IEEE Control Systems Letters paper presented at
2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

# Coherency-Aware Learning Control of Inverter-Dominated Grids: A Distributed Risk-Constrained Approach

Kyung-Bin Kwon†, Ramij Raja Hossain†, Sayak Mukherjee†, Kaustav Chatterjee,
Soumya Kundu, Sameer Nekkalapu, Marcelo Elizondo

*Abstract*— **This paper investigates the importance of integrating the coherency knowledge for designing controllers to dampen sustained oscillations in wide-area power networks with significant penetration of inverter-interfaced resources. Coherency is a fundamental property of power systems, where time-scale separation in frequency dynamics leads to clustered behavior among generators of different groups. Large-scale penetration of inverter-driven low inertia resources replacing conventional synchronous generators (SGs) can lead to perturbation in the coherent partitioning; hence, integrating such information is of utmost importance for oscillation control designs. We present the coherency-aware design of a distributed output feedback-based reinforcement learning method that additionally incorporates risk constraints to capture the uncertainties related to net-load fluctuations. The use of domain-aware coherency information has produced improved training and oscillation performance than the coherency-agnostic control design, hence proving to be effective in controller design. Finally, we validated the proposed method with numerical experiments on the benchmark IEEE 68-bus test system.**

*Index Terms*— **Distributed Control, Power System Coherency, Reinforcement Learning, Risk Constraints.**

## I. Introduction

The need for integrating renewable energy sources to cope with global climate-related issues results in a larger penetration of inverter-interfaced resources in modern power grids. The current state-of-the-art inverter technology includes predominantly two types of control, namely grid-following (GFL) and grid-forming (GFM) [1], [2]. GFLs act as a current source, and on the contrary, GFMs behave as a voltage source, allowing regulations of system voltages and frequency like synchronous generators (SGs). This SG-like characteristics helped GFM technologies to receive major attention in recent times and proved to be effective solutions for integrating utility-scale renewable resources [3].

In conventional power systems, replacing high-inertia synchronous generators (SGs) by low-inertia grid-forming machines (GFMs) presents numerous challenges. One significant issue is low-frequency sustained oscillations among coherent areas of power grids, prompting the design of wide-area damping controllers. Slow coherency, a fundamental characteristic [4], stems from the time-scale separation in synchronous machine dynamics. It involves strongly coupled machines oscillating together, synchronizing rapidly to form coherent groups. Conversely, these groups oscillate against each other and synchronize over a slower time scale due to weaker coupling. Works such as [5]–[8] have demonstrated the utility of grid-forming machines (GFMs) in transmission planning, dynamic equivalencing, controlled islanding, and oscillation damping control. The significant penetration of GFM-interfaced utility-scale distributed energy resources (DERs) affects the intrinsic dynamic behavior of power systems. Therefore, it's crucial to theorize how these resources impact the coupled oscillation structures of the grid at low frequencies. Designing wide-area controls would benefit from understanding the grid's underlying clustering structure with GFM-based renewable resources. We address this gap by first capturing the modified coherency structure of the grid due to GFM integration, then designing oscillation control using this coherency information.

Over the last few decades, Reinforcement Learning (RL) has effectively tackled intricate nonlinear dynamic tasks within the context of a Markov Decision Process (MDP) [9], and control-theoretic state-spaces [10], [11]. In power systems, RL is used for short-term transient voltage control [12], [13], energy storage control in microgrids [14], and wide-area damping control [15]. However, RL's model-free approach results in high sample complexity for training [16]. To address this, we use a hybrid approach: employing the nominal dynamic model of the grid with additional process noise to account for uncertainties from load and renewable fluctuations. We then use a zero-order gradient optimization approach within a risk-constrained linear-quadratic regulator (LQR) setting. In addition to the recent work [17], the control design incorporates a distributed architecture to account for infusing the perturbed coherency information, which improves the oscillation performance.

**Contributions.** The contribution of the work is multi-fold. First, we present the formulation of wide-area oscillation control when large-scale GFM integration replaces some of the conventional synchronous generation. Due to the consideration of the load and renewable fluctuations, the mean-variance risk constraints are enforced within the control framework. Next, we present the theoretical analysis of how large-scale grid-forming penetration can structurally impact grid clustering behavior. Subsequently, the paper presents the design of a distributed output feedback reinforcement learning controller that leverages the domain-aware coherency information, demonstrating improved performance compared to coherency-agnostic control designs. Lastly, numerical experiments conducted on the IEEE 68-bus system validate the efficacy of the proposed approach.

Authors are with the Pacific Northwest National Laboratory (PNNL), Richland, WA, 99354, USA. The research described in this paper is part of the E-COMP (Energy System Co-Design with Multiple Objectives and Power Electronics) Initiative at PNNL (operated by Battelle for the U.S. Department of Energy under Contract DEAC05-76RL01830). † Equal contributions. *Corresponding email: sayak.mukherjee@pnnl.gov*.

