Leveraging Predictions in Power System Frequency Control: an Adaptive Approach

Wenqi Cui, Guanya Shi, Yuanyuan Shi and Baosen Zhang

Abstract—Ensuring the frequency stability of electric grids with increasing renewable resources is a key problem in power system operations. In recent years, a number of advanced controllers have been designed to optimize frequency control. These controllers, however, almost always assume that the net load in the system remains constant over a sufficiently long time. Given the intermittent and uncertain nature of renewable resources, it is becoming important to explicitly consider net load that is time-varying.

This paper proposes an adaptive approach to frequency control in power systems with significant time-varying net load. We leverage the advances in short-term load forecasting, where the net load in the system can be accurately predicted using weather and other features. We integrate these predictions into the design of adaptive controllers, which can be seamlessly combined with most existing controllers including conventional droop control and emerging neural network-based controllers. We prove that the overall control architecture achieves frequency restoration decentralizedly. Case studies verify that the proposed method improves both transient and frequency-restoration performances compared to existing approaches.

I. Introduction

Frequency control plays a fundamental role in power system operations by maintaining the real-time supply and demand balance [1]. With the increase in renewable resources, frequency stability has received significant interest because of the reduction in mechanical inertia and observed performance degradation [2]. Much of the recent works have focused on designing control algorithms that leverage the power electronic interfaces of renewable resources. Since devices like batteries and solar PVs are connected to the grid through inverters, they can quickly adjust their power to implement sophisticated control laws in response to frequency changes. For example, [3]–[9] have been proposed to design optimal nonlinear controllers that outperform linear droop control. These include neural network-based controllers in [3]–[6] to overcome the challenges in optimizing over functional space of stabilizing nonlinear controllers. In addition, steady-state frequency restoration can be realized distributedly through a method called distributed averaging-based integral (DAI) [5]–[9].

A key assumption made in all of the above works is that there exists a separation of timescales. Namely, after a disturbance occurs–e.g., a change in net load–the system will not experience other disturbances until the frequency has recovered and reached a steady state. However, because of the intermittent and uncertain nature of renewable resources, the assumption that the net load will remain constant for a sufficiently long time until the system settles down may no longer be valid. Thus, it becomes critical to explicitly model the time-varying nature of net load in frequency control.

The changes in net load are typically modeled as noises in the system and there are various ways to deal with them. If the controllers for the noiseless system (i.e., the nominal system) are exponentially stable, then bounded noises will lead to bounded state deviations [10]. Alternatively, robust controllers can be designed for the "worst-case" noise, again assuming the noise is bounded [11], [12]. However, as the intermittency and uncertainties grow in the net load, both methods tend to lead to very conservative results.

In this paper, we leverage the fact that although the net load is time-varying, it is largely predictable. In recent years, short-term net load forecasting–spanning seconds to hours– has shown impressive performance [13]–[17]. Using features such as weather data and calendar information, the future net load can be predicted with a high degree of accuracy (see, e.g. [13]–[17] and the references within). Therefore, these predictions could be used in the control problem to mitigate the impact of time-varying net load. Previously, the features have been used in end-to-end learning methods that attempt to learn a mapping directly from feature observations to control actions [18], [19]. However, it is challenging to provide performance guarantees from the learned controllers based on these end-to-end learning approaches.

We bridge the gap between load predictions and frequency control by designing controllers with provable performance guarantees under time-varying net load. Motivated by recent works on adaptive control [20]–[22], we think of load forecasting as finding a set of basis functions representing the time-varying patterns of the net load, e.g., functions of wind speed, solar radiation, temperature, past trajectories of load, etc. Then we build an adaptive control law based on (possibly unknown) combination of basis functions.

The adaptive control law can be seamlessly integrated with conventional droop control or emerging neural network-based controllers [3]–[6], and we prove the frequencies will converge to their nominal value under such compositional controller design. Case studies on IEEE 39-bus and 145-bus power systems demonstrate that the adaptive control approach improves both transient and frequency-restoration performance compared to existing approaches.

Wenqi Cui and Baosen Zhang are with the Department of Electrical and Computer Engineering, University of Washington Seattle, WA 98195, USA {wenqicui, zhangbao}@uw.edu

Guanya Shi is with the Paul G. Allen School of Computer Science and Engineering, University of Washington Seattle, guanyas@cs.washington.edu

Yuanyuan Shi is with the Department of Electrical and Computer Engineering, University of California San Diego, yyshi@eng.ucsd.edu

II. MODEL AND PROBLEM FORMULATION

A. Notations

Throughout this manuscript, vectors are denoted in lower-case bold and matrices are denoted in upper-case bold, scalars are unbolded unless otherwise specified. Vectors of all ones and zeros are denoted as $\mathbb{1}_n, \mathbb{0}_n \in \mathbb{R}^n$, respectively. For $\boldsymbol{\theta} \in \mathbb{R}^n$, θ_i represent its i-th element. Superscript $*$ indicates the equilibrium value of a variable.

B. Model

Consider an n-bus power system modeled as a connected graph $(\mathcal{V}, \mathcal{E})$. Specifically, buses are indexed by $i, j \in \mathcal{V} := [n] := \{1, \ldots, n\}$ and transmission lines are denoted by unordered pairs $\{i, j\} \in \mathcal{E} \subset \{\{i, j\} \,|\, i, j \in \mathcal{V}, i \neq j\}$. We adopt the commonly used lossless power flow model and assume that all voltages are at 1 per unit [1]. Then the state variables are phase angle $\boldsymbol{\theta} := (\theta_i, i \in [n]) \in \mathbb{R}^n$ and frequency deviation from the nominal value $\boldsymbol{\omega} := (\omega_i, i \in [n]) \in \mathbb{R}^n$. Denote M_i and D_i as the inertia constant and damping coefficient at bus i. The susceptance of the line $\{i, j\} \in \mathcal{E}$ is $B_{ij} = B_{ji} > 0$, and the susceptance $B_{kl} = B_{lk} = 0$ when $\{k, l\} \notin \mathcal{E}$. Then, the frequency dynamics is given by the swing equation written as [1],

$$\dot{\theta}_i(t) = \omega_i(t), \tag{1}$$

$$M_i \dot{\omega}_i(t) = p_i(t) - D_i \omega_i(t) - u_i(t) - \sum_{j=1}^{n} B_{ij} \sin(\theta_i(t) - \theta_j(t)),$$

where $u_i(t)$ is the control action that changes the active power injection, $p_i(t)$ is the net power injection (the negative of net load) at bus i, which is time-varying.

