

Learning-Augmented Primal-Dual Control Design for Secondary Frequency Regulation

Yixuan Yu, Rajni K. Bansal, Yan Jiang, and Pengcheng You

Abstract—Frequency stability is fundamental to the secure operation of power systems. With growing uncertainty and volatility introduced by renewable generation, secondary frequency regulation - responsible for restoring the nominal system frequency - must now deliver enhanced performance not only in the steady state but also during transients. This paper presents a systematic framework to embed learning in the design of a primal-dual controller that provides provable (potentially exponential) stability and steady-state optimality, while simultaneously improving key transient metrics, including frequency nadir and control effort, in a data-driven manner. In particular, we employ the primal-dual dynamics of an optimization problem that encodes steady-state objectives to realize secondary frequency control with asymptotic stability guarantee. To augment transient performance of the controller via learning, a change of variables on control inputs, which will be deployed by neural networks, is proposed such that under mild conditions, stability and steady-state optimality are preserved. It further allows us to define a learning goal that accounts for the exponential convergence rate, frequency nadir and accumulated control effort, and use sample trajectories to enhance these metrics. Simulation results validate the theories and demonstrate superior transient performance of the learning-augmented primal-dual controller.

I. INTRODUCTION

The stability of power systems requires maintaining the frequency close to its nominal value (50 Hz or 60 Hz), which is essential for the safe and reliable operation of grid equipment. From a time-scale perspective, frequency control is traditionally divided into three layers [1]: primary control for immediate stabilization, secondary control to restore the frequency to its nominal value, and tertiary control for economic dispatch to optimize steady-state generation.

Recent research has sought to achieve secondary frequency control and economic optimality at the same time, which can be broadly categorized into two classes of approaches. The first class leverages consensus-based algorithms to enhance the efficiency of classical PI controllers [2]–[5]. The second approach models the objectives of frequency regulation and

economic optimality jointly as an optimization problem, where the closed-loop system is interpreted through primal-dual dynamics [6]–[10]. Particularly, the primal-dual framework has attracted significant attention, as it provides a unifying perspective and bridges the control objective of frequency regulation with the economic objective of optimal power system operation. However, they are all focusing on the optimality of the steady state, ignoring the transient behavior of power systems.

In practice, transient performance is also critical for the safety and efficiency of power systems. Some works [11], [12] improve transient performance using model-based feedback control methods, such as LQR and MPC. These methods typically rely on linear feedback since it is challenging to incorporate nonlinear control laws into the optimal control problem. Nevertheless, nonlinear designs become important when state deviations are large, as they can capture richer system dynamics more accurately than linear approximations. This motivates the use of learning-based approaches, which offer strong capability in parameter optimization.

Another line of works has employed neural network (NN) to achieve better transient performance in frequency regulation tasks. Some of them [13], [14] improve transient performance by parameterizing the controller with a monotone neural network and training its parameters. However, these approaches do not account for steady-state economic optimality. To address this limitation, a nonlinear extension of the linear distributed averaging-based integral (DAI) controller in [5] is proposed in [15], where the nonlinear component is parametrized by a monotone neural network. Although this design ensures steady-state optimality and enhances transient performance, it remains restricted in terms of the optimality objectives it can accommodate, such as frequency regulation task with constrained optimality. Taken together, these works consistently rely on monotone structures in their controllers. However, the use of monotonicity is typically assumed rather than derived, lacking a clear rationale or systematic justification of its underlying mechanism. Additionally, they employ similar transient metrics, yet these metrics offer little theoretical guarantee about the destination of convergence in learning. This naturally raises a key question: *how can we systematically design a controller that achieves secondary frequency regulation and steady-state economic optimality, while allowing learning to augment transient performance under theoretical guarantee?*

To address this, we extend the primal-dual framework to design a controller with nonlinear, learnable component. Specifically, we model the optimal steady-state objectives

P. You is the co-corresponding author.

This work was supported in part by the National Natural Science Foundation of China under Grant 72431001, in part by CUHKSZ University Development Fund, and in part by the Brij Disa Centre for Data Science and Artificial Intelligence (CDSA) at the Indian Institute of Management Ahmedabad (IIMA).