## II. GFM Integration for SG Replacement: Grid Control Problem

### A. Grid Dynamics

Consider a power grid consisting of $N$ number of buses. We consider $N_r$ number of resource buses, i.e., where we either consider synchronous generators (SGs) or grid-forming inverter (GFM) interfaced resources. From a conventional viewpoint, all such buses would operate with synchronous generators. The dynamics of SG $i \in \mathcal{G}$ in bus $i$ is represented by the following swing equations [18]:

$$\dot{\delta}_i = \omega_i - \omega_0, \tag{1a}$$

$$\dot{\omega}_i = \frac{1}{M_i}\big[D_i(\omega_0 - \omega_i) + P_i - P_{ei}\big], \tag{1b}$$

where $\delta_i, \omega_i, P_i, P_{ei}$ are the angle, frequency, input mechanical power and generated electrical powers of SG $i$, and parameters $\{M_i, D_i\}$ are the inertia and damping coefficients, respectively. At this point, we replace $p$ number of synchronous generators with grid-forming inverters of the same ratings; therefore, we will have $N_r - p$ number of SGs remaining in the grid with indices contained in the set $\mathcal{G}$, and $p$ number of GFMs integrated with indices contained in the set $\mathcal{F}$. For each GFM $j \in \mathcal{F}$ in bus $j$, the active and reactive powers $P_{ej}$ and $Q_{ej}$ at the terminal are calculated using the terminal voltage/current measurements passing through a low-pass filter. The active power-frequency and reactive power-voltage droop controls are implemented as follows [19], [20]:

$$\dot{\delta}_j = \omega_j - \omega_0, \tag{2a}$$

$$\dot{\omega}_j = \frac{1}{\tau_j}\big[\omega_0 - \omega_j + \lambda_j^p(P_j^{\text{ref}} - P_{ej})\big], \tag{2b}$$

$$\dot{V}_j^e = \frac{1}{\tau_j}\big[V_j^{\text{ref}} - V_j^e - V_j + \lambda_j^q(Q_j^{\text{ref}} - Q_{ej})\big], \tag{2c}$$

$$\dot{E}_j = k_j^{pv}\dot{V}_j^e + k_j^{iv}V_j^e \tag{2d}$$

where $\lambda_j^p, \lambda_j^q, \tau_j$ are the active, reactive power droop coefficients and filter time constant, respectively. The parameters $k_j^{pv}$ and $k_j^{iv}$ in (2d) are the proportional and integral gains associated with the $Q$-$V$ control loop, respectively.

### B. Control Formulation

The system operator would like to minimize oscillatory power grid behavior following faults and other disturbances. The operator can create a linearized grid model by linearizing around a stable operating point (Kron reduction [18]) in normal operations. Let us concatenate all the generator states such that $x_s = [\delta_1, \ldots, \delta_{N_r-p}, \omega_1, \ldots, \omega_{N_r-p}]$, and the state vector per GFM $j$ are denoted as $x_f = \{x_j\}, x_j := [\delta_j, \omega_j, V_j^e, E_j]$. Thereafter, the stacked state-vector $x = [x_s, x_f]^\intercal$ would result in an integrated discrete small-signal dynamics by considering the GFM control time $\Delta t$ as:

$$x_{t+1} = Ax_t + Bu_t + \xi_t, t = 0, 1, \ldots, \tag{3}$$

where $u := [\Delta V^{ref}, \Delta P^{ref}, \Delta Q^{ref}]^\intercal \in \mathbb{R}^{3N_f}$ with $V^{ref} = \{V_j^{ref}\}_j, P^{ref} = \{P_j^{ref}\}_j, Q^{ref} = \{Q_j^{ref}\}_j$. Here, we add process noise $\xi_t$ to account for the variability in the grid due to load and renewable perturbations to formulate a robust design.

Utilizing the system dynamics given in equation (3), we determine an optimal control strategy by solving a linear quadratic regulator (LQR). This objective aims to minimize the overall cost associated with both state and control variables:

$$\min_K \; R_0(K) = \lim_{T\to\infty}\frac{1}{T}\mathbb{E}\sum_{t=0}^{T-1}\big[x_t^\intercal Q x_t + u_t^\intercal R u_t\big], \tag{4}$$

with $Q$ and $R$ being positive semi-definite and positive definite weight matrices for each state and control variables, respectively. Our objective is to find the distributed static controller gain matrix $K$ for $u_t = -Ky_t$. Here, $y_t$ are the outputs $Cx_t$ with output matrix $C$ indicating the list of observable states, which enforces the distributed nature of the wide-area control. Along with (4), we implement a mean-variance risk constraint to improve the worst-case performance:

$$\min_K \; R_0(K) = \lim_{T\to\infty}\frac{1}{T}\mathbb{E}\sum_{t=0}^{T-1}\big[x_t^\intercal Q x_t + u_t^\intercal R u_t\big] \tag{5}$$

$$\text{s.t. } R_c(K) = \lim_{T\to\infty}\frac{1}{T}\mathbb{E}\sum_{t=0}^{T-1}\big(x_t^\intercal Q x_t - \mathbb{E}\big[x_t^\intercal Q x_t | \mathcal{H}_t\big]\big)^2 \leq c$$

where $\mathcal{H}_t = [x_0, u_0, \ldots, x_{t-1}, u_{t-1}]$ being a state and control trajectory up to $t$ and $c$ being a risk tolerance parameter. The mean-variance risk measure indicates deviations between the expected and realized state costs. Incorporating the risk constraint helps address challenges from system variability due to external load disturbances, renewable fluctuations, and imperfect modeling, thus enhancing the controller's worst-case performance. Here, we introduce the distributed consideration in the design.