We consider static local feedback controllers: bus i measures its local frequency deviation ω_i and applies a static function to determine the control action u_i. We envision the control actions coming from inverter-connected resources, which can follow almost arbitrary control law because of their fast-response capabilities.

C. Optimal frequency control

The goal of frequency control is to keep the frequency around their reference value (e.g., 60Hz in the United States). Fig. 1 demonstrates typical frequency dynamics after a step increase in load, and we are interested in optimizing the transient process while realizing frequency restoration:
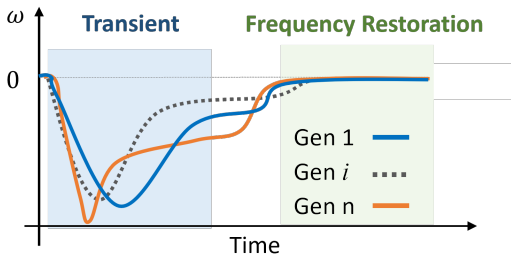

Fig. 1. Frequency dynamics after a step increase of load. We are interested in optimizing the transient process while realizing frequency restoration.

1) Frequency restoration: We seek to restore the frequency to their reference value after disturbances, i.e., $\lim_{t\to\infty} \omega_i(t) = 0, \forall i$. This goal is none trivial because of the time-varying nature of $p_i(t)$. We will later show that under some realistic models of $p_i(t)$, it is possible to realize frequency restoration using an adaptive control law.

2) Transient period: This is the period after a disturbance and before the system settles down. During the transient period, the most important goal is to reduce the maximum frequency deviation (also called the frequency nadir), represented by the infinity norm of $\omega_i(t)$ over the time horizon from 0 to T, i.e., $\|\omega_i\|_\infty := \sup_{0 \le t \le T} |\omega_i(t)|$. We also want to avoid large control efforts, computed as $\int_{t=0}^{T} C_i(u_i(t))dt$ where $C_i(\cdot)$ is a Lipschitz-continuous cost function for $i \in [n]$. The transient optimization problem up to time T is [1]

$$\min_{\boldsymbol{u}(\cdot)} \sum_{i=1}^{n} \left(\|\omega_i\|_\infty + \gamma \int_{t=0}^{T} C_i(u_i(t))dt \right), \tag{2a}$$

$$\text{s.t. dynamics in (1),} \tag{2b}$$

$$\lim_{t\to\infty} \omega_i(t) = 0, \forall i \in [n] \tag{2c}$$

where γ is a coefficient that trades off the cost of action with respect to the frequency deviation. The requirement of frequency restoration is represented as a constraint in (2c).

Even when $p_i(t)$ is time-invariant, the nonlinear dynamics (2b), the frequency restoration requirement (2c) and the nonconvex cost functions make it challenging to solve (2) using conventional optimization techniques. Recent works [3]–[6] propose to learn functions $\boldsymbol{u}(\cdot)$ by parameterizing them as neural networks and train them by minimizing the cost in (2a). These works derive structural properties of stabilizing controllers, and enforce the structures in neural networks to guarantee (2c) by design. However, these analyses rely on the assumption that $p_i(t)$ is time-invariant. The rest of the paper considers what happens when $p_i(t)$ is time-varying.

III. LEVERAGING PREDICTIONS IN POWER SYSTEMS

This section explicitly represents time-varying net power injections for frequency control. Then, we show the intuitions for developing adaptive law in controllers.

A. Modeling time-varying power injections

We model the time-varying net power injection in (1) as

$$p_i(t) = p_i^* + \boldsymbol{\phi}_i(t)^\top \boldsymbol{a}_i, \tag{3}$$

where p_i^* is the setpoint of net power injection (e.g., the economic dispatch solution) that satisfies DC power flow $p_i^* = \sum_{j=1}^{n} B_{ij} \sin(\theta_i^* - \theta_j^*) \forall i \in [n]$. The second term, $\boldsymbol{\phi}_i(t)^\top \boldsymbol{a}_i$, captures the variation in time. More specifically, we think $\boldsymbol{\phi}_i(t)$ as a vector of basis functions of features and $\boldsymbol{a}_i$ as some coefficients (possibly unknown). Note that both $\boldsymbol{\phi}$ and $\boldsymbol{a}$ are indexed by i and different buses may have a different set of basis functions and coefficients.

The model in (3) comes from the fact that the net power injections typically depend on a set of time-varying features

---

[1]The cost function can also include the summation of frequency deviation or other terms, and this does not change the analysis.

(or basis functions), including weather, temperature, time of the day, solar irradiation, wind conditions, historical load and others [13]–[17]. If $\boldsymbol{\phi}_i(t)$ includes these features, then $\boldsymbol{\phi}_i(t)^\top \boldsymbol{a}_i$ is simply a nonlinear time series-based load forecasting algorithm [13]–[17]. In addition, $\boldsymbol{\phi}_i(t)$ can include different kernel functions that have been used in accurate forecasting algorithms [20]–[22]. Note that we always include the scalar 1 as part of the feature kernels to capture the base load. The coefficients, $\boldsymbol{a}$, are the weights assigned to each of the features that make up the variations in the net load. Our approach does not assume that the value of $\boldsymbol{a}$ is known, therefore as long as the functional form of (3) is correct, we can design a controller that will adapt to the time-varying deviations. In practice, this means that we can include a large set of features to help fully capture the behavior of the net load.

B. Intuition on the adaptive control law

Since the frequency dynamics of the system in (1) depends only on the phase angle differences between different nodes, we make the change of coordinates $\delta_i(t) := \theta_i(t) - \frac{1}{n}\sum_{j=1}^n \theta_j(t)$ and write the frequency dynamics in (1) in the center-of-inertia coordinates, which are more convenient for analysis [1], [23]:

$$\dot{\delta}_i(t) = \omega_i(t) - \frac{1}{n}\sum_{j=1}^n \omega_j(t)\,, \tag{4}$$

$$M_i\dot{\omega}_i(t) = p_i^* + \boldsymbol{\phi}_i(t)^\top \boldsymbol{a}_i - D_i\omega_i(t) - u_i(t) - \sum_{j=1}^n B_{ij}\sin(\delta_{ij}(t)),$$

where $\delta_{ij}(t) := \delta_i(t) - \delta_j(t)$ and we plug in the expression of $p_i(t)$ in (3).

