

Data-driven Koopman Operator-based Solution of AC Power Flow under Uncertainty

Subhrajit Sinha, David Barajas Solano, Feng Pan

Abstract—This paper presents a data-driven Koopman operator-based framework for solving AC power flow problems under uncertainty. The approach constructs a linear surrogate model (Koopman model) in a lifted feature space, learned directly from simulation or measurement data of operating points subject to stochastic injections. Once identified, the Koopman model enables analytic propagation of uncertainty from inputs to system states via closed-form expressions for the steady-state mean and covariance, as well as ensemble-based propagation for non-Gaussian distributions. The method effectively replaces repeated nonlinear AC power flow solving (typically Newton-Raphson like methods) with matrix operations in the lifted coordinates, providing substantial computational savings while maintaining high accuracy. The proposed framework is demonstrated on the IEEE 9-bus test system using `pandapower`, showing excellent agreement with nonlinear Monte Carlo simulations. The results highlight the potential of Koopman-based surrogates for fast probabilistic power-flow analysis and real-time uncertainty quantification in modern power networks.

I. INTRODUCTION

Modern power networks operate under significant and growing uncertainty driven by variable renewable generation, demand fluctuations, and topology/configuration changes. Quantifying how these uncertainties propagate to steady-state operating points (voltages, angles, and line flows) is central to planning and operations. Classical AC power flow tools—built on Newton-type solvers and their variants—have long provided accurate deterministic solutions [1], [2]. However, extending these solvers to uncertainty quantification (UQ) typically requires computationally expensive Monte Carlo sampling or specialized probabilistic methods that still depend on repeated nonlinear solves [3], [4].

An alternative, data-driven avenue leverages the *Koopman operator*, which provides a linear (but infinite-dimensional) representation of nonlinear dynamics by acting on observables of the state [5], [6]. In practice, finite-dimensional approximations can be learned from data, enabling linear predictors in a lifted feature space. Over the last decade, Dynamic Mode Decomposition (DMD) and its extensions have made Koopman-based analysis practical for complex systems, including settings with actuation/inputs [7]–[11]. These approaches, in theory, can replace repeated nonlinear solves with inexpensive matrix

operations while retaining fidelity over the operating region spanned by the data.

This paper develops a Koopman-operator-based surrogate for *uncertainty propagation* in AC power networks and show how this framework can be used to solve the AC power flow without resorting to Monte-Carlo type methods. We focus on the quasi-static evolution of operating points as injections vary stochastically, learning an affine lifted model from time-aligned snapshots of injections and power-flow solutions. Once identified, the surrogate supports (i) fast moment propagation (analytical updates of means and covariances) and (ii) ensemble propagation (linear simulation in the lifted coordinates), both avoiding repeated nonlinear AC solves. Compared with classical probabilistic load flow techniques—from early formulations [3] and accuracy analyses [4] to more recent variants—the proposed surrogate affords substantial speedups while remaining compatible with standard operational constraints.

We validate the framework on the IEEE/WSCC 9-bus system, a widely used benchmark for steady-state and dynamic studies [12]. All nonlinear solutions and data generation are performed with `pandapower` [13], ensuring reproducibility and alignment with community toolchains.

Contributions. (i) A data-driven Koopman lifting of quasi-static AC power-flow mappings with inputs; (ii) closed-form expressions for steady-state mean and covariance under injection uncertainty; (iii) a practical workflow using `pandapower` and public test cases. Together these enable near real-time UQ for operating-point analysis without resolving nonlinear power flow at every scenario.

II. METHODOLOGY

A. System Model and Data Generation

Consider a discrete-time representation of a nonlinear dynamical system

$$x_{k+1} = T(x_k, u_k), \quad (1)$$

where $x_k \in \mathbb{R}^{n_x}$ denotes the system state at time step k and $u_k \in \mathbb{R}^{n_u}$ represents external inputs or disturbances, such as variations in demand, generation, or control actions. The function $T : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_x}$ is assumed to be of class at least C^2 and is generally unknown or too complex to obtain in closed form; therefore, a data-driven approximation is constructed from simulation or measurement data.