*Coherent-structure-aware Control:* From a system-level perspective, integrating GFM-dominated resources inherently influences the dynamic behavior of power systems. Slow coherency, a key characteristic, arises from the time-scale separation in the dynamics of synchronous machines. Strongly coupled machines oscillate together, forming slow, coherent groups, while these groups oscillate against each other over a slower time scale due to weaker coupling. The goal of distributed control is to leverage this slow coherency structure for better oscillation damping control. Understanding the inherent clustering pattern, especially with significant GFM-based renewable energy, enhances wide-area control performance. Thus, we first identify the altered coherency structure from GFM integration and then design distributed control accordingly.

## III. Distributed Risk-Constrained Learning Control Solution

### A. Perturbed Coherency Behavior

We start by extracting the Laplacian sub-matrix of the state dynamics, i.e., frequency dynamics governed by the influence of angles involving synchronous generators and inverters. Let us consider the vector concatenating all the angles for SGs is denoted by $\delta_s$, and for GFMs by $\delta_f$, and similarly for frequencies in $\omega_s$, and $\omega_f$. Then, we can write the frequency dynamics in compact form for SG: $M\dot{w}_s = \bar{f}_{\text{sync}}(\delta_s, V)$, where $\bar{f}_{\text{sync}}(.)$ denotes the part of the full functional form corresponding to the frequency dynamics. The dynamics are dependent upon the SG angles as the electrical active power from the SG $i$ is given by $P_{si} = (E_i/x'_{di})\,(V_{i_{Re}}\sin\delta_i - V_{i_{Im}}\cos\delta_i)$ . Now, for the GFM units, we can similarly write in the compact form as, $M_f\dot{w}_f = \bar{f}_{\text{gfm}}(\delta_f, E_f, V)$, where $\bar{f}_{\text{gfm}}(.)$ denotes a part of the full dynamics corresponding to the frequency evolution, and $E_f$ denotes stacked internal voltage states for inverters. The equivalent inertia provided by the GFM unit is dependent upon the GFM droop control parameters: $M_f = blkdiag(\{\frac{\tau_j}{\lambda_j^p}\}_j)$. Now, the power-flow equations in the functional form can be slightly elaborated to encapsulate some of these specific states such that, $0 = g(\delta_s, \delta_f, E_f, V)$, At a stable operating point, linearizing the SG, GFM dynamics, and the power-flow equations, we have,

$$M\Delta\dot{w}_s = A_{11}\Delta\delta_s + A_{12}\Delta V, \tag{6}$$

$$M_f\Delta\dot{w}_f = A_{21}\Delta\delta_f + A_{22}\Delta V, + A_{23}\Delta E_f, \tag{7}$$

$$0 = A_{31}\Delta\delta_s + A_{32}\Delta\delta_f + A_{33}\Delta V + A_{34}\Delta E_f, \tag{8}$$

where the matrices $A_{ij}$ denotes the Jacobians corresponding to the argument variables. Let us also perform some manipulations here, such as denoting $A_1 = blkdiag(A_{11}, A_{21}), A_2 = [A_{12}; A_{22}], M_e = blkdiag(M, M_f), A_3 = [A_{31}A_{32}], A_4 = [\mathbf{0}; A_{34}]$. Then, we will have

$$\Delta V = -A_{33}^{-1}A_3\begin{bmatrix}\delta_s\\ \delta_f\end{bmatrix} - A_{33}^{-1}A_{34}\Delta E_f. \tag{9}$$

Using this expression in (6) and (7), we will have,

$$M_e\begin{bmatrix}\Delta\dot{w}_s\\ \Delta\dot{w}_f\end{bmatrix} = (A_1 - A_2A_{33}^{-1}A_3)\begin{bmatrix}\delta_s\\ \delta_f\end{bmatrix} + (A_4 - A_2A_{33}^{-1}A_{34})\Delta E_f. \tag{10}$$

The Laplacian matrix for this Kron-reduced coupled SG-GFM interaction network is captured by,

$$L = (A_1 - A_2A_{33}^{-1}A_3), \tag{11}$$

and the inertia-weighted sub-matrix that captures the influence of angles on the frequency dynamics is given by $M_e^{-1}L$. The eigen-structure of this weighted Laplacian matrix determines how the resultant dynamic coupling between the conventional SG and the GFM is modified.