Intuitively, if the controller has the form $u_i(t) = \hat{u}_i(\omega_i(t)) + \boldsymbol{\phi}_i(t)^\top \boldsymbol{a}_i$, it will cancel the impact of net load variations in $\boldsymbol{\phi}_i(t)^\top \boldsymbol{a}_i$. However, this is not possible since $\boldsymbol{a}_i$ is not known. Inspired by works on adaptive control [20], [21], [24], we design an adaptation law $\boldsymbol{\phi}_i(t)^\top \hat{\boldsymbol{a}}_i(t)$ where $\hat{\boldsymbol{a}}_i(t)$ is updated according to observations in the next section.

IV. MODULAR DESIGN OF ADAPTIVE CONTROLLERS

In this section, we derive the modular controller design that consists of a base control module and an adaptation law. Then we provide conditions on the control law that guarantees frequency restoration.

A. Modular adaptive control law

We adopt a modular approach for the controller design: the control law $\hat{u}_i(\omega_i(t))$ for the system with time-invariant net load, and the adaptation law $\boldsymbol{\phi}_i(t)^\top \hat{\boldsymbol{a}}_i$ for the variation of the net load in time. Compactly, the control law is

$$u_i(t) = \underbrace{\hat{u}_i(\omega_i(t))}_{\text{Base controller}} + \underbrace{\boldsymbol{\phi}_i(t)^\top \hat{\boldsymbol{a}}_i(t)}_{\text{Adaptive Law}} \tag{5a}$$

$$\dot{\hat{\boldsymbol{a}}}_i(t) = \omega_i(t) \cdot \boldsymbol{A}_i \boldsymbol{\phi}_i(t), \tag{5b}$$

where $\boldsymbol{A}_i \succ 0$ is a tunable matrix and we will specify it later in Section IV-C.

To specify the base controller, we make the following assumption for the ranges of the angle difference.

Assumption 1. $\forall \{i, j\} \in \mathcal{E}$, the bus voltage angle difference is within $\pm \pi/2$, i.e., $\forall \{i, j\} \in \mathcal{E}$, $|\delta_i(t) - \delta_j(t)| \in [0, \pi/2)$.

Since the operation of power systems normally restricts the angle differences to be much smaller than $\pi/2$ [1], this assumption holds in most scenarios. Under Assumption 1, there is a general class of controllers that guarantees the local exponential stability of the system if $p_i(t)$ is time-invariant. We call these controllers as the Base Controller:

Base Controller. Any controller $\hat{u}_i(\omega_i(t)) \forall i \in [n]$ that is monotonically increasing and across the origin is locally exponential stable for the system if $p_i(t)$ is time-invariant [3]. For example, this class of controllers includes

(i) Droop control $\hat{u}_i(\omega_i(t)) = \varphi_i \omega_i(t)$ with $\varphi_i > 0$. This is the most popular form of controllers implemented in primary frequency control.

(ii) Neural network-based control law $\hat{u}_i(\omega_i(t)) = \hat{u}_{\boldsymbol{\psi}_i}(\omega_i(t))$, where $\hat{u}_{\boldsymbol{\psi}_i}(\cdot)$ is the function parameterized by monotone neural networks with $\boldsymbol{\psi}_i$ being trainable parameters in [3]. These monotone neural network-based controllers can offer better performances compared to linear ones [3], [5].

Remark 1. If $\boldsymbol{\phi}_i(t) = 1$, then (5b) reduces to a generic integral control where $\dot{\hat{a}}_i(t) = A_i\omega_i(t)$ and the compositional controller (5) reduces to a proportional-integral (PI) controller [6]. Hence, the adaptation law in (5b) can be viewed as an extended version of integral control that makes use of the basis functions in $\boldsymbol{\phi}_i(t)$. In the experiment, we will show that the control law in (5) can reduce the frequency oscillations compared to a standard PI controller.

B. Frequency restoration

The next theorem states conditions on the system with adaptive control law in (5) to guarantee frequency restoration.

Theorem 1. Suppose Assumption 1 holds and the control law is (5) where $\hat{u}_i(\omega_i(t))$ is a monotonically increasing function cross the origin. For the closed-loop system formed by (4) and (5), there exists a non-empty set $\mathcal{Q}_\rho$ such that if the initial states satisfying $(\boldsymbol{\delta}(0), \boldsymbol{\omega}(0), \hat{\boldsymbol{a}}(0)) \in \mathcal{Q}_\rho$, the frequency deviation $\boldsymbol{\omega}$ of each bus converges to zero, i.e., $\lim_{t\to\infty}\omega_i(t) = 0 \,\forall i \in [n]$.

This theorem covers a special case where $\boldsymbol{\phi}_i(t) = 1$, and the results align with [6] that shows the PI controller guarantees frequency restoration. Here, we want to show that if the time-varying net load has a richer set of features $\boldsymbol{\phi}$, the adaptive control law can still maintain frequency restoration. The region of attraction $\mathcal{Q}_\rho$ is characterized in Proposition 1. The subscript ρ describes the size of this region, and we will show how it can be computed. For convenience, we omit (t) for the time-varying variables in the rest of the paper.