Y. Yu and P. You are with the Department of Control Science and Systems Engineering, Peking University, Beijing, China. yuyixuan834@stu.pku.edu.cn, pcyou@pku.edu.cn

R. K. Bansal is with the Operations and Decision Sciences Area, Indian Institute of Management Ahmedabad, Gujarat, India. rajnikantb@iima.ac.in

Y. Jiang is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. yjiang@cuhk.edu.cn

jointly as an optimization problem and reinterpret the controller as part of its associated primal-dual dynamics. This interpretation provides a principled way to incorporate nonlinear control law without compromising asymptotic stability or steady-state economic optimality. Building upon this, the nonlinear learnable component is parameterized by a monotone neural network, enabling the enhancement of transient performance. To further ensure the effectiveness of learning, we introduce metrics for transient performance and, in turn, provide theoretical potential of exponential convergence.

Our contributions are summarized as follows.

- 1) **Unified framework for steady state and transient performance:** We develop a unified framework for controller design that simultaneously guarantees optimal steady-state objectives and augments transient performance. Specifically, we reinterpret the controller as a learning embedded primal-dual dynamics. The asymptotic stability and economic optimality are ensured by the underlying primal-dual structure, while transient enhancement is achieved through learning embedded in the framework.
- 2) **Learning flexibility via nonlinear change of variables:** The learnable component in the controller originates from a change of variables from an optimization perspective, systematically embedding flexibility for learning into the primal-dual based controller design. Importantly, strict monotonicity of the change of variables ensures the preservation of steady-state stability and optimality.
- 3) **Convergence rate metric for transient improvement:** We propose a metric to evaluate the convergence rate of the controller and guide the learning process. We further prove, theoretically, that the closed-loop system achieves potentially exponential stability after learning, based on this metric.

II. PROBLEM STATEMENT

A. Power System Model

We consider a n -bus power network with the topology characterized by a directed and connected graph $(\mathcal{V}, \mathcal{E})$, where the buses indexed by $i \in \mathcal{V} := \{1, \dots, n\}$ are connected through the transmission lines indexed by ordered pairs $(i, j) \in \mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. Without loss of generality, we assume there is only one aggregate controllable generator at each bus subject to a linear approximation of swing dynamics [16]. More precisely, given the power disturbance p_i at bus $i \in \mathcal{V}$, represented by a step change scalar, the dynamics of the phase angle θ_i and the frequency ω_i from the nominal frequency are given by

$$\dot{\theta}_i = \omega_i, \quad (1a)$$

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

where $M_i > 0$ is the generator inertia constant, $D_i > 0$ summarizes the generator damping and frequency-dependent load, and $B_{ij} > 0$ characterizes the sensitivity of the power

flow on line $(i, j) \in \mathcal{E}$ to the phase angle differences between its two end buses. u_i is a controllable power injection, by generator, to be designed for frequency control. Note that the network model (1) implicitly assumes that the variables θ_i , ω_i , u_i are deviations from their nominal values.

For convenience in analysis, we define $\bar{\theta}_{ij} := \theta_i - \theta_j$ as phase angle difference on line $(i, j) \in \mathcal{E}$ and rewrite the swing dynamics in a compact form:

$$\dot{\bar{\theta}} = C^T \omega, \quad (2a)$$

$$M \dot{\omega} = u - p - D \omega - C B \bar{\theta}, \quad (2b)$$

where $\bar{\theta} := (\bar{\theta}_{ij}, (i, j) \in \mathcal{E})$, $\omega := (\omega_i, i \in \mathcal{V})$, $u := (u_i, i \in \mathcal{V})$, and $p := (p_i, i \in \mathcal{V})$. $M := \text{diag}(M_i, i \in \mathcal{V})$, $D := \text{diag}(D_i, i \in \mathcal{V})$, and $B := \text{diag}(B_{ij}, (i, j) \in \mathcal{E})$ are diagonal positive definite matrices. $C \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{E}|}$ is the incidence matrix of the directed graph $(\mathcal{V}, \mathcal{E})$.

B. Optimal Steady-State Objectives

The basic goal of secondary frequency control for the power system is to not only stabilize the frequencies, i.e., $\dot{\omega} = 0$, but also restore them to the nominal value such that $\omega = 0$. Under these steady-state conditions, (2b) reduces to a characterization of the nodal power balance over the network:

$$u - p - C B \bar{\theta} = 0. \quad (3)$$

Beyond this, we further require that at the steady state, the control efforts have to be optimally allocated according to a given cost function $F(u) := \sum_{i=1}^n F_i(u_i)$ for the power injection at each bus i . We assume that $F(u) : \mathbb{R}^n \rightarrow \mathbb{R}$ is strictly convex and second-order continuously differentiable with its gradient satisfying $\nabla F(u) = 0$ at $u = 0$.¹

Therefore, given p , the optimal steady-state problem that minimizes the aggregate control cost to achieve secondary frequency control is given by

$$\min_{\bar{\theta}, \omega, u} F(u) = \sum_{i=1}^n F_i(u_i) \quad (4a)$$

$$\text{s. t. } u - p - D \omega - C B \bar{\theta} = 0 \quad (4b)$$

$$u - p - C B \bar{\theta} = 0, \quad (4c)$$

where (4b) and (4c) can jointly enforce $\dot{\omega} = 0$ and $\omega = 0$ at optimal.