**Proposition 1:** *(i) GFMs preserve the grid Laplacian structure $L$, such that $L.\mathbf{1} = \mathbf{0}$; (ii) Moreover, a considerable replacement can cause perturbation in coherent boundaries.*

*Proof: Part (i)*: With only conventional synchronous machines integrated into the grid, the linearized model can be recalled as follows: $M\Delta\dot{w}_s = A_{11}\Delta\delta_s + A_{12}\Delta V, 0 = A_{31}\Delta\delta_s + A_{33}\Delta V$, and the Kron-reduced frequency dynamics would lead to, $M\Delta\dot{\omega}_s = (A_{11} - A_{12}A_{33}^{-1}A_{31})\Delta\delta_s$. Thus, the inertia-weighted Laplacian for the sync-only grid is expressed as $M^{-1}\bar{L}$ where: $\bar{L} = (A_{11} - A_{12}A_{33}^{-1}A_{31})$. Consequently, the perturbation resulting from extensive GFM integration, leading to $M_e^{-1}L$ will reflect the intrinsic clustering behavior. The GFM frequency dynamics enable the synchronous-like behavior with the modification to the inertia parameters ($\tau_j/\lambda_j^p$ for inverter $j$). As a result, after the admittance matrices are Kron-reduced to account for the all-to-all coupling between generation resources, including SGs and GFMs, the active power generated by the $i^{th}$ GFM unit toward another GFM or SG (say, denoted by $j^{th}$ unit) would be equivalently given by $E_iE_jB_{ij}\sin(\delta_i - \delta_j)$ where $B_{ij}$ is the equivalent Kron-reduced admittance between SG/GFM resources. The elements of the matrix $L$ are:

$$\begin{aligned} L(i,j) &= E_iE_jB_{ij}\cos(\delta_{i0} - \delta_{j0}) \text{ if } i \sim j, \\ &= 0, \text{if disconnected.} \end{aligned} \tag{12}$$

Considering the active power-balance structure between the SG/GFMs we will have, $L(i,i) = -\sum_{j\in\mathcal{N}_i} L(i,j)$. Therefore, we will have $\mathbf{1}$ as one of its eigenvectors with zero eigenvalue resulting in the condition: $L.\mathbf{1} = \mathbf{0}$. □

*Part (ii)*: We consider $r$ coherent areas, recalling the slow coherent behaviors as in [21]. We consider matrices $W_r$, and $\bar{W}_r$ whose columns are the eigenvectors of the zero eigenvalues and the $(r-1)$ slow eigenvalues of the weighted Laplacians $M_e^{-1}L$, and $M^{-1}\bar{L}$, respectively. A Gaussian elimination-based method is proposed in [5] to create permuted versions of $W_r$ and $\bar{W}_r$ by identifying the group reference machines, denoted as $\mathcal{W}_r$, and $\bar{\mathcal{W}}_r$. The canonical angles between these two sub-spaces are defined as $\theta_i = \cos^{-1}\sigma_i, i = 1, \ldots, r$ where $\sigma_i, i = 1, \ldots, r$ are the $r$ smallest singular values of $\bar{\mathcal{W}}_r^T\mathcal{W}_r$. Let $e_i$ and $\bar{e}_i$ be the $i^{th}$ eigenvalues of $M_e^{-1}L$, $M^{-1}\bar{L}$, respectively. If there is a gap such that $|e_r - \bar{e}_{r+1}| > \beta$ for some $\beta > 0$, then the perturbation in the sub-spaces can be bounded by $||\sin(\Theta)||_F \leq \frac{1}{\beta}||M^{-1}\bar{L}\mathcal{W}_r - \mathcal{W}_r\Sigma_r||_F$, where $\Theta = diag(\theta_1, \ldots, \theta_r)$ and $\Sigma_r = diag(e_1, \ldots, e_r)$; more details can be found in [22], [23]. As a result of this perturbation, the row vectors of the permuted GFM-integrated slow-subspaces ($\alpha_i$'s) will deviate from those of the SG-based subspaces ($\bar{\alpha_i}$'s) as follows:

$$||\alpha_i - \bar{\alpha}_iQ||_F \leq \frac{1+\sqrt{2}}{\beta}||M_e^{-1}L - M^{-1}\bar{L}||_F, \tag{13}$$

where $Q$ is orthogonal matrix that minimizes $||\mathcal{W}_r - \bar{\mathcal{W}}_rQ||_F$. However, since these row vectors are constrained by the hyperplane $\sum_{j=1:r}\alpha_{ij} = 1, i = 1, \ldots, n$, [21] they will exhibit a rotational shift. This results in a new clustering structure, significantly altering $||M_e^{-1}L - M^{-1}\bar{L}||_F$. This captures the impact of GFM integration, introducing a perturbation in the slow eigenspace that ripples through the row vectors of the slow subspace, consequently altering the clustering. □

### *B. Distributed Output Feedback*

*1) Pre-tuning:* The pre-tuning is designed to select appropriate feedback variables based on the modified clustering structure. Here, the main steps are as follows:

- *Step-1: Compute* small-signal model for the grid with GFM penetration.
- *Step-2: Examine* eigen-vectors corresponding to slow eigenvalues (inter-area oscillations) and compute mode-shapes (an example will be given in numerical experiments).
- *Step-3: Compute* the all-new coherent groupings encompassing SGs and GFMs based on the mode-shapes.
- *Step-4: Construct* angular differences along the new partitions, which are used as distributed feedback. This step signifies capturing tie-line active power flows as they are dependent upon angular differences of the boundary resources.

Therefore, feedback are taken from the generating resources assuming they are equipped with decentralized estimators near the coherent boundaries within the nearest possible hops of the form $\Delta\delta_i - \Delta\delta_j$ (signifying $P_{ij}$) where $i^{th}$ and $j^{th}$ resources belong to neighboring areas. Next, we describe the algorithm to compute $u_t = -Ky_t$.