The main tool to prove Theorem 1 is the Invariance-like Theorem [10, Theorem 8.4]. We define $\boldsymbol{x} :=$

$(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \in \mathbb{R}^{2n+\sum_{i=1}^{n} l_i}$ and write the closed-loop system formed by (4) and (5) as $\dot{\boldsymbol{x}} = \boldsymbol{f}(t, \boldsymbol{x})$. Define the set for the states satisfying Asusmption 1 as $\mathcal{X} := \{(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \,|\, |\delta_i - \delta_j| \in [0, \pi/2] \,\forall \{i, j\} \in \mathcal{E}\}$, which contains $\boldsymbol{x}^* = (\boldsymbol{\delta}^*, \mathbb{0}_n, \boldsymbol{a})$. The Invariance-like Theorem (in the notation of this paper) is:

Proposition 1 (Invariance-like Theorem [10]). *Consider the system $\dot{\boldsymbol{x}} = \boldsymbol{f}(t, \boldsymbol{x})$ where $\boldsymbol{x} \in \mathbb{R}^{2n+\sum_{i=1}^{n} l_i}$. Let $V : \mathcal{X} \to R$ be a continuously differentiable functions such that*

$$W_1(\boldsymbol{x}) \leq V(\boldsymbol{x}) \leq W_2(\boldsymbol{x}) \tag{6}$$

$$\dot{V}(\boldsymbol{x}) \leq -W_0(\boldsymbol{x}) \tag{7}$$

hold on $\mathcal{X}$, where $W_1(\boldsymbol{x})$ and $W_2(\boldsymbol{x})$ are continuous positive definite functions and $W_0(\boldsymbol{x})$ is a continuous positive semidefinite function on $\mathcal{X}$. Choose $r > 0$ such that $\mathcal{B}_r := \{\boldsymbol{x} \in \mathbb{R}^{2n+\sum_{i=1}^{n} l_i} \,|\, \|\boldsymbol{x}\| < r\} \subset \mathcal{X}$ and let $\rho < \min_{\|x\|=r} W_1(\boldsymbol{x})$. Then, all solutions of the system with an initial condition $\boldsymbol{x}(0) \in \mathcal{Q}_\rho := \{x \in \mathcal{B}_r \,|\, W_2(\boldsymbol{x}) \leq \rho\}$ are bounded and satisfy

$$\lim_{t \to \infty} W_0(\boldsymbol{x}(t)) \to 0 \,.$$

To use the above result, we construct a function $V(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}})$ that is bounded by continuous positive definite functions and satisfying $\dot{V}(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \leq -\sum_{i=1}^{n} D_i \omega_i^2$. Then the convergence of $-\sum_{i=1}^{n} D_i \omega_i^2$ to zero follows directly from Proposition 1. The rest of this section outlines the proof of Theorem 1.

We use the following energy function

$$V(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) = \hat{V}(\boldsymbol{\delta}, \boldsymbol{\omega}) + \frac{1}{2} \sum_{i=1}^{n} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i)^\top \boldsymbol{A}_i^{-1} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i) \tag{8}$$

where $\hat{V}(\boldsymbol{\delta}, \boldsymbol{\omega}) =: \frac{1}{2} \sum_{i=1}^{n} M_i \omega_i^2 + W_\mathrm{p}(\boldsymbol{\delta})$ with

$$W_\mathrm{p}(\boldsymbol{\delta}) := -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} \left(\cos(\delta_{ij}) - \cos(\delta_{ij}^*)\right)$$
$$- \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} \sin(\delta_{ij}^*)(\delta_i - \delta_i^*) \,.$$

The next Lemma shows how V can be bounded:

Lemma 1 (Bounds of the energy function). *For all $(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \in \mathcal{X}$, the energy function $V(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}})$ in (8) satisfies*

$$W_1(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \leq V(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \leq W_2(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}})$$

with

$$W_1(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) := \gamma_1(\|\boldsymbol{\delta} - \boldsymbol{\delta}^*\|_2^2 + \|\boldsymbol{\omega}\|_2^2 + \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_2^2) \,,$$
$$W_2(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) := \gamma_2(\|\boldsymbol{\delta} - \boldsymbol{\delta}^*\|_2^2 + \|\boldsymbol{\omega}\|_2^2 + \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_2^2) \,,$$

for some constants $\gamma_1 > 0$ and $\gamma_2 > 0$.

Proof. We start by bounding the term $W_\mathrm{p}(\boldsymbol{\delta})$. Define $S(\boldsymbol{\delta}) := -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} B_{ij} \cos(\delta_{ij})$, then $\nabla S(\boldsymbol{\delta}) = \left([\nabla S(\boldsymbol{\delta})]_i := \sum_{j=1}^{n} B_{ij} \sin(\delta_{ij}), i \in [n]\right) \in \mathbb{R}^n$

and $\nabla^2 S(\boldsymbol{\delta}) := \left([\nabla^2 S(\boldsymbol{\delta})]_{i,j} = \frac{\partial^2 S(\boldsymbol{\delta})}{\partial \delta_i \delta_j}, i, j \in [n]\right) \in \mathbb{R}^{n \times n}$ where

$$\frac{\partial^2 S(\boldsymbol{\delta})}{\partial \delta_i \delta_j} = \begin{cases} -B_{ij} \cos(\delta_{ij}) & \text{if } i \neq j \\ \sum_{j'=1, j' \neq i}^{n} B_{ij'} \cos(\delta_{ij'}) & \text{if } i = j \end{cases}, \forall i, j \in [n] \,.$$

For $(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \in \mathcal{X}$, $\nabla^2 S(\boldsymbol{\delta})$ is the Laplacian matrix that satisfying $\nabla^2 S(\boldsymbol{\delta}) \succ 0$ and thus $S(\boldsymbol{\delta})$ is convex. Note that

$$W_\mathrm{p}(\boldsymbol{\delta}) = S(\boldsymbol{\delta}) - S(\boldsymbol{\delta}^*) - \nabla S(\boldsymbol{\delta}^*)^\top (\boldsymbol{\delta} - \boldsymbol{\delta}^*) \tag{9}$$

is the Bregman distance of the convex function $S(\boldsymbol{\delta})$ at the point $\boldsymbol{\delta}^*$. According to [9, Lemma 4], (9) can be bounded by $\beta_1 \|\boldsymbol{\delta} - \boldsymbol{\delta}^*\|_2^2 \leq W_\mathrm{p}(\boldsymbol{\delta}) \leq \beta_2 \|\boldsymbol{\delta} - \boldsymbol{\delta}^*\|_2^2$ for $\boldsymbol{\delta}$ satisfying Assumption 1 and some constants $\beta_1, \beta_2 > 0$.

