\mathcal{T}_{\boldsymbol{u}}$ satisfying (22), the controller $\mathcal{T}_{\boldsymbol{u}}\mathcal{T}_{\boldsymbol{g}}^{-1}\in\mathbb{R}^{qT\times pT}$ achieves the desired response.

$$(\mathcal{I}-\boldsymbol{Z}\boldsymbol{\mathcal{A}}-\boldsymbol{Z}\boldsymbol{\mathcal{B}})\begin{bmatrix}\mathcal{T}_{\boldsymbol{g}}\\ \mathcal{T}_{\boldsymbol{u}}\end{bmatrix}=\mathcal{I} \tag{22}$$

where $\boldsymbol{\mathcal{A}}$ is formed by diagonally concatenating T instances of $\boldsymbol{A}$ along with a $p\times p$ zero matrix, expressed as $\boldsymbol{\mathcal{A}}=\mathrm{blkdiag}(\boldsymbol{A},\dots,\boldsymbol{A},\boldsymbol{0}_{p\times p})$; $\boldsymbol{\mathcal{B}}$ is constructed through diagonal concatenation of T instances of $\boldsymbol{B}$ and a $p\times q$ zero matrix, expressed as $\boldsymbol{\mathcal{B}}=\mathrm{blkdiag}(\boldsymbol{B},\dots,\boldsymbol{B},\boldsymbol{0}_{p\times q})$; and $\boldsymbol{Z}$ is the block-downshift operator, i.e., a matrix with the identity matrix on the first subdiagonal block and zeros elsewhere.

Based on the definition above, we further examine the effect of inaccurate Koopman linear representations $\bar{\boldsymbol{A}}$, $\bar{\boldsymbol{B}}$, and $\bar{\boldsymbol{g}}$ on the controlled system dynamics and provide an estimation of the upper error bound given as (23)-(29).

For the identified model $\bar{\boldsymbol{\mathcal{A}}}=\mathrm{blkdiag}(\bar{\boldsymbol{A}},\dots,\bar{\boldsymbol{A}},\boldsymbol{0}_{p\times p})$ and $\bar{\boldsymbol{\mathcal{B}}}=\mathrm{blkdiag}(\bar{\boldsymbol{B}},\dots,\bar{\boldsymbol{B}},\boldsymbol{0}_{p\times q})$, the block-lower triangular matrices $\{\bar{\mathcal{T}}_{\boldsymbol{g}}, \bar{\mathcal{T}}_{\boldsymbol{u}}\}$ satisfy:

$$(\mathcal{I}-\boldsymbol{Z}\bar{\boldsymbol{\mathcal{A}}}-\boldsymbol{Z}\bar{\boldsymbol{\mathcal{B}}})\begin{bmatrix}\bar{\mathcal{T}}_{\boldsymbol{g}}\\ \bar{\mathcal{T}}_{\boldsymbol{u}}\end{bmatrix}=\mathcal{I} \tag{23}$$

By rewriting (23), we can obtain:

$$\begin{cases}(\mathcal{I}-\boldsymbol{Z}\boldsymbol{\mathcal{A}}-\boldsymbol{Z}\boldsymbol{\mathcal{B}})\bar{\mathcal{T}}=\mathcal{I}-\boldsymbol{Z}\Delta\bar{\mathcal{T}}\\ \Delta=\boldsymbol{Z}[\Delta_{\mathcal{A}}\quad\Delta_{\mathcal{B}}]\\ \bar{\mathcal{T}}^{\mathrm{T}}=[\bar{\mathcal{T}}_{\boldsymbol{g}}^{\mathrm{T}}\quad\bar{\mathcal{T}}_{\boldsymbol{u}}^{\mathrm{T}}]\end{cases} \tag{24}$$

where $\Delta_{\mathcal{A}}$ and $\Delta_{\mathcal{B}}$ are the block diagonal matrices satisfying $\Delta_{\mathcal{A}}=\bar{\boldsymbol{\mathcal{A}}}-\boldsymbol{\mathcal{A}}$ and $\Delta_{\mathcal{B}}=\bar{\boldsymbol{\mathcal{B}}}-\boldsymbol{\mathcal{B}}$, respectively. The response of $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$ with the controller $\bar{\mathcal{T}}_{\boldsymbol{u}}\bar{\mathcal{T}}_{\boldsymbol{g}}^{-1}$ is given by:

$$\begin{bmatrix}\mathcal{G}^{\mathrm{T}}(\boldsymbol{x})\\ \boldsymbol{u}_*\end{bmatrix}=[\bar{\mathcal{T}}+\bar{\mathcal{T}}\Delta(\mathcal{I}-\bar{\mathcal{T}}\Delta)^{-1}\bar{\mathcal{T}}\,]\boldsymbol{\gamma} \tag{25}$$

We decompose $\mathcal{T}$, Δ, and $\boldsymbol{\gamma}$ as follows to separate the effects of the known initial condition of x_0 from the unknown future disturbances:

$$\begin{cases}\mathcal{T}=\begin{bmatrix}\bar{\mathcal{T}}_{\boldsymbol{g}}\\ \bar{\mathcal{T}}_{\boldsymbol{u}}\end{bmatrix}=[\bar{\mathcal{T}}^{0}|\,\bar{\mathcal{T}}^{\tilde{\gamma}}]\\ \boldsymbol{\gamma}=\begin{bmatrix}x_0\\ \tilde{\boldsymbol{\gamma}}\end{bmatrix}\\ \Delta=\begin{bmatrix}\Delta^{0}\\ \Delta^{\tilde{\gamma}}\end{bmatrix}\end{cases} \tag{26}$$

where $\bar{\mathcal{T}}^{0}$ is the first block column of $\mathcal{T}$; the symbol ~ indicates portions of that matrix or vector; and Δ^0 is the first block row of Δ. Then, we have:

$$\begin{bmatrix}\bar{\mathcal{G}}(\boldsymbol{x})\\ \bar{\boldsymbol{u}}_*\end{bmatrix}-\begin{bmatrix}\mathcal{G}(\boldsymbol{x})\\ \boldsymbol{u}_*\end{bmatrix}=(\bar{\mathcal{T}}-\mathcal{T})\boldsymbol{\gamma}+\bar{\mathcal{T}}^{\tilde{\gamma}}\Delta^{\tilde{\gamma}}(\mathcal{I}-\bar{\mathcal{T}}^{\tilde{\gamma}}\Delta^{\tilde{\gamma}})^{-1}\bar{\mathcal{T}}\boldsymbol{\gamma} \tag{27}$$

$$(\mathcal{I}-\boldsymbol{Z}\boldsymbol{\mathcal{A}}-\boldsymbol{Z}\boldsymbol{\mathcal{B}})(\bar{\mathcal{T}}-\mathcal{T})=\Delta\bar{\mathcal{T}}=\bar{\mathcal{T}}^{\tilde{\gamma}}\Delta^{\tilde{\gamma}} \tag{28}$$

$$(\mathcal{I}-\bar{\mathcal{T}}^{\tilde{\gamma}}\Delta^{\tilde{\gamma}})^{-1}=\sum_{k=0}^{k}(\bar{\mathcal{T}}^{\tilde{\gamma}}\Delta^{\tilde{\gamma}})^{\mathrm{T}} \tag{29}$$

$$\|\Delta^{\tilde{\gamma}}\|\le\|\Delta\|\le\|\Delta_{\mathcal{A}}\,\Delta_{\mathcal{B}}\|\le\varepsilon_{\boldsymbol{A}}+\varepsilon_{\boldsymbol{B}} \tag{30}$$

Therefore, it can be concluded that when $\varepsilon_{\boldsymbol{A}}$ and $\varepsilon_{\boldsymbol{B}}$ converge to 0, $\|[\bar{\mathcal{G}}^{\mathrm{T}}(\boldsymbol{x})\bar{\boldsymbol{u}}_*^{\mathrm{T}}]^{\mathrm{T}}-[\bar{\mathcal{G}}^{\mathrm{T}}(\boldsymbol{x})\,\boldsymbol{u}_*^{\mathrm{T}}]^{\mathrm{T}}\|$ converges to 0. In other words, the error of the open-loop dynamics is limited by the representation errors of Koopman eigenpairs. Consequently, by incorporating a finite safety margin into the frequency safety constraints, we can ensure the frequency safety despite the presence of errors in the frequency dynamic characteristics in (4).

B. Safety Margin Tuning Scheme

In Section III-A, we examine the errors in frequency trajectory prediction, and the optimal control measures that stem from inaccuracies in Koopman linear representations. Given these considerations, it is important to incorporate a finite safety margin within the frequency safety constraints in (6) to ensure that the system frequency trajectory remains within the prescribed hard limits.

Although (30) provides the upper bound of errors in frequency trajectory and control measures caused by Koopman linear representation error, the acquisition of $\varepsilon_{\boldsymbol{A}}$ and $\varepsilon_{\boldsymbol{B}}$ is challenging. Hence, this subsection proposes an analytical method for calculating the safety margin to prevent both the Koopman linear representation error and the rounded optimal control strategy from causing the system frequency to exceed the prescribed hard limits.

Proposition 1 By replacing the frequency limits in (6) with (31), it ensures that the optimal control strategy obtained from KLS, when rounded to the nearest larger value, does not violate the prescribed hard limits on the system frequency.

$$\bar{\omega}_t\ge\omega_{\min}+\sigma \tag{31}$$

$$\sigma\ge\left\|\boldsymbol{C}\sum_{k=0}^{t-1}\boldsymbol{A}^{(t-1)-k}\boldsymbol{B}\right\|\frac{d}{2}+\max_{\alpha\in\mathfrak{A},\beta\in\mathfrak{B},\boldsymbol{u}\in\mathcal{U}}\|\bar{\omega}_t^{\alpha,\beta}(\boldsymbol{u})-\omega_t^{\alpha,\beta}(\boldsymbol{u})\| \tag{32}$$

where σ is the safety margin; and $\boldsymbol{C}=[C_1, C_2, \dots, C_p]\in\mathbb{R}^p$ is an observation matrix with the first element C_1 equal to 1 and the remaining elements equal to 0.

Remark 4 The values of $\omega_t^{\alpha,\beta}(\cdot)$ can be obtained through power system simulation experiments, while the values of

$\bar{\omega}_t^{\alpha,\beta}(\cdot)$ can be calculated using the linear prediction system (4). Since α, β, and $\boldsymbol{u}$ are in the training set for (4), the last term on the right hand side of (32) is the largest prediction error of (4) on the frequency trajectories in the training set. Hence, the last term on the right hand side of (32) is acquired upon the completion of the training process for (4).

Proof The constraints in the optimal control problem (6) guarantee that the minimum value of $\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})$ is no less than $\omega_{\min}+\sigma$, and the steady-state value is no less than $\omega_{\infty\min}+\sigma$. Thus, it is crucial to find an upper bound on the difference between $\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})$ and $\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))$ in order to determine the value of σ. The estimation of this upper bound is given as:

$$
\begin{aligned}
&\|\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})-\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\| \le \|\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})-\bar{\omega}_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\|+\\
&\quad \|\bar{\omega}_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))-\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\| \le \left\|\boldsymbol{C}\sum_{k=0}^{t-1}\boldsymbol{A}^{(t-1)-k}\boldsymbol{B}\right\|\frac{d}{2}+\\
&\quad \|\bar{\omega}_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))-\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\| \le \left\|\boldsymbol{C}\sum_{k=0}^{t-1}\boldsymbol{A}^{(t-1)-k}\boldsymbol{B}\right\|\frac{d}{2}+\\
&\quad \max_{\alpha\in\mathfrak{A},\beta\in\mathfrak{B},\boldsymbol{u}\in\mathcal{U}}\|\bar{\omega}_t^{\alpha,\beta}(\boldsymbol{u})-\omega_t^{\alpha,\beta}(\boldsymbol{u})\|
\end{aligned}
\tag{33}
$$

where $\bar{\boldsymbol{u}}_*^{\alpha,\beta}$ denotes the optimal load shedding solution obtained by solving (6); $Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta})$ denotes the actual load shedding amounts at each load node i; $\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})$ denotes the predicted frequency of the linear prediction system at time t when the load shedding amount is $\bar{\boldsymbol{u}}_*^{\alpha,\beta}$; and $\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))$ and $\bar{\omega}_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))$ denote the actual and predicted system frequencies at time t, when the load shedding amount is $Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta})$, respectively. The values of $\omega_t^{\alpha,\beta}(\cdot)$ can be obtained through power system simulation, while the values of $\bar{\omega}_t^{\alpha,\beta}(\cdot)$ can be calculated using the linear prediction system (4).