*2) Zeroth-order RL:* Problem (5) poses difficulty due to the risk-aware constraints, and static output feedback, therefore, we utilize the policy gradient based reinforcement learning framework. To solve (5), we first reformulate the constraint to a quadratic form while removing the conditional expectation with respect to $\mathcal{H}_t$ [24]. Specifically, with an assumption that the noise $\xi_t$ has a finite fourth-order moment, the constraint is reformulated as

$$R_c(K) = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\sum_{t=0}^{T-1}\Big(4x_t^{\intercal}QWQx_t + 4x_t^{\intercal}QM_3\Big) \leq \bar{c} \quad (14)$$

where $\bar{c} := c - m_4 + 4\mathrm{tr}\{(WQ)^2\}$, $\bar{\xi} := \mathbb{E}[\xi_t]$, $W := \mathbb{E}[(\xi_t - \bar{\xi})(\xi_t - \bar{\xi})^{\intercal}]$, $M_3 := \mathbb{E}[(\xi_t - \bar{\xi})(\xi_t - \bar{\xi})^{\intercal}Q(\xi_t - \bar{\xi})]$, and $m_4 := \mathbb{E}[(\xi_t - \bar{\xi})^{\intercal}Q(\xi_t - \bar{\xi}) - \mathrm{tr}(WQ)]^2$. Here, we consider random noise samples for $\xi$ to represent worst-case variability in actual real-world load and generation variations. Hence, the required moment values are computed by gathering samples of $\xi$, which are used in offline trajectory generation. Once learning is complete, the learned $K$ is utilized to generate control actions during online implementation. In practice, designers can adopt more systematic methods to generate $\xi$ using historical measurements gathered by utilities. From these datasets, measures of the noise statistics can be obtained.

Next, we implement a stochastic gradient-descent method with a max-oracle (SGDmax), which can consider both the risk constraint and structured feedback using dual approach [25]. First, we formulate the Lagrangian function with the multiplier $\mu \geq 0$, as given by,

$$\mathcal{L}(K,\mu) := R_0(K) + \mu[R_c(K) - \bar{c}]. \quad (15)$$

Note that the Lagrangian retains the same quadratic structure as the original LQR objective. This characteristic allows us to apply a gradient-based approach, which is widely used for solving the unconstrained LQR problem [24]. Specifically, we employ the policy gradient method with the Lagrangian function considered as a value function. We directly update our policy $K$ by computing the estimated gradient of the value function, which, in our case, is represented by the Lagrangian function [9]. With (15), the dual problem is formulated as the maximin problem:

$$\max_{\mu\in\mathcal{Y}} \min_{K} \mathcal{L}(K,\mu), \quad (16)$$

where $\mathcal{Y} := [0, \bar{\mu}]$ is a bounded set for $\mu$ by assuming that (5) is feasible and thus $\mu$ is finite. Here, we consider minimax counterpart of (16) instead, i.e.

$$\min_{K} \Phi(K), \text{ where } \Phi(K) := \max_{\mu\in\mathcal{Y}} \mathcal{L}(K,\mu), \quad (17)$$

which enables us to find the stationary point of (16) since the KKT stationary condition is the same in both problems. Note that we can directly find optimal $\mu$ by choosing $\mu = 0$ if the constraint is satisfied and $\mu = \bar{\mu}$ otherwise as $\mathcal{L}(K,\mu)$ is a linear function of $\mu$. In order to further minimize $\Phi(K)$, we can implement the gradient descent (GD) method with the zero-order policy gradient (ZOPG) [25]. Specifically, we can estimate the ZOPG by utilizing the expression:

$$\hat{\nabla}_K\mathcal{L}(K;U) = \frac{n_K}{r}\mathcal{L}(K + rU, \mu')U, \quad (18)$$

where $U$ represents a random perturbation following the same structure as $K$, with $\|U\| = 1$, while $r$ and $n_K$ indicates the smoothing radius and the number of non-zero entries and $\mu' = \arg\max_{\mu\in\mathcal{Y}} \mathcal{L}(K + rU, \mu)$. Note that incorporating ZOPG eliminates the necessity of computing the first-order gradient $\nabla_K\mathcal{L}(K,\mu)$ as ZOPG can directly utilize the function value $\mathcal{L}(K,\mu)$ to estimate the gradient.

By utilizing the ZOPG, we implement the stochastic gradient-descent with max-oracle (SGDmax) algorithm as indicated in Algorithm 1. From the initial policy $K_0$, we perform iterative gradient descent updates on $K$. To mitigate estimation variance, we leverage $\hat{G}(K)$, which denotes the average of $N_i$ estimates obtained through ZOPG. It is noteworthy that by selecting appropriate values for the smoothing radius $r$, step-size $\eta$, and the number of iterations $M$ along with the Lipschitz and smoothness constants of $\Phi(K)$, we can achieve a high convergence probability to the stationary point, as stated in Lemma 1. The detailed proof with selection of parameters can be found in [25, Appendix B].