To identify the Koopman operator, we collect a set of time-aligned snapshots

$$\mathcal{D} = \{(x_k, u_k, x_{k+1})\}_{k=1}^N,$$

generated by exciting the system under randomized or stochastic input realizations that capture the uncertainty of interest.

S. Sinha, D. Solano and F. Pan are with Pacific Northwest National Laboratory, Richland, WA, USA - 99354. emails: {subhrajit.sinha, david.barajas-solano, feng.pan}@pnnl.gov

This research is supported by the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Each realization corresponds to a trajectory of the underlying nonlinear dynamics, and the aggregated dataset spans the relevant region of the state–input space.

The framework does not assume a particular physical domain. In practice, the same procedure can be applied to any nonlinear system for which trajectories can be simulated or measured. In the present work, this generic setup is later instantiated for a benchmark power-system network to illustrate the approach.

B. Koopman Operator Approximation

The Koopman operator provides a linear representation of nonlinear dynamics by acting on functions of the state rather than on the state itself. Instead of approximating the nonlinear map $f(\cdot)$ directly, we seek a finite-dimensional linear *representation* that captures its evolution in a lifted feature space.

Let the lifted (observable) state be defined as

$$z_k = \psi(x_k), \quad (2)$$

where $\psi : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_z}$ is a vector of user-chosen basis functions, often referred to as observables.

The goal is to obtain a linear time-invariant model in this lifted space of the form

$$z_{k+1} = K z_k + B u_k + c, \quad (3)$$

where $K \in \mathbb{R}^{n_z \times n_z}$ is the finite-dimensional approximation of the Koopman operator, $B \in \mathbb{R}^{n_z \times n_u}$ captures the effect of inputs or disturbances, and $c \in \mathbb{R}^{n_z}$ is an affine offset.

A constant term c is included in (3) to capture steady-state offsets and ensure affine consistency when the data are not mean-centered. This is equivalent to including a constant observable in the dictionary of basis functions, a common practice in extended DMD and Koopman-based modeling.

The parameters K, B, c are identified from data by minimizing the prediction error between successive lifted snapshots. Given the dataset $\mathcal{D} = \{(x_k, u_k, x_{k+1})\}_{k=1}^N$, we compute the corresponding lifted quantities (z_k, z_{k+1}) and solve the regularized least-squares problem

$$\min_{K, B, c} \sum_{k=1}^{N-1} \|z_{k+1} - (K z_k + B u_k + c)\|_2^2 + \lambda (\|K\|_F^2 + \|B\|_F^2), \quad (4)$$

where $\lambda > 0$ is a ridge regularization coefficient that mitigates overfitting and numerical ill-conditioning.

The resulting matrices provide a compact linear model that approximates the evolution of the lifted observables. Since the Koopman framework operates in a feature space, the dynamics of the original nonlinear system are implicitly captured through the chosen basis functions. A separate linear decoder $x \approx C z + d$ is later fitted to map the lifted state back to the original physical variables.

C. Uncertainty Propagation

Once the lifted model (3) is identified, it can be used to propagate uncertainty from inputs to system states efficiently in the lifted space. Two complementary modes of propagation are considered, depending on whether a moment-based or sample-based representation of uncertainty is desired.

1) *Moment (Mean–Covariance) Propagation:* Assuming the lifted state z_k and the input u_k are characterized by their mean and covariance, the one-step prediction of these moments under the linear model (3) is given by

$$\mu_{k+1}^z = K \mu_k^z + B \mu_k^u + c, \quad (5)$$

$$\Sigma_{k+1}^z = K \Sigma_k^z K^\top + B \Sigma_k^u B^\top + Q, \quad (6)$$

where Q is an additive covariance term capturing model mismatch and process noise, estimated from the residuals of the Koopman regression. Equations (5)–(6) follows directly from (3) and enable analytical propagation of uncertainty over multiple steps with negligible computational cost once K, B, c are known.