C. Transient Performance Objectives

Although the optimal steady-state objectives guarantee desirable long-run behavior, it lacks a mechanism to shape transient dynamics. In practice, the quality of power system operation strongly depends on transient behaviors, such as the speed of recovery (convergence rate), the maximum frequency deviation (frequency nadir), and the accumulated control effort required throughout the process.

To employ learning for transient performance improvement, the key challenge is integrating sufficient flexibility

¹The condition $\nabla F(0) = 0$ is natural from a physical standpoint, since it implies that zero control input corresponds to the minimum operating cost.

into the controller design to apply learning effectively, while still achieving optimal steady-state objectives. Furthermore, we aim to design appropriate transient performance metrics that can guide the learning process, ensuring the resulting controller achieves favorable performance.

III. CONTROLLER DESIGN AND ANALYSIS

In this section, we build upon the primal-dual framework for secondary frequency regulation to design a controller that not only achieves the optimal steady-state objectives but also provides sufficient flexibility for learning. We start by presenting the proposed controller. Then we analyze the controller from a primal-dual perspective, revealing the underlying structure that enables flexibility for learning without compromising stability and steady-state economic optimality.

To achieve the optimal steady-state objectives and better transient performance at the same time, we propose the following controller²:

$$u = f(s) \quad (5)$$

with

$$\dot{s} = -[\nabla F(f(s)) + \omega + \lambda], \quad (6a)$$

$$\dot{\lambda} = \Gamma^\lambda [f(s) - p - CB\tilde{\phi}], \quad (6b)$$

$$\dot{\tilde{\phi}} = \Gamma^{\tilde{\phi}} [BC^T \lambda], \quad (6c)$$

where $s := (s_i, i \in \mathcal{V})$, $\lambda := (\lambda_i, i \in \mathcal{V})$, $\tilde{\phi} := (\tilde{\phi}_{ij}, (i, j) \in \mathcal{V})$ are newly introduced internal variables and $\Gamma^\lambda, \Gamma^{\tilde{\phi}} \succ 0$ denote diagonal constant control gains. $f(s) := (f_i(s_i), i \in \mathcal{V})$ is a nonlinear function.

The update rules (6) of these variables are demonstrated to enforce the optimal steady-state objectives based on the system dynamics, i.e., in a model-based manner. At the same time, the nonlinear $u = f(s)$ is left unspecified to be learned, aiming at enhancing the transient performance.

A. Primal-Dual Interpretation

To systematically analyze the proposed learnable controller and reveal the underlying mechanism that provides flexibility for learning, we interpret it from a primal-dual perspective. In particular, we construct the following optimization problem, whose primal-dual dynamics aligns with the closed-loop system (2), (5), and (6).

$$\min_{\tilde{\theta}, \omega, s, \tilde{\phi}} F(f(s)) + \frac{1}{2} \omega^T D \omega \quad (7a)$$

$$\text{s. t. } f(s) - p - D\omega - CB\tilde{\theta} = 0 \quad : \nu \quad (7b)$$

$$f(s) - p - CB\tilde{\phi} = 0 \quad : \lambda \quad (7c)$$

where $\nu \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^n$ are Lagrange multipliers of constraints (7b) and (7c), respectively. Accordingly, the Lagrangian of (7) is

$$\begin{aligned} L(\tilde{\theta}, \omega, s, \nu, \lambda, \tilde{\phi}) := & F(f(s)) + \frac{1}{2} \omega^T D \omega \\ & + \nu^T (f(s) - p - D\omega - CB\tilde{\theta}) \\ & + \lambda^T (f(s) - p - CB\tilde{\phi}). \end{aligned} \quad (8)$$

²We avoid using p directly in practice, as disturbances are unmeasurable. Following [10], it can be replaced with $-M\dot{\omega} + f(s) - D\omega - CB\tilde{\phi}$.