Using the Rayleigh-Ritz theorem [25], the term $\frac{1}{2} \sum_{i=1}^{n} M_i \omega_i^2$ is lower bounded by $\frac{1}{2} (\min_i M_i) \|\boldsymbol{\omega}\|_2^2$ and is upper bounded by $\frac{1}{2} (\max_i M_i) \|\boldsymbol{\omega}\|_2^2$. The term $\frac{1}{2} \sum_{i=1}^{n} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i)^\top \boldsymbol{A}_i^{-1} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i)$ is lower bounded by $\frac{1}{2} (\min_i \lambda_{min}(\boldsymbol{A}_i)) \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_2^2$ and upper bounded by $\frac{1}{2} (\max_i \lambda_{max}(\boldsymbol{A}_i)) \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_2^2$. Therefore, we can bound the energy function in (8) with

$$\gamma_1 := \frac{1}{2} \min \left(\min_i M_i, 2\beta_1, \min_i \lambda_{min}(\boldsymbol{A}_i)\right) > 0 \,,$$
$$\gamma_2 := \frac{1}{2} \max \left(\max_i M_i, 2\beta_2, \max_i \lambda_{max}(\boldsymbol{A}_i)\right) > 0 \,.$$
$\qquad \square$

The next lemma shows an important inequality associated with the time derivative of the energy function and the positive definite function $W_0(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) := \sum_{i=1}^{n} D_i \omega_i^2$.

Lemma 2 (Time derivative of the energy function). *The time derivative of the energy function satify*

$$\dot{V}(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \leq -W_0(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) := -\sum_{i=1}^{n} D_i \omega_i^2 \,.$$

Proof. We start by computing the partial derivatives of $V(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}})$ with respect to each state. From (8) and (9), we have

$$\nabla_{\boldsymbol{\delta}} V = \nabla S(\boldsymbol{\delta}) - \nabla S(\boldsymbol{\delta}^*) \,,$$
$$\nabla_{\boldsymbol{\omega}} V = \boldsymbol{M} \boldsymbol{\omega} \,, \tag{10}$$
$$\nabla_{\hat{\boldsymbol{a}}_i} V = \boldsymbol{A}_i^{-1} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i) \,\forall i \in [n] \,.$$

The time derivative of $V(\boldsymbol{\delta}, \boldsymbol{\omega})$ defined in (8) is given by

$$\dot{V}(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}})$$
$$= (\nabla_{\boldsymbol{\delta}} V)^\top \dot{\boldsymbol{\delta}} + (\nabla_{\boldsymbol{\omega}} V)^\top \dot{\boldsymbol{\omega}} + \sum_{i=1}^{n} (\nabla_{\hat{\boldsymbol{a}}_i} V)^\top \dot{\hat{\boldsymbol{a}}}_i$$
$$\overset{\textcircled{1}}{=} (\nabla S(\boldsymbol{\delta}) - \nabla S(\boldsymbol{\delta}^*))^\top \left(\boldsymbol{\omega} - \mathbb{1}_n \frac{\mathbb{1}_n^\top \boldsymbol{\omega}}{n}\right) \tag{11}$$
$$+ \sum_{i=1}^{n} \omega_i \left(p_i^* - \boldsymbol{\phi}_i(t)^\top \tilde{\boldsymbol{a}}_i - D_i \omega_i - \hat{u}_i(\omega_i) - [\nabla S(\boldsymbol{\delta})]_i\right)$$
$$- \sum_{i=1}^{n} \omega_i \underbrace{(p_i^* - [\nabla S(\boldsymbol{\delta}^*)]_i)}_{=0} + \sum_{i=1}^{n} \omega_i \cdot \tilde{\boldsymbol{a}}_i^\top \boldsymbol{A}_i^{-1} \boldsymbol{A}_i \boldsymbol{\phi}_i(t)$$

$$= -\sum_{i=1}^{n} D_i \omega_i^2 - \sum_{i=1}^{n} \omega_i \hat{u}_i(\omega_i) - (\nabla S(\boldsymbol{\delta}) - \nabla S(\boldsymbol{\delta}^*))^\top \left(\mathbb{1}_n \frac{\mathbb{1}_n^\top \boldsymbol{\omega}}{n} \right)$$

$$\overset{\text{②}}{=} -\sum_{i=1}^{n} D_i \omega_i^2 - \sum_{i=1}^{n} \omega_i \hat{u}_i(\omega_i)$$

$$\overset{\text{③}}{\leq} -\sum_{i=1}^{n} D_i \omega_i^2,$$

where $\tilde{\boldsymbol{a}}_i := \hat{\boldsymbol{a}}_i - \boldsymbol{a}_i$. The equality ① plugs in (10) and adds the extra term using $p_i^* = \sum_{j=1}^{n} B_{ij} \sin(\delta_i^* - \delta_j^*) \equiv [\nabla S(\boldsymbol{\delta}^*)]_i \forall i \in [n]$. Since $B_{ij} \sin(\delta_i - \delta_j) = -B_{ji} \sin(\delta_j - \delta_i) \forall i, j$, we have $\sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} \sin(\delta_i - \delta_j) = 0$. Then the equality ② follows from the property that $\mathbb{1}_n^\top \nabla S(\boldsymbol{\delta}) = \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} \sin(\delta_i - \delta_j) = 0$ for all $\boldsymbol{\delta} \in \mathbb{R}^n$. Because $\hat{u}_i(\omega_i)$ is monotonically increasing with ω_i, we have $\hat{u}_i(\omega_i)\omega_i \geq 0 \forall i$ and thus the inequality ③ holds. $\square$

By Proposition 1, the bound of the energy function in Lemma 1 and $\dot{V}_1(\boldsymbol{\delta}, \boldsymbol{\omega}, \hat{\boldsymbol{a}}) \leq -\sum_{i=1}^{n} D_i \omega_i^2$ in Lemma 2 imply that $\sum_{i=1}^{n} D_i \omega_i^2 \to 0$ as $t \to \infty$ for initial states satisfying $(\boldsymbol{\delta}(0), \boldsymbol{\omega}(0), \hat{\boldsymbol{a}}(0)) \in \mathcal{Q}_\rho$. Since $\sum_{i=1}^{n} D_i \omega_i^2 = 0$ if and only if $\omega_i = 0 \forall i \in [n]$, we complete the proof of Theorem 1.