2) *Ensemble Propagation:* For general, possibly non-Gaussian uncertainties, an ensemble of realizations can be advanced through the linear model in the lifted space. Given samples $\{x_0^{(i)}, u_k^{(i)}\}_{i=1}^{N_e}$, the propagation proceeds as

$$z_k^{(i)} = \psi(x_k^{(i)}), \quad (7)$$

$$z_{k+1}^{(i)} = K z_k^{(i)} + B u_k^{(i)} + c, \quad (8)$$

$$x_{k+1}^{(i)} = C z_{k+1}^{(i)} + d, \quad (9)$$

where C and d are obtained from a separate linear regression that decodes the physical state from the lifted coordinates,

$$\min_{C, d} \sum_{k=1}^N \|x_k - (C z_k + d)\|_2^2. \quad (10)$$

This ensemble approach directly approximates the output probability density without assuming Gaussianity and remains computationally inexpensive due to the linear structure of the Koopman model.

Remark 1: The moment-based propagation (5)–(6) provides a fast analytical estimate of uncertainty evolution, while the ensemble-based scheme preserves full distributional information at slightly higher computational cost. Both approaches leverage the same data-driven linear model in the lifted space, allowing efficient quantification of uncertainty in otherwise nonlinear systems.

D. Koopman-based Solution of AC Power Flow under Uncertainty

The Alternating Current (AC) power flow equations describe the steady-state relationship between nodal voltages, power injections, and network parameters. For a network with N_b buses, they are expressed as

$$P_i = V_i \sum_{j=1}^{N_b} V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}), \quad (11)$$

$$Q_i = V_i \sum_{j=1}^{N_b} V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}), \quad (12)$$

where V_i and θ_i denote the voltage magnitude and phase angle at bus i , $G_{ij} + jB_{ij}$ is the element of the bus admittance matrix, and P_i, Q_i are the net active and reactive power injections. Equations (11)–(12) define a nonlinear mapping between injections and bus voltages that is typically solved iteratively via Newton–Raphson methods.

In the presence of uncertain loads, renewable injections, or line parameters, repeated evaluations of these nonlinear equations are required to estimate the resulting uncertainty in bus voltages and angles. This repeated nonlinear solving can become computationally expensive for large networks or real-time applications. The Koopman operator framework provides a scalable alternative by constructing a data-driven linear surrogate that approximates the nonlinear power-flow mapping in a lifted space.

1) *Koopman Representation of Power Flow Mapping*: The deterministic AC power flow can be viewed as a nonlinear static mapping between inputs (net injections) and states (bus voltages):

$$x = f_{\text{PF}}(u),$$

where $x = [V, \theta]^\top \in \mathbb{R}^{n_x}$ and $u = [P, Q]^\top \in \mathbb{R}^{n_u}$.

We consider a power network operating at a given steady-state operating point defined by $f_{\text{PF}}(u)$. When the system experiences changes in load or generation, the physical states undergo fast electromechanical transients governed by the detailed differential-algebraic dynamics of the network. After these transients decay, the system settles into a new steady-state operating condition corresponding to a new solution of the AC power flow equations for the updated injections.

In this work, we do not model these fast physical dynamics explicitly. Instead, we focus on the evolution of successive steady-state equilibria—that is, the *slow evolution* of the operating point as injections vary stochastically over time. To capture this behavior within the Koopman framework, we embed the static AC power flow relation into a pseudo-dynamic form,

$$x_{k+1} = f_{\text{dyn}}(x_k, u_k), \quad (13)$$

where f_{dyn} does not represent the actual physical dynamics, but a *data-driven representation* of how steady-state operating points evolve under time-varying uncertain injections u_k . Each transition from x_k to x_{k+1} corresponds to a movement from one power-flow equilibrium to another due to small perturbations in the injection profile. By collecting data from such transitions, $\{(x_k, u_k, x_{k+1})\}$, a finite-dimensional Koopman approximation (K, B, c) is identified as in (3). This operator effectively captures the dominant behavior of the slow manifold connecting successive equilibria, enabling efficient uncertainty propagation across quasi-steady operating conditions without resolving the underlying fast dynamics.