In (33), the first inequality is based on the triangle inequality of the induced norm. The third inequality utilizes the largest prediction error from the training set to estimate the upper limit of the prediction error in the testing set. The proof for the second inequality in (33) is given as:

$$
\begin{aligned}
&\bar{\boldsymbol{\omega}}_t(\bar{\boldsymbol{u}})-\bar{\boldsymbol{\omega}}_t(Q(\bar{\boldsymbol{u}}))=\boldsymbol{C}\bar{\boldsymbol{A}}^t\boldsymbol{g}(\boldsymbol{x}_0)+\boldsymbol{C}\sum_{k=0}^{t-1}\bar{\boldsymbol{A}}^{(t-1)-k}\bar{\boldsymbol{B}}\boldsymbol{u}_k-\boldsymbol{C}\bar{\boldsymbol{A}}^t\boldsymbol{g}(\boldsymbol{x}_0)-\\
&\quad \boldsymbol{C}\sum_{k=0}^{t-1}\bar{\boldsymbol{A}}^{(t-1)-k}\bar{\boldsymbol{B}}Q(\boldsymbol{u}_k)=\boldsymbol{C}\sum_{k=0}^{t-1}\bar{\boldsymbol{A}}^{(t-1)-k}\bar{\boldsymbol{B}}(\boldsymbol{u}_k-Q(\boldsymbol{u}_k))\le\\
&\quad \left\|\boldsymbol{C}\sum_{k=0}^{t-1}\bar{\boldsymbol{A}}^{(t-1)-k}\bar{\boldsymbol{B}}\right\|\frac{d}{2}
\end{aligned}
\tag{34}
$$

σ is expected to ensure that the KLS guarantees the system frequency to remain within the safety range under anticipated operating conditions $\alpha\in\mathfrak{A}$ and faults $\beta\in\mathfrak{B}$.

Therefore, when the inequality (31) holds, it ensures that the frequency trajectory of the actual power system is within the safety range, under the load shedding amount $Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta})$.

Remark 5 Manual adjustment of σ is possible, which entails the following steps. First, simulating the system frequency by rounding the optimal control strategy to the nearest feasible value for each operating condition and fault in the training set. Then, instances where the system frequency fails to meet the hard limits are selected. σ is then increased and the system frequency is simulated under the new σ. This step is repeated until the system frequency complies with the limits. By contrast, our method avoids the extensive simulation required to find suitable values of σ. Instead, it relies on $\boldsymbol{A}$ and $\boldsymbol{B}$ in (4), and the prediction error already computed during the training of the encoder, thus enhancing the efficiency of safety margin design.

Here, we further discuss when the equality holds in the inequality (33).

In the deviation of the upper bound for $\|\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})-\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\|$, the first inequality in (33) and the last inequality in (34) are based on the sub-multiplicative inequality and the triangle inequality of the induced norm, respectively. For any two arrays $\boldsymbol{\mathcal{M}}$ and $\boldsymbol{\mathcal{N}}$, equality for the triangle inequality holds when the two arrays are linearly dependent, while the equality for the sub-multiplicative inequality holds if and only if each row of $\boldsymbol{\mathcal{M}}$ and each column of $\boldsymbol{\mathcal{N}}$ are linearly dependent.

$\|\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})-\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\|$ is often strictly lower than the upper bound derived in (33). The reason is the equality conditions of the triangle inequality and the sub-multiplicative inequality in (33) and (34) are hard to satisfy. The gap between $\|\bar{\omega}_t^{\alpha,\beta}(\bar{\boldsymbol{u}}_*^{\alpha,\beta})-\omega_t^{\alpha,\beta}(Q_d(\bar{\boldsymbol{u}}_*^{\alpha,\beta}))\|$ and its upper bound will be further illustrated in the simulation results in Section IV.

IV. CASE STUDY

In this section, the prediction capability and control effectiveness of KLS are illustrated using simulation data through a case study on the CloudPSS platform [40], [41].

A. Test System and Datasets

To validate the effectiveness of the proposed KLS, we conducted simulation experiments on the CEPRI-FS test system. The electromechanical transient model for this case study is available for access at [42]. The CEPRI-FS test system consists of 102 500 kV buses, and possesses a load level of 2600 MW, with installed capacities of 2400 MW and 5400 MW for renewable and conventional energy sources, respectively. The load composition consists of 40% dynamic load and 60% constant impedance loads. In CEPRI-FS test system, all synchronous machines use the 5th-order synchronous generator model, with each having different parameters for generators and governors. The full electromagnetic transient (EMT) model of the test system is built on the CloudPSS platform [43].

Based on the model of synchronous generator, system inertia is an important parameter that affects frequency safety. The influence of operating conditions on frequency dynamics can be attributed, in part, to the variations in system inertia caused by changes under operating conditions. Therefore, in this subsection, we assume a certain level of randomness in system inertia to capture unanticipated operating conditions.

The training and testing sets are generated as follows. In CEPRI-LF test system, synchronous generators in a region are represented as an equivalent single unit. Variations in system operating conditions may result in changes in the unit commitment, potentially leading to variations in the system inertia of the equivalent unit. Therefore, to represent variations under the operating conditions, we obtain a set of possible system inertia by enumerating different commitment of the generating units. When obtaining different system trajectories, different inertia values are employed. We assume that u_i at bus i is a uniformly distributed random number between 0 and 1. The fault set of the system is generated by traversing $N-1$, $N-2$, and $N-3$ generator trippings. For simulating each trajectory, each fault in the fault set is chosen at the same probability.