Similarly to [17], the swing dynamics (2) can be expressed as

$$\dot{\tilde{\theta}} = -\Gamma^{\tilde{\theta}} \nabla_{\tilde{\theta}} L(\tilde{\theta}, \omega, s, \nu, \lambda, \tilde{\phi}), \quad (9a)$$

$$\omega = \arg \min_{\omega} L(\tilde{\theta}, \omega, s, \nu, \lambda, \tilde{\phi}), \quad (9b)$$

$$\dot{\nu} = \Gamma^{\nu} \nabla_{\nu} L(\tilde{\theta}, \omega, s, \nu, \lambda, \tilde{\phi}), \quad (9c)$$

with $\Gamma^{\tilde{\theta}} = B^{-1}$ and $\Gamma^{\nu} = M^{-1}$. Note that (9b) enforces $\omega \equiv \nu$ along the trajectory. Hence, we can use $\omega \leftrightarrow \nu$ interchangeably and define the following reduced Lagrangian:

$$\begin{aligned} \tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}) := & F(f(s)) + \frac{1}{2} \omega^T D \omega \\ & + \omega^T (f(s) - p - D\omega - CB\tilde{\theta}) \\ & + \lambda^T (f(s) - p - CB\tilde{\phi}). \end{aligned} \quad (10)$$

Then, the closed-loop system can be equivalently rewritten as

$$\dot{\tilde{\theta}} = -\Gamma^{\tilde{\theta}} \nabla_{\tilde{\theta}} \tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}), \quad (11a)$$

$$\dot{\omega} = \Gamma^{\omega} \nabla_{\omega} \tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}), \quad (11b)$$

$$\dot{s} = -(J_f(s))^{-1} \nabla_s \tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}), \quad (11c)$$

$$\dot{\lambda} = \Gamma^{\lambda} \nabla_{\lambda} \tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}), \quad (11d)$$

$$\dot{\tilde{\phi}} = -\Gamma^{\tilde{\phi}} \nabla_{\tilde{\phi}} \tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}), \quad (11e)$$

with $\Gamma^{\omega} = \Gamma^{\nu}$, $\Gamma^{\lambda}, \Gamma^{\tilde{\phi}} \succ 0$ being diagonal step-size matrices. $J_f(s) := \text{diag}(f'_i(s_i), i \in \mathcal{V})$ is the Jacobian of $f(\cdot)$ to s .

Remark 1 (Precondition): Interpreting the controller within the primal-dual framework reveals a key role of nonlinear structure $u = f(s)$. From an optimization perspective, the change of variables $u = f(s)$ acts as a preconditioning mechanism: This technique improves the conditioning of a optimization problem and accelerates convergence to the optimal solution by modifying the gradient's magnitude and direction [18]. In our formulation, the gradient flow of problem (7) is rescaled, as reflected in the additional step-size matrix $(J_f(s))^{-1}$ in (11c).

We also highlight the role of introducing $\tilde{\phi}$ as virtual phase angle differences in (7c) instead of directly using $\tilde{\theta}$. Basically, this substitution enables the swing dynamics (2) to be embedded into the primal-dual dynamics of problem (7). Otherwise, the gradient of $\tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi})$ with respect to $\tilde{\theta}$ would not align with the frequency dynamics (2a) that are determined by the physics.

B. Stability Analysis

Having reinterpreted the closed-loop system (2), (5), and (6) as the primal-dual gradient flow of $\tilde{L}(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi})$, now we proceed to characterize the equilibrium of primal-dual dynamics, i.e., the equilibrium of the closed-loop system. Note that the designed optimization problem (7) is inherently non-convex, since the function $f(s)$ in objective function is non-convex and the equality constraints (7b), (7c) are non-affine. Therefore, it is nontrivial to ensure that the primal-dual dynamics converge to a primal-dual optimal solution to (8). To this end, we first make an assumption for the nonlinear change of variables $u = f(s)$ as follows:

Assumption 1 (Change of variables f): $\forall i \in \mathcal{V}, f_i : \mathbb{R} \rightarrow \mathbb{R}$ is a strictly increasing and Lipschitz continuous function passing through the origin.

With Assumption 1 and the strict convexity of F , the composition $\nabla F(f(s))$ is also a strictly monotone function passing through the origin. This structure generalizes the principle of linear negative feedback, which underlies many classical controllers, including droop-like laws in power systems. With this, the following lemma shows that problem (7) has a unique global optimal solution. Then Lemma 2 proves that problem (7) is equivalent to (4). Finally, Theorem 1 establishes the equivalence between closed-loop equilibrium and the optimal solution to (7), thereby ensuring the achievement of optimal steady-state objectives due to Lemma 2.