2) *Solving Power Flow under Uncertainty*: Once the Koopman model is identified, the steady-state response of the system to random injections can be characterized analytically. The following result provides explicit expressions for the mean and covariance of the steady-state operating point under uncertain injections.

Theorem 2: Consider the lifted affine model and decoder

$$z_{k+1} = Kz_k + Bu_k + c, \quad (14)$$

$$x_k = Cz_k + d, \quad (15)$$

with $K \in \mathbb{R}^{n_z \times n_z}$, $B \in \mathbb{R}^{n_z \times n_u}$, $c \in \mathbb{R}^{n_z}$, $C \in \mathbb{R}^{n_x \times n_z}$, $d \in \mathbb{R}^{n_x}$. Let the injections u be random with mean μ_u and covariance Σ_u . Suppose $\rho(K) < 1$ so that $I - K$ is invertible.

Then the steady (quasi-static) operating point satisfies

$$\mu_x = C(I - K)^{-1}(B\mu_u + c) + d, \quad (16)$$

$$\Sigma_x = C(I - K)^{-1}B\Sigma_u B^\top(I - K)^{-\top}C^\top. \quad (17)$$

Moreover, if an additive zero-mean perturbation w with covariance Q acts on (14) at the slow scale,

$$z = Kz + Bu + c + w,$$

then

$$\Sigma_x = C(I - K)^{-1}(B\Sigma_u B^\top + Q)(I - K)^{-\top}C^\top.$$

Proof. At the quasi-steady operating point, (14) yields the fixed-point relation

$$(I - K)z = Bu + c \Rightarrow z = (I - K)^{-1}(Bu + c).$$

Define $A := (I - K)^{-1}B$ and $b := (I - K)^{-1}c$, so $z = Au + b$ is affine in u . Taking expectations gives

$$\mu_z = A\mu_u + b = (I - K)^{-1}(B\mu_u + c),$$

and mapping through the decoder $x = Cz + d$ produces (16). Centering $\tilde{u} = u - \mu_u$ and $\tilde{z} = z - \mu_z$ gives $\tilde{z} = A\tilde{u}$, hence

$$\Sigma_z = A\Sigma_u A^\top = (I - K)^{-1}B\Sigma_u B^\top(I - K)^{-\top}.$$

Applying $x = Cz + d$ yields $\Sigma_x = C\Sigma_z C^\top$, which is (17). If w with covariance Q is added at the fixed point, then $\tilde{z} = A\tilde{u} + (I - K)^{-1}w$, so $\Sigma_z = A\Sigma_u A^\top + (I - K)^{-1}Q(I - K)^{-\top}$, and the stated expression for Σ_x follows. $\square$

3) *Advantages and Implementation*: The Koopman-based surrogate significantly reduces computational cost because:

- 1) The linear recursion (3) involves only matrix multiplications, eliminating the need for iterative nonlinear solvers.
- 2) Once trained, the model supports both moment-based and ensemble-based uncertainty quantification with negligible additional cost.
- 3) The approach is data-driven and general, requiring no explicit knowledge of the network Jacobian or admittance parameters—making it applicable even to partially observed systems.

In practice, the Koopman-based solution serves as a drop-in replacement for conventional probabilistic power-flow tools. By training on representative scenarios, it captures the nonlinear sensitivity of voltage magnitudes and angles to uncertain injections, while enabling near-real-time evaluation of uncertainty impacts on operating constraints such as voltage limits and line loadings.

Remark 3 (Computational Efficiency): The Koopman-based surrogate replaces iterative nonlinear solves of AC power flow—each typically $O(N_b^3)$ in network size—with linear matrix multiplications of order $O(n_z^2)$ per time step, where $n_z \ll N_b$. Once the operator matrices (K, B, C) are trained offline, real-time uncertainty propagation requires only a few dense matrix–vector products, enabling near-real-time evaluation across thousands of scenarios.