The training set consists of 600 frequency trajectories, while the testing set consists of 300 frequency trajectories. Each frequency trajectory has a length of 1 min. The time intervals in $t=1,2,\ldots,T$ is set to be 1 s, while in $t-\tau$, $t-\tau+\Delta t,\ldots,t$, Δt is set to be 1 ms. When generating a frequency trajectory, u_i and faults are randomly generated according to their respective distributions. When tuning the hyperparameters of the neural network, the training data are further split into a training set and a validation set in an 8:2 ratio.

Frequency trajectories from the testing set are used to assess the prediction accuracy of KLS. Furthermore, 300 test scenarios are created, with the same operating conditions and faults as in the 300 frequency trajectories. In the 300 test scenarios, the control effect of different load shedding strategies is compared.

The frequency nadir and steady-state frequency values *SSV* in various power imbalance scenarios are presented in Fig. 1, where TP and WT are short for thermal plant and wind turbine, respectively. The system inertia is based on the original manufacturer data. A nadir of 40 Hz indicates that the system frequency has reached the lower limit of the simulation software. The proposed KLS focuses on power system frequency safety. For trip events that lead to voltage or angle stability issues, we exclude them from the test scenarios.

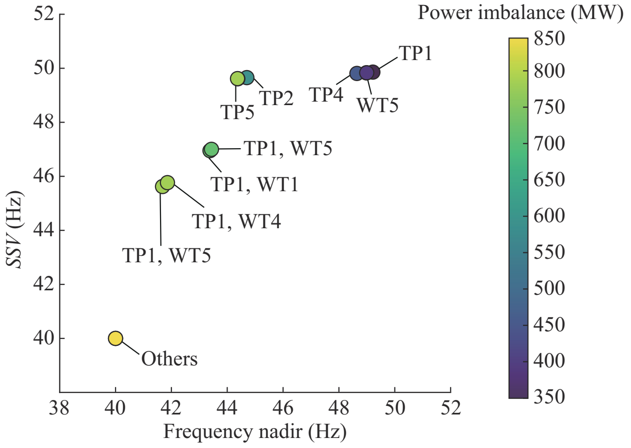


Fig. 1. Frequency nadir and *SSV* in various power imbalance scenarios.

It is important to note that system inertia is not the sole critical factor influencing frequency dynamics. Other significant elements, such as governors, also play a role. This subsection uses system inertia as an example to demonstrate that the introduction of time-delay embedding enables the identification of hidden variables, which is not directly measured, corresponding to unexpected operating conditions and faults.

B. Adaptability

Figure 2 demonstrates that KLS is capable of extracting latent variables strongly correlated with the system inertia and power imbalance from the frequency trajectory within a 300 ms time window after a fault occurs. Here, the system inertia equals 1 when the system inertia for each equivalent unit equals the corresponding original manufacturer data. Specifically, Latent 1 and Latent 2 exhibit a remarkably high correlation with the system inertia and power imbalance, respectively, providing the evidence that Koopman linear representations are capable of capturing parameter variations in the original state space system model using measurements. This result validates the adaptability of KLS to diverse operating conditions and faults. Furthermore, as the results presented in Fig. 1 are based on the training set, it confirms that KLS is able to adapt to unanticipated operating conditions and faults.

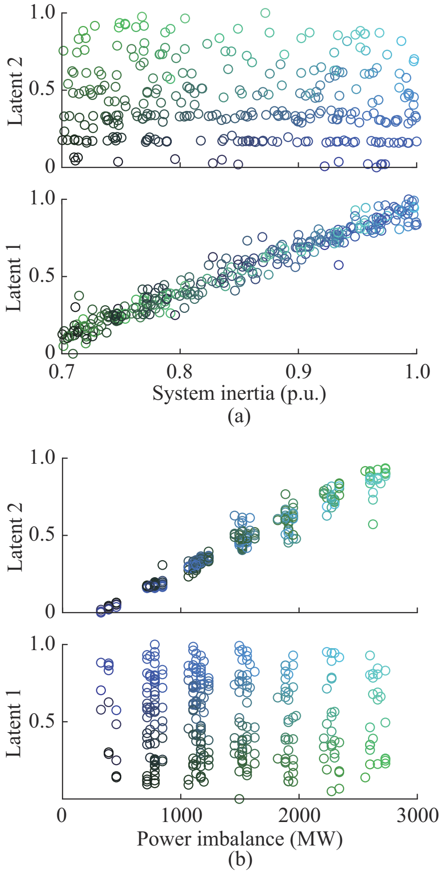


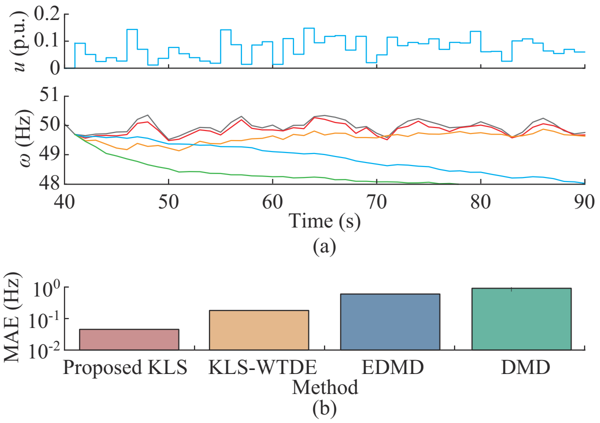
Fig. 2. Correlation between system inertia and outputs of latent extractor and between power imbalance and outputs of latent extractor. (a) System inertia. (b) Power imbalance.

C. Prediction Capability

The strong nonlinearity of dynamics makes the local linearization method hard to fit the model accurately in global horizon, and the piecewise linearization has difficulty in remaining a balance between accuracy and simplicity. Therefore, only Koopman-based methods are compared. As the benchmark of the learning algorithm, the dynamic mode decomposition (DMD) and extended dynamic mode decomposition (EDMD) are implemented with 100 radial basis functions (RBFs) as observables. To illustrate the effectiveness of incorporating time delay embedding, KLS without time delay embedding (KLS-WTDE), i.e., when $\tau=0$ in (5), is also tested as a benchmark. In subsequent discussions, we refer to KLS-WTDE, DMD, and EDMD as the state-of-the-art methods (SOTAMs) for brevity and clarity.