**Lemma 1.** *[25] By appropriately selecting parameters $r, \eta$, and $M$ in accordance with the local Lipschitz and smoothness properties of $\Phi(K)$, and initializing a feasible $K_0$, Algorithm 1 is demonstrated to converge to the stationary point of* (17) *with a high probability of approximately 90%.*

## IV. Numerical Experiments

We conducted numerical tests using the standard IEEE 68-bus system. As described in Section II, we replaced $p = 3$ SGs out of a total of 16 SGs with GFMs, positioned at nodes 63, 64, and 65. The parameters for the GFMs were configured as follows: $\tau = 0.01$, $k^{pv} = 1.00$, $k^{iv} = 5.86$, $\lambda^p = 0.05$, $\lambda^q = 0.05$. For the distributed feedback, we considered that only angle differences $\Delta\delta_{ij}$ could be observed between SGs and GFMs in different areas, within nearest-possible hops from the area boundary, and set the output matrix $C$

**Algorithm 1:** Coherency-aware Stochastic gradient-descent with max-oracle (SGDmax)

**Inputs:** A feasible and stable policy $K_0$, upper bound $\bar{\mu}$ for $\mu$, step- size $\eta$, the number of ZOPG samples $N_i$ and the number of iterations $M$.
Run steps 1-4 as of pre-tuning to select the modified coherency structure and distributed feedback.
**for** $m = 0, 1, \ldots, M-1$ **do**
    **for** $s = 1, \ldots, N_i$ **do**
        Sample the random $U_s$ where $\|U_s\| = 1$.
        Use (18) to return $\hat{\nabla}_K \mathcal{L}(K_m; U_s)$.
    **end**
    Compute the averaged stochastic gradient $\hat{G}(K_m) = \frac{1}{N_i}\sum_{s=1}^{N} \hat{\nabla}_K \mathcal{L}(K_m; U_s)$.
    Update $K_{m+1} \leftarrow K_m - \eta \hat{G}(K_m)$.
**end**
**Return:** the final iterate $K_m$.

accordingly. This captures the tie-line active power flows as they are dependent upon angular differences. Fig. 1 shows how the GFM integration impacts the coherent area boundaries where SGs at buses $53, 54, 55, 60$, and $61$ are transitioned from the nominal area grouping to form an all-new area partitioning. Fig. 2 shows how the mode-shapes (eigenvector elements corresponding to the low-frequency(slow) modes) of the resources signify this shift in area boundaries. More details on such changes can be found in the pre-print [26].

We compared two scenarios with different area partitioning strategies aimed at enhancing oscillation control, as illustrated in Fig. 1. In scenario 1 (*coherency-agnostic*), default areas were utilized as in the left figure of Fig. 1, while scenario 2 (*coherency-aware*) implemented a coherency-aware distributed design for area partitioning by following the right one. Consequently, scenario 1 involved buses $\{53, 54, 60, 61\}$ in Area 1 and $\{62, 63, 64\}$ in Area 2, whereas scenario 2 considered $\{56, 57, 58, 59\}$ in Area 1 and $\{53, 55, 60, 61\}$ in Area 2. Both scenarios underwent training using the SGDmax algorithm outlined in Algorithm 1, with the following parameters: $r = 0.01$, $M = 10$, and $\eta = 10^{-8}$. The control time step was set to $\Delta t = 0.01s$, with observations conducted from $0-100s$. At $t = 0$, we generate the impulse inputs to each SG and GFM, while $\xi_t$ follows the Gaussian distribution.

We first present training results to verify the convergence

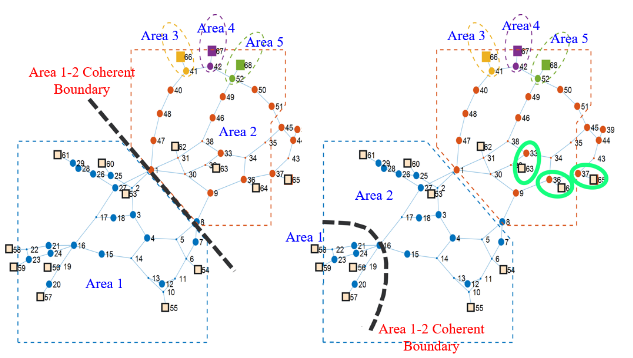


Fig. 1. Transition in Coherent Area Partitioning

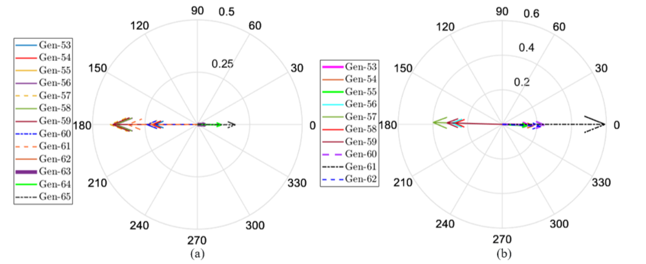

Fig. 2. Mode shape plots for (a) Nominal (Base) and, (b) Modified GFM integrated power system model