C. Training

The matrix $\boldsymbol{A}_i \succ 0$ and the base control law $\hat{u}_i(\omega_i)$ should be chosen to optimize the behavior of the system. For $\boldsymbol{A}_i$'s, similar to how LQR controllers are tuned, we parameterize them as diagonal matrices and optimize these entries. For $\hat{u}_i(\cdot)$, we adopt the learning-based method to parameterize $\hat{u}_i(\cdot)$ as a monotone neural network. Note that this construction of $\hat{u}_i(\cdot)$ and $\boldsymbol{A}_i$ satisfies the conditions in Theorem 1, then the frequency restoration always holds by design. We optimize $\hat{u}_i(\cdot)$ and $\boldsymbol{A}_i$ by training it with the loss function the same as the optimization objective (2a). Detailed formulation of the monotone neural network and the training framework can be found in our previous work [3].

V. Case Study

Case studies are conducted on the IEEE New England 10-machine 39-bus (NE39) and the IEEE 50-machine 145-bus test systems [26], [27]. We use TensorFlow 2.0 framework in Google Colab with a single Nvidia Tesla T4 GPU with 16GB memory to train $\boldsymbol{A}_i$ and $\hat{u}_i(\cdot)$ for all $i \in [n]$.

A. IEEE 39-bus test system.

We conduct experiments on the IEEE-NE39 power network with parameters given in [3], [26]. For training and testing the controller, we generate 300 trajectories by randomly picking at most three buses to experience a step load change uniformly distributed in uniform$[-1, 1]$ p.u., where 1p.u.=100 MW is the base unit of power. We discretize time and use k to index the time step, with an interval of 0.01s between each step. The number of time steps in a trajectory in the training set is $K = 400$, i.e., each trajectory spans the time horizon of 4s. For simplicity and illustrative purposes, we use sinusoidal signals to represent time-varying oscillations of the net load and demonstrate

the performance of the proposed method. More advanced load forecasting algorithms (e.g., see [14]–[16] and references within) can also be used without changes to the algorithm. The basis $\phi_i(k) = (sin(\eta_i^1 k), sin(\eta_i^2 k), 1)$, where $\eta_i^1, \eta_i^2 \sim$ uniform$[0.005\pi, 0.02\pi]$. The coefficients $[\boldsymbol{a}_i]_j \sim$ uniform$[0.1, 0.2]$ for each item j in $\boldsymbol{a}$. The training epoch number is 600. The loss function is (2a) where $\gamma = 0.1$, $C_i(u_i) = c_i(u_i)^2$ with $c_i \sim$ uniform$[0.025, 0.075]$.

We compare the performance of 1) NN-Adaptive: the proposed adaptive controller (5) where $\hat{u}_i(\cdot)$ is parameterized by the monotone neural network with 20 neurons in the hidden layer, 2) NN-Integral: a standard integral controller (see Remark 1) is used without considering the other features in ϕ, and $\hat{u}_i(\cdot)$ is the same as NN-Adaptive - parameterized by the monotone neural network with 20 neurons in the hidden layer; 3) Linear Droop: Conventional linear droop control with $\hat{u}_i(\omega_i) = \varphi_i \omega_i$ for all $i \in [n]$. The values of φ_i are optimized through learning.

The average batch loss during epochs of training is shown in Fig. 2. All three methods converge, with the linear droop control having the highest loss. Fig. 3 compares the transient and frequency-restoration cost on the test set. Specifically, the transient cost is (2a) for $T = 4$s. The frequency-restoration cost is the average frequency deviation in 10s-15s after the step change, where the dynamics approximately enter the steady state after $t = 10$s as we will show later in the simulation. The transient cost of NN-Adaptive is 1.5% lower than NN-Integral and 67% lower than Linear Droop, respectively. The frequency-restoration cost of NN-Adaptive is 60% lower than NN-Integral and 94% lower than Linear Droop, respectively. Hence, linear droop control has significantly higher transient and frequency-restoration costs than others.
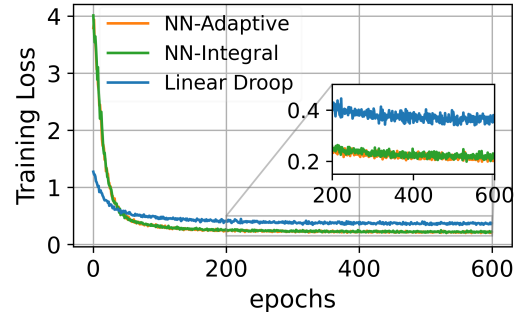

Fig. 2. Average batch loss along epochs for IEEE-NE39 test case. All converge, with NN-Adaptive and NN-Integral achieving much lower loss than Linear Droop. The training loss only reflects transient cost but cannot reflect the performance of frequency restoration.

It is noteworthy that the training loss (in Fig. 2) as well as the transient cost on the test set (in Fig. 3) for NN-Adaptive and NN-Integral is similar. But the frequency-restoration cost of NN-Adaptive on the test set is *significantly lower* than that of NN-Integral. The reason is that training only covers a finite length of trajectories where the transient cost is dominant, and it is difficult to quantify how long the trajectory should be to improve frequency-restoration performances.
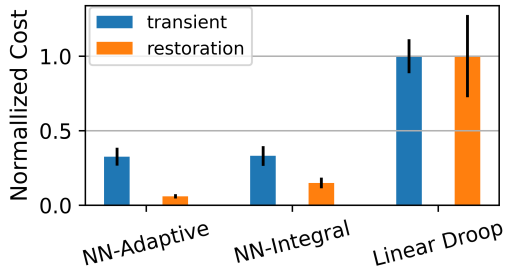
Fig. 3. The average transient cost and frequency-restoration cost with error bar on the randomly generated test set with size 300. NN-Adaptive achieves the lowest transient and frequency-restoration cost.