To illustrate the capability of the proposed KLS to learn the dynamics of system frequency from the online measurement, the frequency measurement of 1 min after a fault occurs is used to fit the linear model in (4) for the system with different system inertia, control inputs, and faults.

Figure 3(a) presents the true and predicted frequency trajectories under random control inputs. The red line represents the simulated dynamic process, while the blue, green, orange, and black lines represent the predicted results obtained from EDMD, DMD, KLS-WTDE, and KLS, respectively.



(a)

(b)

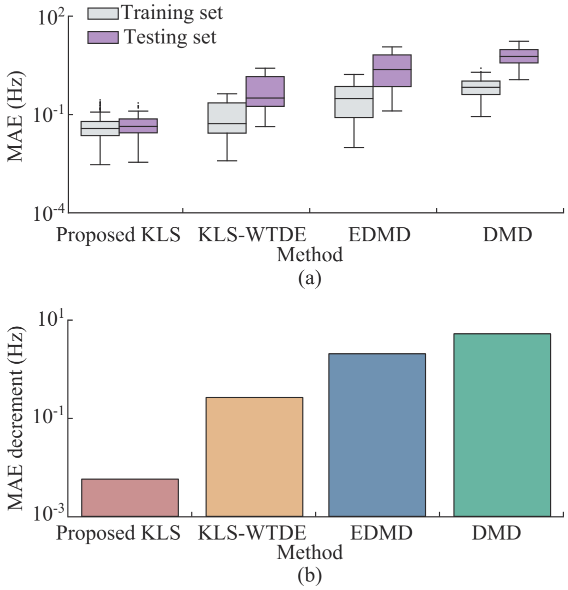
Fig. 3. Comparison between true frequency trajectory and predicted frequency trajectory during future 60 s with proposed KLS, KLS-WTDE, EDMD, and DMD. (a) True and predicted frequency trajectories under random control inputs. (b) Mean average error (MAE) of prediction errors.

Based on the frequency sequence observed within 300 ms after the generator tripping at 40 s, the proposed KLS demonstrates accurate prediction of the evolving frequency for the subsequent 60 s under different control inputs. This highlights the effectiveness of the latent extractor combined with time-delayed measurements in capturing the dynamics of the system frequency. In contrast, DMD and EDMD exhibit poorer performance, which can be attributed to their limited capability in incorporating time-delay information and harnessing the powerful non-linear representation offered by deep learning techniques.

The prediction accuracy of the proposed KLS and SOTAMs on the training and testing sets is illustrated, as shown in Fig. 4. The MAE is employed as a measure of prediction accuracy. Due to the capability of the proposed KLS to track parameter variations in the original state space system model, the prediction accuracy is consistently high on both the training and testing sets.

D. Control Effect

In this subsection, we focus on investigating the impact of prediction accuracy on control effectiveness.



(a)

(b)

Fig. 4. Prediction accuracy of proposed KLS and SOTAMs on training and testing sets. (a) MAE. (b) MAE decrement.

It is assumed that a continuous adjustment of load shedding can be achieved, and σ is set to be zero. The effectiveness of the proposed KLS and SOTAMs is evaluated by assessing the power system frequency safety after the implementation of these control strategies. Furthermore, to demonstrate the adaptability of the proposed KLS to unanticipated operating conditions and faults, all results in this subsection are computed using the test dataset. The normalized safety metric *Safety* is calculated as:

$$Safety=\alpha\frac{Nadir^{*}-Nadir_0^S}{Nadir_1^S-Nadir_0^S}+\beta\frac{SSV^{*}-SSV_0^S}{SSV_1^S-SSV_0^S} \tag{35}$$

where the superscripts $*$ and S denote the frequency nadir and steady-state frequency, respectively.

In (35), when the frequency nadir is no less than $Nadir_1^S$ and the steady-state frequency is no less than SSV_1^S, *Safety* is assigned a value of 1; when the frequency nadir $Nadir^{*}$ is less than $Nadir_0^S$ and the steady-state frequency SSV^{*} is less than SSV_0^S, *Safety* is assigned a value of 0. The weights for measuring the safety indicators of nadir and steady-state value are represented by α and β, respectively. In this paper, the values of $Nadir_1^S$ and SSV_1^S are set to be 9.0 Hz and 49.5 Hz, respectively, which are equal to $\omega_{\min}$ and $\omega_{\infty\min}$. $Nadir_0^S$ and SSV_0^S are set to be 48.5 Hz and 49.0 Hz, respectively. In Fig. 4, *Safety* is calculated by setting the values of α and β as 0.5 without loss of generality. Any α and β satisfying $\alpha+\beta=1$, $\alpha>0$, $\beta>0$ can calculate this metric, balancing the importance of not violating the minimum allowable value versus keeping the steady-state frequency within the safety range.

The average control safety metrics of proposed KLS, KLS-WTDE, EDMD, and DMD on the test scenarios are shown in Fig. 5. T-tests are conducted to compare the effectiveness of various methods. Among the 300 test scenarios, the proportions of *Safety* exceeding 0.9 for the proposed KLS, KLS-WTDE, EDMD, and DMD are 97.5%, 49.3%, 41.6%, and

4%, respectively. KLS, with its superior predictive accuracy compared with SOTAMs, enhances the power system frequency safety. However, without incorporating a safety margin into the frequency constraints of the optimal control problem (6), the average control safety metric of the proposed KLS fails to reach a value of 1.

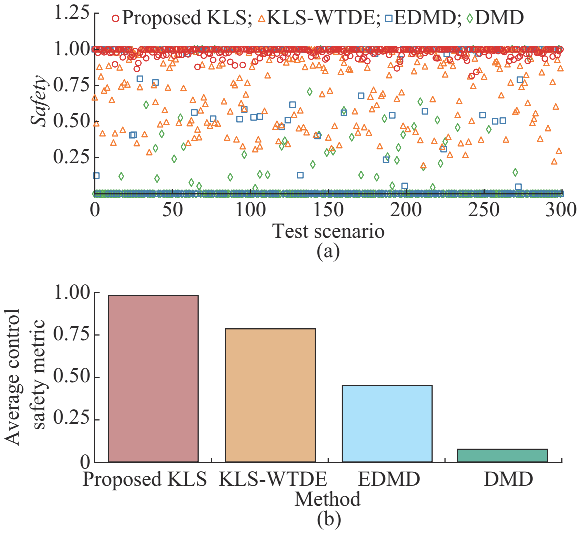


Fig. 5. *Safety* and average control safety metric of proposed KLS, KLS-WTDE, EDMD, and DMD on test scenarios. (a) *Safety*. (b) Average control safety metric.