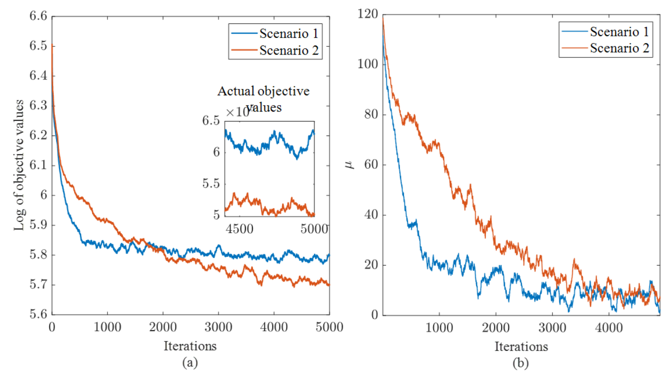


Fig. 3. (a) Training performance and (b) Average $\mu$ for Scenario 1 and Scenario 2

of Algorithm 1 and check the solution satisfies the constraint. Fig. 3(a) depicts the evolution of the objective values over time. Along with convergence, notably, the coherency-aware design consistently yields smaller objective values compared to the coherency-agnostic approach, achieving approximately a 30% enhancement. In addition, during the convergence phase, after the initial transients settle around 2000 iterations, the cost objectives of the coherency-aware design consistently remain lower bounds compared to those of the coherency-agnostic design. This observation suggests that integrating perturbed area-partitioning information into the distributed feedback significantly enhances the learning process for wide-area control. Next, Fig. 3(b) illustrates the trajectory of average $\mu$ from $N_i$ ZOPG samples during training. It is evident that starting near $\bar{\mu}$, the average $\mu$ gradually converges towards 0. This outcome suggests that as the iterations continue, we approach the optimal solution that satisfies the constraint as $\mu = 0$ when the constraint is satisfied, whereas $\mu = \bar{\mu}$ otherwise.

Using the converged policies in both scenarios, we conduct tests between scenarios 1, 2, and the one without control. Fig. 4 illustrates $\Delta\omega_i$ and $\Delta\delta_i - \Delta\delta_j$ for some typical scenarios. Specifically, Fig. 4(a) and Fig. 4(b) compare $\Delta\omega_{53}$ and $\Delta\omega_{66}$ between the one without control (yellow) and scenario 2 (black), respectively. Clearly, the control policy of scenario 2 outperforms the one without control. More specifically, the damping of a couple of slow modes 0.49 and 0.56 Hz have been improved from $1.2\%, 1.8\%$ to $13.4\%$ and $4.5\%$, respectively. Similarly, Fig. 4(c) and Fig. 4(d) depict the comparison of $\Delta\omega_{53}$ and $\Delta\omega_{66}$ between scenario 1 (orange) and scenario 2 (black). Still, scenario 2 exhibits better performance than scenario 1, indicating that coherent

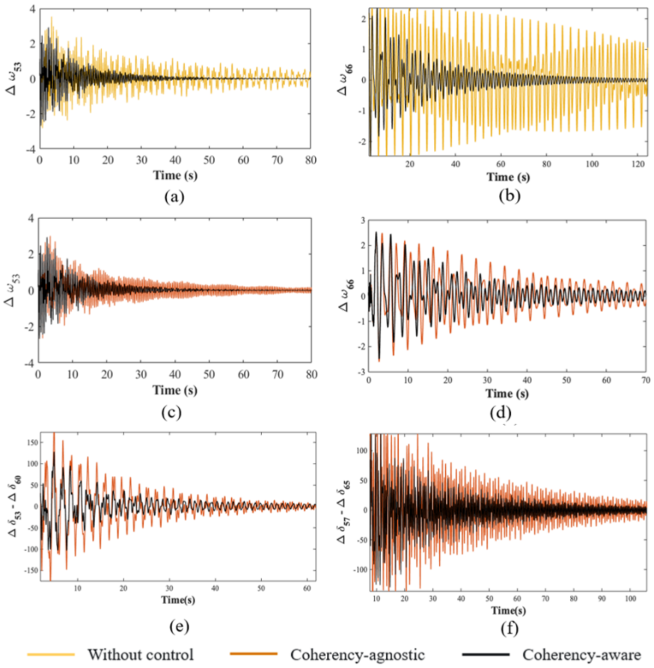


Fig. 4. Plots for *without control vs coherency-aware* (a) frequency at SG-53, (b) frequency at SG-66; plots for *coherency-agnostic vs coherency-aware* (c) frequency at SG-53, (d) frequency at SG-66, (e) angle difference between SG-53 and SG-60, and (f) angle difference between SG-57 and GFM-65.

area partitioning enhances damping performance. This is further demonstrated in Fig. 4(e) and Fig. 4(f), which illustrate $\Delta\delta_{53} - \Delta\delta_{60}$ and $\Delta\delta_{57} - \Delta\delta_{65}$, respectively. Similar to $\Delta\omega$, scenario 2 displays superior performance in reducing deviations in both cases. Since these angle differences are closely related to tie-line power flow, the result highlights the effectiveness of the proposed method in mitigating the power flow deviations across the coherent areas.