Therefore, the adaptive control approach provides an easy and efficient way to enforce frequency restoration by design, which also does not degrade transient performances.
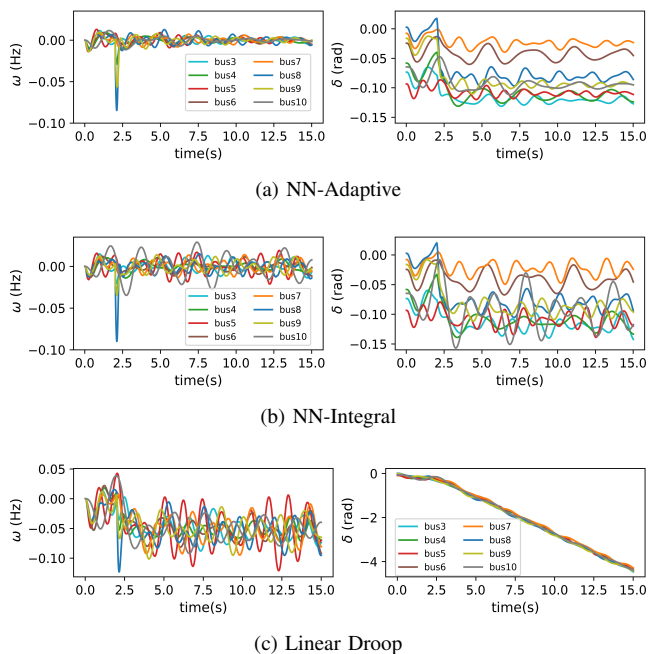

(a) NN-Adaptive

(b) NN-Integral

(c) Linear Droop

Fig. 4. Dynamics of the frequency deviation ω and angle δ on 8 nodes with a step load change at 2s. (a) NN-Adaptive achieves fast restoration of frequency after disturbances and the lowest oscillations. (b) NN-Integral has higher oscillations. (c) Linear droop control cannot restore the frequency to its nominal value and has significantly larger oscillations.

To demonstrate the performance of different controllers, we show the dynamics of the frequency deviation ω and angle δ after a step increase of load at the time of 2s in Fig. 4. As guaranteed in Theorem 1, ω of NN-Adaptive in Fig. 4(a) converges to approximately zero after the step load change. NN-Adaptive in Fig. 4(b) also pulls back the frequency deviation to the level close to zero, but it has much larger oscillations compared with NN-Adaptive. By contrast, linear droop control in Fig. 4(c) cannot recover frequency at the nominal value and experiences significantly large oscillations. Therefore, the adaptive control law can greatly improve both transient and frequency-restoration performances.

B. IEEE 50-machine, 145-bus test case

To validate the performance of the proposed method in larger systems, we further conduct case studies on IEEE 145-bus, 50-machine dynamic test case [27]. The parameter setting for the neural network controller and the training process is the same as the 39-bus system. Apart from the step change of load, here we add noise uniformly distributed in uniform$[-0.03, 0.03]$ p.u. (where 1p.u.=100 MW) on the net power injections $p_i(t), i \in [n]$ to account for the time-varying deviations that are not covered in $\phi_i(t)^\top a_i$.

The average batch loss during epochs of training is shown in Fig. 5 and the transient & frequency-restoration cost on the test set is shown in Fig. 6. Similar to the observations in IEEE 39-bus test system, NN-Adaptive has the lowest transient and frequency-restoration cost. The transient cost of NN-Adaptive is 8% lower than NN-Integral and 69% lower than Linear Droop, respectively. The frequency-restoration cost of NN-Adaptive is 49% lower than NN-Integral and 88% lower than Linear Droop, respectively. To visualize the performance improvement, we show the dynamics of the frequency deviation ω and angle δ after a step increase of load at the time of 2s in Fig. 7. Despite the existence of noise, ω of NN-Adaptive in Fig. 7(a) converges to approximately zero. NN-Integral in Fig. 7(b) and linear droop control in Fig. 7(c) experience much larger oscillations compared with NN-Adaptive. Moreover, droop control also cannot restore the frequency to its nominal value. Therefore, conventional droop control may not be enough to achieve satisfactory frequency response when there exists more frequent fluctuations of net load.
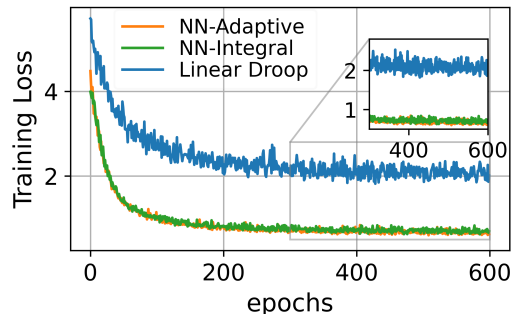

Fig. 5. Average batch loss along epochs for IEEE 145-bus test case. All converge, with NN-Adaptive and NN-Integral achieving much lower loss than Linear Droop. The training loss only reflects transient cost but cannot reflect the performance of frequency restoration.

VI. CONCLUSION

This paper proposes an adaptive approach that can be built on conventional droop control or emerging neural network-based controllers to efficiently adapt to time-varying net load through predictions. We prove that frequencies will converge to their nominal value under the compositional adaptive controller design. Case studies on IEEE 39-bus and 145-bus systems demonstrate that the proposed method improves both transient and frequency-restoration performances compared
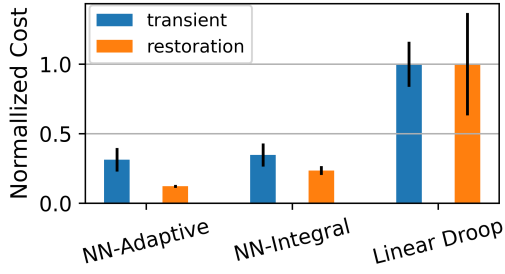
Fig. 6. The average transient cost and frequency-restoration cost with error bar on the randomly generated test set with size 300. NN-Adaptive achieves the lowest transient and frequency-restoration cost.
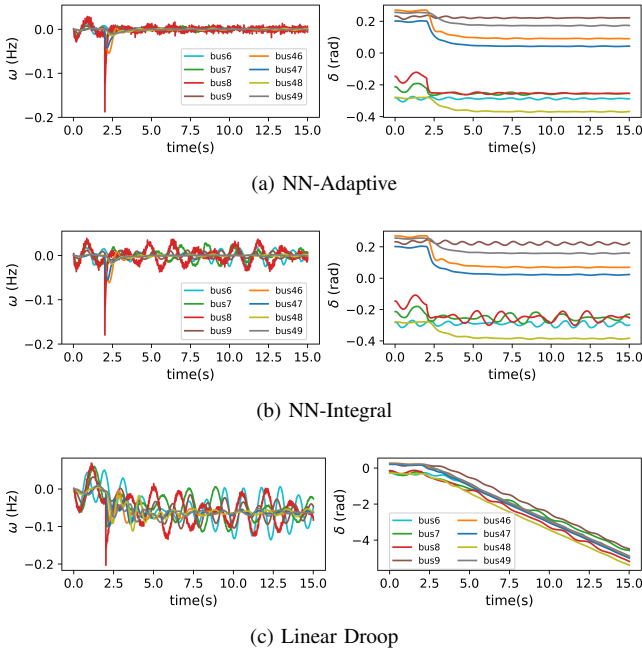