In Fig. 6, the predictive results and control effectiveness of the proposed KLS are demonstrated, compared with SOTAMs in one of the test scenarios. It is evident that the proposed KLS offers higher frequency prediction accuracy. Although KLS-WTDE is less accurate than the proposed KLS, it outperforms both DMD and EDMD. In terms of control effectiveness, the proposed KLS achieves the minimum one-shot load shedding and frequency safety. While KLS-WTDE ensures frequency safety, its control measures are not optimal. DMD and EDMD, due to the inability to make accurate predictions, fail to guarantee the power system frequency safety.

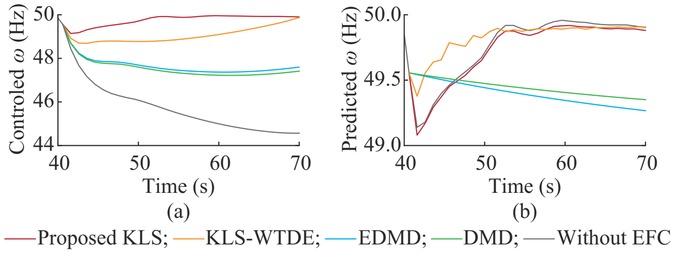


Fig. 6. Comparison between proposed KLS and SOTAMs on a scenario in testing set. (a) Control effectiveness. (b) True and predicted frequency trajectories under optimal control inputs calculated by proposed KLS.

E. Safety Margin

This subsection analyzes the control effectiveness after the introduction of a safety margin in Section III-B. The evaluation of control measures is based on two indicators: the power system frequency safety and the control cost. Increasing the amount of load shedding typically results in a higher system frequency. If the system frequency remains within the safety range (i.e., $Safety=1$), the greater the deviation of the nadir and the steady-state value of the system frequency from the specified hard limits are, the higher the associated control cost becomes. Hence, a normalized economic metric *Economy*, ranging from 0 to 1, is introduced in (36) as an indicator to measure the control cost. $Nadir_1^E=49.0$ Hz and $SSV_1^E=49.5$ Hz indicate that when the nadir is equal to 49.0 Hz and the *SSV* is equal to 49.5 Hz, the *Economy* of load shedding measures is assigned a value of 1. Similarly, $Nadir_0^E=49.5$ Hz and $SSV_0^E=50.0$ Hz indicate that when the nadir is no less than 49.5 Hz and the *SSV* is no less than 50.0 Hz, the *Economy* of load shedding measures is assigned a value of 0.

$$Economy=\min\left(\frac{Nadir^*-Nadir_0^E}{Nadir_1^E-Nadir_0^E},\frac{SSV^*-SSV_0^E}{SSV_1^E-SSV_0^E}\right) \tag{36}$$

Figure 7 presents the safety improvement when incorporating the safety margin calculated using (32) when d takes values of 10 MW, 25 MW, and 50 MW. It can be observed that the introduction of the safety margin enhances the power system frequency safety for different load levels associated with a feeder. Moreover, as d increases, the safety improvement becomes more significant.

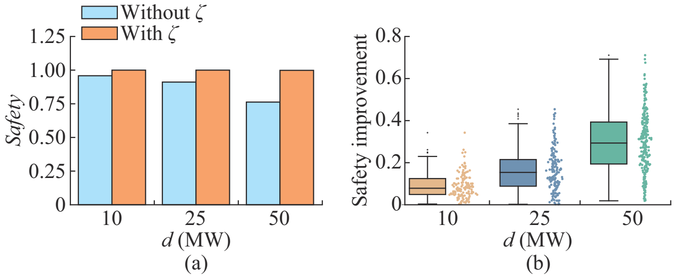


Fig. 7. Safety improvement when incorporating safety margin. (a) *Safety*. (b) Safety improvement.

For the sake of clarity, we will refer to the method in (8) as the ceiled KLS (KLS-C). The *Economy* of KLS-C and the proposed KLS is demonstrated in Fig. 8. It can be observed that by incorporating the load shedding amount in (7) along with the safety margin, the proposed KLS achieves a reduced level of load shedding amount compared with that of KLS-C. This reduction is achieved while ensuring the power system frequency safety.

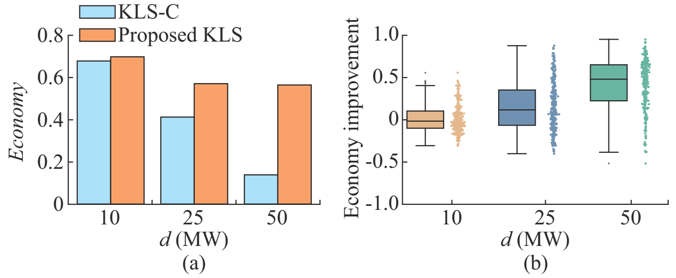


Fig. 8. *Economy* of KLS-C and proposed KLS. (a) *Economy*. (b) Economy improvement.

F. Comparison with Traditional UFLS

Frequency thresholds and load shedding proportions at each stage for traditional UFLS are determined according to [44]. The control effectiveness of traditional UFLS and the proposed KLS across 300 test scenarios is then compared. The results in two of the scenarios are shown in Fig. 9. In Fig. 9(a), it is demonstrated that in an envisioned scenario, both the proposed KLS and traditional UFLS ensure the power system frequency safety. In Fig. 9(b), in a non-envisioned scenario, traditional UFLS fails to guarantee the minimum frequency within the specified range, and additionally, overshooting occurs.