Lemma 1 (Uniqueness): Suppose Assumption 1 holds. The optimization problem (7) admits a unique global optimal solution $(\hat{\theta}^*, \omega^*, s^*, \hat{\phi}^*)$. Furthermore, there exists a unique Lagrange multiplier (ν^*, λ^*) such that $(\hat{\theta}^*, \omega^*, s^*, \hat{\phi}^*, \nu^*, \lambda^*)$ is the only solution satisfying the KKT conditions of problem (7).³

Lemma 1 ensures that $(\hat{\theta}^*, \omega^*, s^*, \hat{\phi}^*, \nu^*, \lambda^*)$ is the unique primal-dual optimal solution to problem (7), and that no other feasible point $(\hat{\theta}, \hat{\omega}, \hat{s}, \hat{\phi}) \neq (\hat{\theta}^*, \omega^*, s^*, \hat{\phi}^*)$, together with any multipliers (ν, λ) , can satisfy the KKT conditions, even if problem (7) is nonconvex.

Remark 2 (Hidden convexity): With Assumption 1, the inverse change of variables $s = f^{-1}(u)$ effectively transforms (7) into a convex optimization problem, as shown in the proof of Lemma 1. This highlights that problem (7) is not arbitrarily constructed, but rather embeds a hidden convex structure [20, Chapter 4.2.5] - can be converted to a convex problem with variable bijection.

Upon Lemma 1, we now turn to investigate the relationship between problem (7) and the optimal steady-state problem (4).

Intuitively, the two equality constraints (7b) and (7c) jointly ensure that the optimal solution to problem (7) satisfies $\omega_i^* = 0$ if one temporarily views $\hat{\theta}$ and $\hat{\phi}$ as the same. Then the additional quadratic term $\omega^T D \omega / 2$ in the objective function does not alter the optimal solution. This observation indicates that the designed optimization problem (7) is equivalent to the optimal steady-state problem (4). The following lemma formalizes this equivalence.

Lemma 2 (Equivalence of optimization problems): Suppose Assumption 1 holds. A point $(\hat{\theta}^*, \omega^*, s^*, \hat{\phi}^*)$ is an optimal solution to (7) if and only if $(\hat{\theta}^*, \omega^*, u^*)$ is an optimal solution to (4), with $u^* = f(s^*)$ and $\hat{\theta}^* = \hat{\phi}^*$.

With Lemmas 1 and 2 in mind, we now claim that the closed-loop equilibrium attains the optimal solution to problem (7), as stated in the following theorem, thereby achieving the optimal steady-state objective.

Theorem 1 (Closed-loop equilibrium): Given $\omega^* = \nu^*$, a point $(\hat{\theta}^*, \omega^*, s^*, \lambda^*, \hat{\phi}^*)$ is an equilibrium of the closed-loop system (2), (5), and (6) if and only if $(\hat{\theta}^*, \omega^*, s^*, \nu^*, \lambda^*, \hat{\phi}^*)$ is a primal-dual optimal solution to problem (7).

³Due to space limit, all the proofs of this paper are provided in [19]

With the characterization of the closed-loop equilibrium, we define the equilibrium set as

$$E := \{(\tilde{\theta}, \omega, s, \lambda, \tilde{\phi}) \mid \dot{\tilde{\theta}}, \dot{\omega}, \dot{s}, \dot{\lambda}, \dot{\tilde{\phi}} = 0\},$$

and establish the stability of controller (5) and (6) through the following theorem.

Theorem 2 (Asymptotic stability): Suppose Assumption 1 holds. Starting from any initial point in $\mathbb{R}^{3|\mathcal{V}|+2|\mathcal{E}|}$, the trajectory of (2), (5), and (6) asymptotically converges to the equilibrium in E .

Remark 3 (Change of variables): With theorem 1 and 2, we have shown that the nonlinear change of variables preserves both asymptotic stability and steady-state economic optimality, provided that the transformation is strictly monotone. More importantly, it is this nonlinear change of variables that embeds flexibility for learning within the primal-dual framework systematically. Furthermore, this interpretation reveals the potential to extend the learning-augmented control design towards a broader range of optimization objectives, including, for instance, the operational constraints.

IV. LEARNING FOR TRANSIENT IMPROVEMENT

We have proposed a class of controllers that achieve asymptotic stability with economically optimal steady-state operation. Within this class, we can search for the best controller under transient metrics of interest by parameterizing the change of variables $u = f(s)$ as a monotone neural network and training it through reinforcement learning. In this section, we first design a new transient performance metric that provides theoretical guarantee for learning, followed by neural network training.

A. Transient Performance Metrics

For transient performance, we first focus on the convergence rate. Theorem 2 only guarantees that the system converges to the equilibrium as $t \rightarrow \infty$, but it cannot guarantee the convergence rate. Therefore, we introduce a metric to explicitly quantify the transient decay behavior

$$R_{i,T}^\alpha := \int_0^T e^{\alpha t} |\omega_i(t)|^2 dt, \quad (12)$$

where $\alpha > 0$ denotes the exponential rate, reflecting the theoretical convergence speed of system frequency.