(a) NN-Adaptive

(b) NN-Integral

(c) Linear Droop

Fig. 7. Dynamics of the frequency deviation ω and angle δ on 8 nodes with a step load change at 2s. (a) NN-Adaptive achieves fast restoration of frequency after disturbances and the lowest oscillations. (b) NN-Integral has higher oscillations. (c) Linear droop control cannot restore frequency to its nominal value and has large oscillations.

with existing approaches, while conventional droop control is not able to remove large oscillations and cannot always drive frequency to its nominal value. Future directions include rigorous analysis of the performance guarantee subject to noises and prediction errors, and case studies using predictions from real system data.

REFERENCES

[1] P. W. Sauer, M. A. Pai, and J. H. Chow, *Power system dynamics and stability: with synchrophasor measurement and power system toolbox.* John Wiley & Sons, 2017.
[2] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan, "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE Power and Energy Magazine*, vol. 15, no. 2, pp. 61–73, 2017.
[3] 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, 2023.
[4] Z. Yuan, C. Zhao, and J. Cortes, "Reinforcement learning for distributed transient frequency control with stability and safety guarantees," *arXiv preprint arXiv:2207.03329*, 2022.
[5] Y. Jiang, W. Cui, B. Zhang, and J. Cortés, "Stable reinforcement learning for optimal frequency control: A distributed averaging-based integral approach," *IEEE Open Journal of Control Systems*, vol. 1, pp. 194–209, 2022.
[6] W. Cui, Y. Jiang, B. Zhang, and Y. Shi, "Structured neural-pi control for networked systems: Stability and steady-state optimality guarantees," *arXiv preprint arXiv:2206.00261*, 2022.
[7] C. Zhao, E. Mallada, and F. Dörfler, "Distributed frequency control for stability and economic dispatch in power networks," in *Proc. of American Control Conference*, July 2015, pp. 2359–2364.
[8] J. Schiffer, F. Dörfler, and E. Fridman, "Robustness of distributed averaging control in power systems: Time delays & dynamic communication topology," *Automatica*, vol. 80, pp. 261–271, June 2017.
[9] E. Weitenberg, C. De Persis, and N. Monshizadeh, "Exponential convergence under distributed averaging integral frequency control," *Automatica*, vol. 98, pp. 103–113, Dec. 2018.
[10] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.
[11] H. Bevrani, M. R. Feizi, and S. Ataee, "Robust frequency control in an islanded microgrid: H_∞ and μ-synthesis approaches," *IEEE transactions on smart grid*, vol. 7, no. 2, pp. 706–717, 2015.
[12] D. Tabas and B. Zhang, "Computationally efficient safe reinforcement learning for power systems," in *2022 American Control Conference (ACC).* IEEE, 2022, pp. 3303–3310.
[13] A. Kusiak and Z. Zhang, "Short-horizon prediction of wind power: A data-driven approach," *IEEE Transactions on Energy Conversion*, vol. 25, no. 4, pp. 1112–1122, 2010.
[14] Y. Wang, N. Zhang, Q. Chen, D. S. Kirschen, P. Li, and Q. Xia, "Data-driven probabilistic net load forecasting with high penetration of behind-the-meter pv," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3255–3264, 2017.
[15] W. Charytoniuk, M. Chen, and P. Van Olinda, "Nonparametric regression based short-term load forecasting," *IEEE transactions on Power Systems*, vol. 13, no. 3, pp. 725–730, 1998.
[16] P. Lu, L. Ye, Y. Tang, Y. Zhao, W. Zhong, Y. Qu, and B. Zhai, "Ultra-short-term combined prediction approach based on kernel function switch mechanism," *Renewable Energy*, vol. 164, pp. 842–866, 2021.
[17] M. Sun, T. Zhang, Y. Wang, G. Strbac, and C. Kang, "Using bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 188–201, 2019.
[18] Y. Su, H. Li, Y. Cui, S. You, Y. Ma, J. Wang, and Y. Liu, "An adaptive pv frequency control strategy based on real-time inertia estimation," *IEEE Transactions on Smart Grid*, vol. 12, no. 3, pp. 2355–2364, 2020.
[19] X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement learning for selective key applications in power systems: Recent advances and future challenges," *IEEE Transactions on Smart Grid*, 2022.
[20] G. Shi, K. Azizzadenesheli, M. O'Connell, S.-J. Chung, and Y. Yue, "Meta-adaptive nonlinear control: Theory and algorithms," *Advances in Neural Information Processing Systems*, vol. 34, pp. 10 013–10 025, 2021.
[21] M. O'Connell, G. Shi, X. Shi, K. Azizzadenesheli, A. Anandkumar, Y. Yue, and S.-J. Chung, "Neural-fly enables rapid learning for agile flight in strong winds," *Science Robotics*, vol. 7, no. 66, p. eabm6597, 2022.
[22] L. Huang, J. Lygeros, and F. Dörfler, "Robust and kernelized data-enabled predictive control for nonlinear systems," *arXiv preprint arXiv:2206.01866*, 2022.
[23] E. Weitenberg, Y. Jiang, C. Zhao, E. Mallada, C. De Persis, and F. Dörfler, "Robust decentralized secondary frequency control in power systems: Merits and tradeoffs," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 3967–3982, Oct. 2019.
[24] J.-J. E. Slotine, W. Li *et al.*, *Applied nonlinear control.* Prentice hall Englewood Cliffs, NJ, 1991, vol. 199, no. 1.
[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2012.
[26] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Transactions on Power Apparatus and Systems*, no. 2, pp. 573–584, 1979.
[27] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE transactions on Power Systems*, vol. 7, no. 4, pp. 1559–1564, 1992.