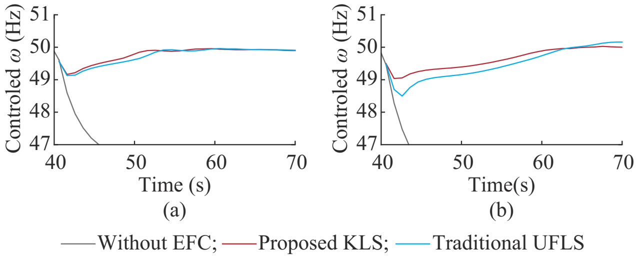


Fig. 9. Comparison between traditional UFLS and proposed KLS in an envisioned scenario and a non-envisioned scenario. (a) Envisioned scenario. (b) Non-envisioned scenario.

The comparison between traditional UFLS and proposed KLS in 300 test scenarios (non-envisioned) is presented in Fig. 10. It is evident that shedding preassigned loads alone cannot ensure frequency safety under various operating conditions and power imbalances. However, the proposed KLS demonstrates its capability to adapt to such intricate variations in the inertia and the power imbalance.

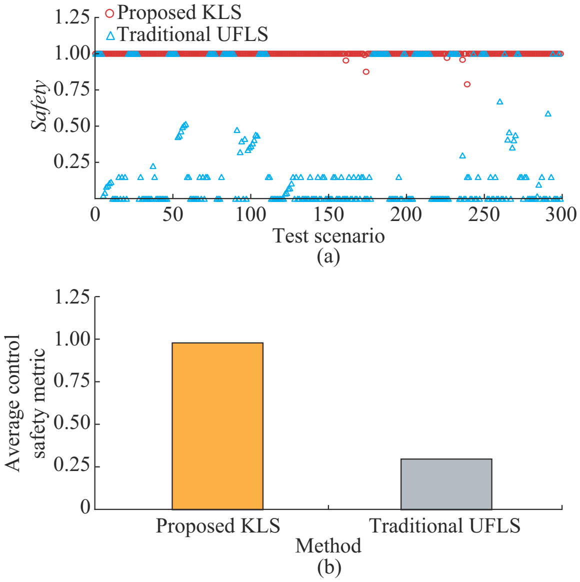


Fig. 10. Comparison between traditional UFLS and proposed KLS in 300 test scenarios. (a) *Safety*. (b) Average control safety metric.

V. CONCLUSION

In this paper, the proposed KLS, which adapts to diverse operating conditions and under frequency events, is introduced to achieve the optimal one-shot load shedding for power system frequency safety. To address approximation inaccuracies and the restriction of load shedding to discrete values, a safety margin tuning scheme is incorporated within KLS framework. Simulation results demonstrate that the proposed KLS effectively captures latent variables strongly correlated with the system inertia and power imbalance within a 300 ms time window after a fault occurs. The proposed KLS exhibits high prediction accuracy on both the training and testing sets, indicating its generalizability beyond the training set. Furthermore, the safety margin tuning scheme enhances the power system frequency safety.

REFERENCES

[1] L. Sigrist, I. Egido, and L. Rouco, "A method for the design of UFLS schemes of small isolated power systems," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 951-958, May 2012.

[2] Y.-Y. Hong and C.-Y. Hsiao, "Under-frequency load shedding in a standalone power system with wind-turbine generators using fuzzy PSO," *IEEE Transactions on Power Delivery*, vol. 37, no. 2, pp. 1140-1150, May 2022.

[3] S. S. Banijamali and T. Amraee, "Semi-adaptive setting of under frequency load shedding relays considering credible generation outage scenarios," *IEEE Transactions on Power Delivery*, vol. 34, no. 3, pp. 1098-1108, Nov. 2018.

[4] Y. Chen and Y. Xue, "On-line refreshing emergency decision table by using EEAC and expert system techniques," *Power System Technology*, vol. 20, no. 1, pp. 7-9, Dec. 1996.

[5] Y. Tofis, S. Timotheou, and E. Kyriakides, "Minimal load shedding using the swing equation," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2466-2467, Oct. 2016.

[6] S. Xu, P. Wu, B. Zhao *et al.*, "Coordinated control strategy of interconnected grid integrated with UHVDC transmission line from Hami to Zhengzhou," *Power System Technology*, vol. 39, no. 7, pp. 1174-1178, Jul. 2015.

[7] T. Amraee, M. G. Darebaghi, A. Soroudi *et al.*, "Probabilistic under frequency load shedding considering RoCoF relays of distributed generators," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3587-3598, Dec. 2017.

[8] F. Ceja-Gomez, S. S. Qadri, and F. D. Galiana, "Under-frequency load shedding via integer programming," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1387-1394, Mar. 2012.

[9] X. Chen, G. Qu, Y. Tang *et al.*, "Reinforcement learning for selective key applications in power systems: recent advances and future challenges," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2935-2958, Feb. 2022.

[10] H. Golpira, H. Bevrani, A. R. Messina *et al.*, "A data-driven under frequency load shedding scheme in power systems," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1138-1150, May 2022.

[11] Q. Shi, F. Li, and H. Cui, "Analytical method to aggregate multi-machine SFR model with applications in power system dynamic studies," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6355-6367, Apr. 2018.

[12] Z. Hou and Z. Wang, "From model-based control to data-driven control: survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3-35, Jun. 2013.

[13] J. Xie and W. Sun, "Distributional deep reinforcement learning-based emergency frequency control," *IEEE Transactions on Power Systems*, vol. 37, no. 4, pp. 2720-2730, Nov. 2021.

[14] W. Cui, Y. Jiang, and B. Zhang, "Reinforcement learning for optimal primary frequency control: a Lyapunov approach," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1676-1688, Dec. 2021.

[15] K. Champion, B. Lusch, J. N. Kutz *et al.*, "Data-driven discovery of coordinates and governing equations," *Proceedings of the National Academy of Sciences*, vol. 116, no. 45, pp. 22445-22451, Oct. 2019.

[16] S. L. Brunton, M. Budišić, E. Kaiser *et al.*, "Modern Koopman theory for dynamical systems," *SIAM Review*, vol. 64, no. 2, pp. 229-340, Feb. 2021.

[17] Z. Ma, Z. Wang, and R. Cheng, "Analytical large-signal modeling of inverter-based microgrids with Koopman operator theory for autonomous control," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, pp. 1376-1387, Sept. 2023.