Theorem 3 (Potentially exponential stability): Starting from any initial point in $\mathbb{R}^{3|\mathcal{V}|+2|\mathcal{E}|}$, suppose $T \rightarrow \infty$ and $R_T^\alpha := \sum_{i=1}^n R_{i,T}^\alpha$ converges during training. Then the learned controller will exponentially stabilize the frequency deviation ω with rate at least $\alpha/2$.

Theorem 3 provides a theoretical guarantee that the controller drives the system to the equilibrium with an exponential response, under the assumption of ideal condition and learnability, i.e., $T \rightarrow \infty$ and R_T^α converges. Although such an ideal condition cannot be fully achieved in practice, the metric (12) still accelerates convergence and improves the controller's response rate.

Besides the convergence rate, we are also concerned with the maximum frequency deviation (frequency nadir)

$$\|\omega_i\|_\infty := \max_{t \geq 0} |\omega_i(t)| \quad (13)$$

and accumulated control effort over the transient period

$$\bar{C}_{i,T} := \frac{1}{T} \int_0^T F_i(u_i) dt. \quad (14)$$

Finally, we define the following loss function to evaluate the transient performance of the controller:

$$J := \sum_{i=1}^n (\rho_r R_{i,T}^\alpha + \rho_n \|\omega_i\|_\infty + \rho_c \bar{C}_{i,T}), \quad (15)$$

where $\rho_r, \rho_n, \rho_c > 0$ are tunable weights that balance the three transient metrics.

B. Strictly Monotone Neural Network Training

According to [13, Theorem 2] and universal approximation property [21], any strictly increasing function can be parameterized by a single-hidden-layer neural network with ReLU activation. To parameterize the strictly monotone function $u = f(s)$, we employ the stacked ReLU monotonic neural network in [15]. By properly designing the relationship between the trainable parameters, the piece-wise linear network is guaranteed to be strictly monotone [13, Lemma 5].

To emulate the closed-loop dynamics, we adopt the recurrent neural network (RNN) structure. Specifically, we follow [15] to discretize (2), (5), and (6) with a time step h .

In training procedure, one batch contains $\lfloor T/h \rfloor$ RNN cells. Each cell receives $(\tilde{\theta}^{(\tau)}, \omega^{(\tau)}, s^{(\tau)}, \lambda^{(\tau)}, \tilde{\phi}^{(\tau)})$ as input and outputs $(\tilde{\theta}^{(\tau+1)}, \omega^{(\tau+1)}, s^{(\tau+1)}, \lambda^{(\tau+1)}, \tilde{\phi}^{(\tau+1)})$ to the next cell with $\tau = 0, 1, \dots, \lfloor T/h \rfloor$. Over $|\mathcal{B}|$ batches, we compute the loss

$$J(\Theta_i(d)) := \frac{1}{|\mathcal{B}|} \sum_{b=1}^{|\mathcal{B}|} \sum_{i=1}^n (\rho_r \hat{R}_{i,T}^\alpha + \rho_n \|\omega_i^{(T)}\|_\infty + \rho_c \hat{C}_{i,T}) \quad (16)$$

and update the trainable parameter set $\Theta_i(d)$ via gradient descent. One such update constitutes an epoch, and the training proceeds for N epochs until $J(\Theta_i(d))$ converges.

V. NUMERICAL RESULTS

In this section, we conduct simulations on the IEEE 39-bus New England test system, which comprises 39 buses interconnected by 46 transmission lines. Among them, 10 buses (indexed by $i \in \{30, \dots, 39\}$) host generators, each with one generator per bus. The system parameters, except D , are adopted from the Power System Toolbox [22], and the droop coefficients are set as $D_i = 150$ p.u. for $i \in \{30, \dots, 39\}$ and $D_i = 100$ p.u. for $i \in \{1, \dots, 29\}$. The generation cost functions are specified as $F_i(u_i) = \frac{1}{4} c_i u_i^4 + b_i$, $\forall i \in \{30, \dots, 39\}$, where the coefficients c_i and b_i are randomly drawn from $(0, 1)$. In what follows, we first present the training results of the monotone neural network used to parameterize the function $u_i = f_i(s_i)$. We then present the performance of the trained controller in comparison with a traditional linear primal-dual based controller.