## V. Conclusions

This paper explores the significance of incorporating coherency-aware control design for mitigating persistent oscillations in large-scale power grids with substantial presence of inverter-based resources. Our proposed coherency-aware design employing distributed output feedback-based reinforcement learning has demonstrated superior training and oscillation performance compared to coherency-agnostic control designs. Future research will include high-fidelity transmission-scale models, different types of power electronics interfaced resources causing coherent clustering behavior of the bulk power grids. Additionally, robust data-driven control designs in presence of net-load uncertainties and incorporation of other risk measures, such as conditional value at risk (CVaR) will be investigated.

## References

[1] D. Pattabiraman, R. Lasseter, and T. Jahns, "Comparison of grid following and grid forming control for a high inverter penetration power system," in *IEEE Power & Energy Society General Meeting*, 2018.

[2] Y. Lin, J. H. Eto, B. B. Johnson, J. D. Flicker, R. H. Lasseter, H. N. Villegas Pico, G.-S. Seo, B. J. Pierre, and A. Ellis, "Research roadmap on grid-forming inverters," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2020.

[3] R. H. Lasseter, Z. Chen, and D. Pattabiraman, "Grid-forming inverters: A critical asset for the power grid," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 2, pp. 925–935, 2019.

[4] J. Chow and P. Kokotovic, "Time scale modeling of sparse dynamic networks," *IEEE Trans. on Automatic Control*, vol. 30(8), 1985.

[5] J. H. Chow, *Power System Coherency and Model Reduction*. Springer New York, Jan. 2013, ch. Slow Coherency and Aggregation, pp. 39–72.

[6] F. Wu and N. Narasimhamurthi, "Coherency identification for power system dynamic equivalents," *IEEE Trans. on Circuits and Systems*, vol. 30, no. 3, pp. 140–147, 1993.

[7] G. Xu and V. Vittal, "Slow coherency based cutset determination algorithm for large power systems," *IEEE Tran. on Power Systems*, vol. 25, no. 2, pp. 877–884, May 2010.

[8] A. Chakrabortty, "Wide-area damping control of power systems using dynamic clustering and tcsc-based redesigns," *IEEE Trans. on Smart Grid*, vol. 3, no. 3, pp. 1503–1514, Sep. 2012.

[9] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*. MIT press Cambridge, 1998, vol. 135.

[10] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, pp. 477–484, 2009.

[11] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[12] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, and Z. Huang, "Adaptive power system emergency control using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1171–1182, 2019.

[13] R. R. Hossain, Q. Huang, and R. Huang, "Graph convolutional network-based topology embedded deep reinforcement learning for voltage stability control," *IEEE Trans. on Power Systems*, vol. 36(5), 2021.

[14] J. Duan, Z. Yi, D. Shi, C. Lin, X. Lu, and Z. Wang, "Reinforcement-learning-based optimal control of hybrid energy storage systems in hybrid ac–dc microgrids," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5355–5364, 2019.

[15] S. Mukherjee, A. Chakrabortty, H. Bai, A. Darvishi, and B. Fardanesh, "Scalable designs for reinforcement learning-based wide-area damping control," *IEEE Trans. on Smart Grid*, vol. 12(3), pp. 2389–2401, 2021.

[16] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning." in *International Conference on Learning Representations*, 2016.

[17] K.-B. Kwon, S. Mukherjee, T. L. Vu, and H. Zhu, "Risk-constrained reinforcement learning for inverter-dominated power system controls," *IEEE Control Systems Letters*, vol. 7, pp. 3854–3859, 2023.

[18] P. Kundur, *Power System Stability and Control*. McGraw-Hill, 1994.

[19] W. Du, F. K. Tuffner, K. P. Schneider, R. H. Lasseter, J. Xie, Z. Chen, and B. Bhattarai, "Modeling of grid-forming and grid-following inverters for dynamic simulation of large-scale distribution systems," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2035–2045, 2021.

[20] S. Kundu and K. Kalsi, "Transient safety filter design for grid-forming inverters," in *American Control Conference (ACC)*, 2020, pp. 1299–1304.

[21] J. H. Chow, "Time-scale modeling of dynamic networks with applications to power systems," *Springer-Verlag: Lecture Notes in Control and Information Sciences*, 1982.

[22] C. Davis and W. M. Kahan, "The rotation of eigenvectors by a perturbation. III," *SIAM Journal on Numerical Analysis*, vol. 7, no. 1, pp. 1–46, 1970.

[23] B. Hunter and T. Strohmer, "Performance analysis of spectral clustering on compressed, incomplete and inaccurate measurements," *arXiv preprint arXiv:1011.0997*, 2010.

[24] A. Tsiamis, D. S. Kalogerias, L. F. O. Chamon, A. Ribeiro, and G. J. Pappas, "Risk-Constrained Linear-Quadratic Regulators," in *59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 3040–3047.

[25] K.-b. Kwon, L. Ye, V. Gupta, and H. Zhu, "Model-free Learning for Risk-constrained Linear Quadratic Regulator with Structured Feedback in Networked Systems," *arXiv preprint arXiv:2204.01779*, 2022.

[26] S. Mukherjee, R. Hossain, K. Chatterjee, S. Kundu, K. Kwon, S. Nekkalapu, and M. Elizondo, "Structural impact of grid-forming inverters on power system coherency," *arXiv:2405.09675*, 2024.