[18] X. Gong, X. 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, Jul. 2022.

[19] M. Korda, Y. Susuki, and I. Mezic, “Power grid transient stabilization using Koopman model predictive control,” *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 1-10, Mar. 2018.

[20] F. Raak, Y. Susuki, and T. Hikihara, “Data-driven partitioning of power networks via koopman mode analysis,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2799-2808, Sept. 2015.

[21] M. Netto and L. Mili, “A robust data-driven Koopman Kalman filter for power systems dynamic state estimation,” *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7228-7237, Jun. 2018.

[22] Z. Ping, Z. Yin, X. Li *et al*., “Deep Koopman model predictive control for enhancing transient stability in power grids,” *International Journal of Robust and Nonlinear Control*, vol. 31, no. 6, pp. 1964-1978, Jun. 2020.

[23] X. Li, J. De La Ree, and C. Mishra, “Frequency control of decoupled synchronous machine using Koopman operator based model predictive,” in *Proceedings of 2019 IEEE PES General Meeting (PESGM)*, Atlanta, USA, Aug. 2019, pp. 1-5.

[24] M. Quade, M. Abel, J. N. Kutz *et al*., “Sparse identification of nonlinear dynamics for rapid model recovery,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 6, pp. 1-8, Jun. 2018.

[25] Y. Susuki, T. Shimomura, Y. Ota *et al*. (2022, Jan.). Online Koopman mode decomposition for power system synchrophasor data. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896322003950

[26] E. Kaiser, J. N. Kutz, and S. L. Brunton, “Data-driven discovery of Koopman eigenfunctions for control,” *Machine Learning: Science and Technology*, vol. 2, p. 35023, Jun. 2021.

[27] Q. Cao, C. Shen, and Y. Liu, “Data-driven emergency frequency control for multi-infeed hybrid AC-DC system,” *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 3880-3894, May 2023.

[28] C. Shen, Q. Cao, M. Jia *et al*., “Concepts, characteristics and prospects of application of digital twin in power system,” *Proceedings of the CSEE*, vol. 42, no. 2, pp. 487-498, Nov. 2021.

[29] A. A. Zadeh, A. Sheikhi, and W. Sun, “A novel probabilistic method for under frequency load shedding setting considering wind turbine response,” *IEEE Transactions on Power Delivery*, vol. 37, no. 4, pp. 2640-2649, Sept. 2021.

[30] D. Q. Mayne, “Model predictive control: recent developments and future promise,” *Automatica*, vol. 50, no. 12, pp. 2967-2986, Dec. 2014.

[31] S. Azuma and T. Sugie, “Optimal dynamic quantizers for discrete-valued input control,” *Automatica*, vol. 44, no. 2, pp. 396-406, Feb. 2008.

[32] R. Larson, “Optimum quantization in dynamic systems,” *IEEE Transactions on Automatic Control*, vol. 12, no. 2, pp. 162-168, Apr. 1967.

[33] B. O. Koopman, “Hamiltonian systems and transformation in hilbert space,” *Proceedings of the National Academy of Sciences*, vol. 17, no. 5, pp. 315-318, May 1931.

[34] Y. Xu, M. Netto, and L. Mili, “Propagating parameter uncertainty in power system nonlinear dynamic simulations using a Koopman operator-based surrogate model,” *IEEE Transactions on Power Systems*, vol. 37, no. 4, pp. 3157-3160, Apr. 2022.

[35] F. Takens, “Detecting strange attractors in turbulence,” *Lecture Notes in Mathematics*, vol. 898, pp. 366-381, Oct. 1981.

[36] J. Bakarji, K. Champion, J. N. Kutz *et al*., “Discovering governing equations from partial measurements with deep delay autoencoders,” *Proceedings of the Royal Society*, vol. 479, p. 20230422, Aug. 2022.

[37] M. Korda and I. Mezic, “Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control,” *Automatica*, vol. 93, pp. 149-160, Jul. 2018.

[38] S. Chen, V. M. Preciado, M. Morari *et al*., “Robust model predictive control with polytopic model uncertainty through system level synthesis,” *Automatica*, vol. 162, p. 111431, Apr. 2024.

[39] J. Anderson, J. C. Doyle, S. H. Low *et al*., “System level synthesis,” *Annual Reviews in Control*, vol. 47, pp. 364-393, Jun. 2019.

[40] Y. Song, Y. Chen, Z. Yu *et al*., “Cloudpss: a high-performance power system simulator based on cloud computing,” *Energy Reports*, vol. 6, pp. 1611-1618, Jan. 2020.

[41] CloudPSS. (2022, Jun.). MS windows NT kernel description. [Online]. Available: https://www.cloudpss.net/

[42] China Electric Power Research Institute Co., Ltd. (2024, Jan.). Electromechanical transient model – CSEE benchmark. [Online]. Available: https://github.com/lbl-hub/CSEE-Benchmark.

[43] Y. Liu and Y. Song, “Modeling and simulation of hybrid AC-DC system on a cloud computing based simulation platform-cloudpss,” in *Proceedings of 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing, China, Oct. 2018, pp. 1-6.

[44] Western Power. (2024, Jan.). Under frequency load shedding (UFLS) scheme connection guidelines. [Online]. Available: https://www.westernpower. com. au/siteassets/documents/under-frequency-load-shedding-UFLS-scheme-connection-guidelines.pdf

Chen Shen received the B.E. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1993 and 1998, respectively. From 1998 to 2001, he is a Postdoc in the Department of Electrical Engineering and Computer Science at University of Missouri Rolla, Rolla, USA. From 2001 to 2002, he was a Senior Application Developer with ISO New England Inc., Holyoke, USA, USA. He has been a Professor in the Department of Electrical Engineering at Tsinghua University since 2009. Currently, he is the Director of Research Center of Cloud Simulation and Intelligent Decision-making at Energy Internet Research Institute, Tsinghua University. His research interests include power system analysis and control, and renewable energy generation and smart grid.

Qianni Cao received the B.E. degree in electrical engineering from Wuhan University, Wuhan, China, in 2020. She is pursuing the Ph.D. degree in electrical engineering at Tsinghua University, Beijing, China. Her research interests include power system probabilistic analysis and data-driven control.