A. Training of Monotone Neural Network

To implement the training illustrated in Section IV-B, each trajectory sample is obtained by applying random step changes p to the discretized system model. For each bus i , the disturbance p_i is independently drawn from a uniform distribution over $(-5, 5)$ p.u.. The hyperparameters of neural network with 20 hidden neurons are set as follows: $|\mathcal{B}| = 64$, $N = 50$, $\alpha = 3$, $\rho_r = 0.1$, $\rho_n = 1$, and $\rho_c = 1$. The initial learning rate is 0.4, which decays exponentially by a factor of 0.5 every 3 steps. The time series length is 6000, and the optimizer used is Adam.

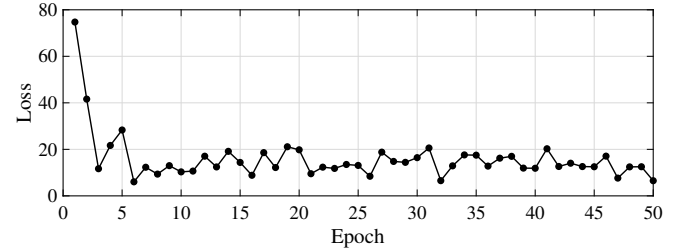


Fig. 1: Loss over epochs in training

The training is conducted with TensorFlow 2.20.0 on an A800 GPU, which takes approximately 16,320 seconds. The evolution of the loss during training is shown in Fig. 1 that the loss J is reduced by approximately 86.7%.

B. Performance of the Learned Controller

We now evaluate the performance of the trained controller and compare it with the traditional primal-dual based linear controller without learning. Disturbance $p_i = 3$ p.u. is applied on bus $i \in \{14, 22, 28, 36, 38\}$ at $t = 0$. Fig. 2 shows the frequency responses and convergence on generator buses $i \in \{30, \dots, 39\}$, under both controller. Fig. 3 compares the control effort of both controllers during transient. According to the results, the learned controller achieves significant improvements in convergence speed, with smaller frequency nadir and reduced control effort during transients. This advantage is further confirmed by the numerical comparison provided in Table I.

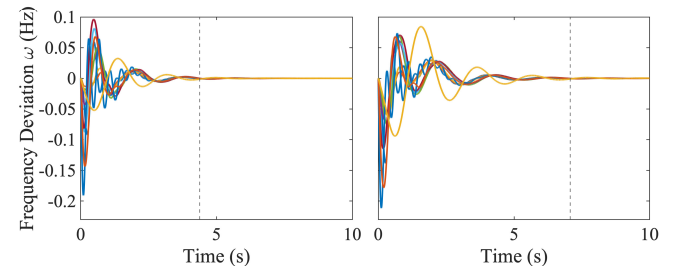


Fig. 2: Frequency deviation of generator buses under a disturbance $p_i = 3$ p.u. on buses $\{14, 22, 28, 36, 38\}$: learned controller (left); traditional primal-dual based controller (right).

To demonstrate that the learned controller achieves economic optimality at steady state, we plot the marginal cost of each generator in Fig. 4. In unconstrained optimization,

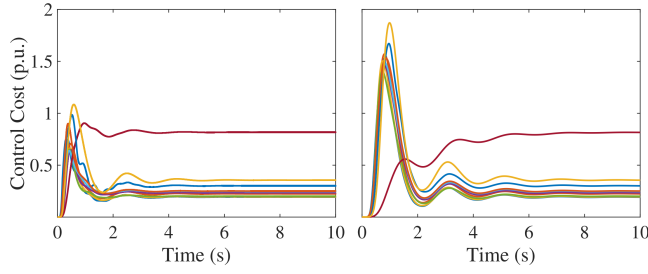


Fig. 3: Cost of generators under a disturbance $p_i = 3$ p.u. on buses $\{14, 22, 28, 36, 38\}$: learned controller (left); traditional primal-dual based controller (right).

	rate (s)	nadir	cost
learned controller	4.3825	0.1680	123.3
linear controller	7.0695	0.1796	133.8

TABLE I: Transient performance comparison of the controllers.

identical marginal costs across generators indicate steady-state optimality [20, Chapter 5.5]. As shown in Fig. 4, the baseline primal-dual controller (right) inherently guarantees economic optimality. Our proposed controller (left) preserves economic optimality while incorporating a nonlinear learnable component, as evidenced by the identical marginal costs.

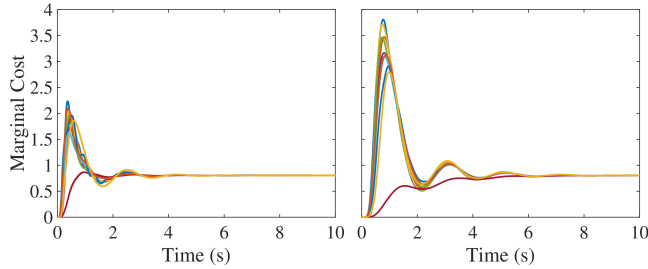


Fig. 4: Identical marginal cost among generators

VI. CONCLUSION

In this work, we propose a controller that simultaneously achieves secondary frequency regulation, steady-state economic optimality, and improved transient performance. By extending the primal-dual framework, we reinterpret the closed-loop system from an optimization perspective, which offers a unified framework for integrating learning flexibility into controller design. This systematic approach also enables addressing broader objectives, including operational constraints. To guarantee effectiveness of learning, we further propose a metric for convergence rate that provides theoretical assurance of potentially exponential convergence of the system under the learned controller. Experiments on the IEEE 39-bus system show that the controller achieves faster convergence, smaller frequency nadir and less accumulated control cost during transients. These results validate that our approach effectively integrates stability, optimality, and learnability in a unified control design.

REFERENCES

- [1] J. Machowski, J. W. Bialek, J. Bialek, and J. R. Bumby, *Power system dynamics and stability*. John Wiley & Sons, 1997.
- [2] M. Andreasson, D. V. Dimarogonas, H. Sandberg, and K. H. Johansson, "Distributed control of networked dynamical systems: Static feedback, integral action and consensus," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1750–1764, 2014.
- [3] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 241–253, 2015.
- [4] Q. Shafiee, J. M. Guerrero, and J. C. Vasquez, "Distributed secondary control for islanded microgrids—a novel approach," *IEEE Transactions on power electronics*, vol. 29, no. 2, pp. 1018–1031, 2013.
- [5] C. Zhao, E. Mallada, and F. Dörfler, "Distributed frequency control for stability and economic dispatch in power networks," in *Proc. of American Control Conference*, 2015, pp. 2359–2364.
- [6] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.
- [7] N. Li, C. Zhao, and L. Chen, "Connecting automatic generation control and economic dispatch from an optimization view," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 254–264, 2016.
- [8] D. Cai, E. Mallada, and A. Wierman, "Distributed optimization decomposition for joint economic dispatch and frequency regulation," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4370–4385, 2017.
- [9] E. Mallada, C. Zhao, and S. Low, "Optimal load-side control for frequency regulation in smart grids," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6294–6309, 2017.
- [10] Z. Wang, F. Liu, S. H. Low, C. Zhao, and S. Mei, "Distributed frequency control with operational constraints, part II: Network power balance," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 53–64, 2019.
- [11] F. R. Badal, S. K. Sarker, Z. Nayem, and S. K. Das, "Robust design of high performance controller for transient stability enhancement of a single machine infinite bus power system," *International Journal of Power and Energy Conversion*, vol. 12, no. 4, pp. 294–313, 2021.
- [12] P. Das, S. P. Biswas, S. Mondal, and M. R. Islam, "Frequency fluctuation mitigation in a single-area power system using LQR-based proportional damping compensator," *Energies*, vol. 16, no. 12, p. 4804, 2023.
- [13] 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.
- [14] Z. Sun, Z. Yuan, C. Zhao, and J. Cortés, "Learning decentralized frequency controllers for energy storage systems," *IEEE Control Systems Letters*, vol. 7, pp. 3459–3464, 2023.
- [15] 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.
- [16] P. You, J. Z. Pang, and E. Yeung, "Stabilization of power networks via market dynamics," in *Proceedings of the Ninth International Conference on Future Energy Systems*, 2018, pp. 139–145.
- [17] P. You, Y. Jiang, E. Yeung, D. F. Gayme, and E. Mallada, "On the stability, economic efficiency and incentive compatibility of electricity market dynamics," *IEEE Transactions on Automatic Control*, pp. 1–16, 2025 (early access).
- [18] J. A. Meijerink and H. A. Van Der Vorst, "An iterative solution method for linear systems of which the coefficient matrix is a symmetric M -matrix," *Mathematics of Computation*, vol. 31, no. 137, pp. 148–162, 1977.
- [19] Y. Yu, R. K. Bansal, Y. Jiang, and P. You, "Learning-augmented primal-dual control design for secondary frequency regulation," 2026. [Online]. Available: <https://arxiv.org/abs/2603.09814>
- [20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [21] Z. Lu, H. Pu, F. Wang, Z. Hu, and L. Wang, "The expressive power of neural networks: A view from the width," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [22] 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.