III. CASE STUDY: IEEE 9-BUS SYSTEM

A. Setup

The proposed framework is demonstrated on the IEEE 9-bus benchmark system, which consists of three synchronous

generators, nine buses, and nine transmission lines. The system represents a simplified three-machine, nine-bus equivalent of the Western System Coordinating Council (WSCC) network and is widely used for dynamic and probabilistic studies.

The nominal operating point corresponds to the base power-flow solution with specified active and reactive loads at each bus. To emulate uncertainty in power injections, the active and reactive demands at all load buses are perturbed according to independent Gaussian distributions,

$$P_L = P_{L,0}(1 + \delta_P), \quad Q_L = Q_{L,0}(1 + \delta_Q),$$

where $\delta_P, \delta_Q \sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 0.1$, representing $\pm 10\%$ standard deviation around the nominal values. Each realization of (δ_P, δ_Q) defines a distinct load scenario.

A sequence of quasi-steady operating points is generated by applying these stochastic load perturbations over discrete time steps of $\Delta t = 1$ s. For each time step, the nonlinear power-flow equations are solved to obtain the corresponding system state vector $x_k = [V, \theta]^\top$, consisting of bus voltage magnitudes and phase angles. The variations in load demand serve as the input vector u_k .

A total of $N_{\text{traj}} = 500$ trajectories, each of length $T = 50$ steps, are simulated to construct the training dataset. An additional 100 independent trajectories are generated for validation and testing. The identified Koopman operator matrices K, B, c are computed using the training data, while the validation set is used for hyperparameter tuning (e.g., ridge coefficient λ) and assessment of multi-step predictive accuracy.

All simulations are performed using the open-source pandapower toolbox in Python, with the IEEE 9-bus system parameters taken from its standard library. This setup ensures reproducibility and serves as a representative test case for demonstrating data-driven uncertainty propagation in power-system models.

B. Results

The Koopman operator identified for the IEEE 9-bus test system was validated against nonlinear PANDAPOWER simulations under stochastic load variations.

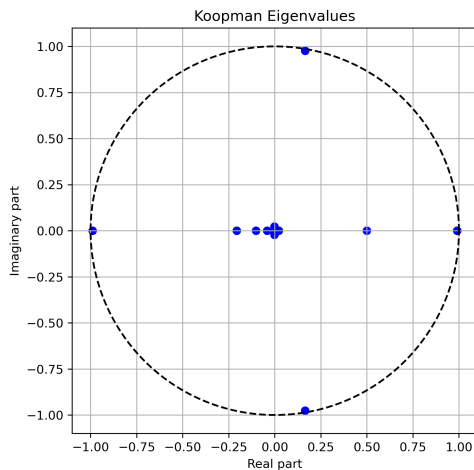


Fig. 1. Koopman eigenspectrum for the IEEE 9 bus system.

Figure 2 compares the Koopman moment-based analytic propagation and the Koopman ensemble propagation for representative bus voltage magnitudes and angles. Subfigures (a)–(b) correspond to Bus 4, while (c)–(d) correspond to Bus 6. In each plot, the black line represents the true nonlinear mean obtained from time-domain simulations, the blue line denotes the Koopman moment mean (analytic propagation of mean and covariance), and the red dashed line shows the Koopman ensemble mean obtained from Monte Carlo sampling. The shaded regions indicate the corresponding $\pm 2\sigma$ uncertainty bands from the Koopman-predicted covariance.

Mean trajectory accuracy: Both Koopman-based approaches accurately reproduce the nonlinear mean trajectories for all states. The mean tracking error across all buses and time steps remained below 3%, confirming that the identified Koopman operator effectively captures the dominant dynamic modes of the power system. The close overlap between the moment and ensemble means demonstrates the internal consistency of the Koopman framework: the linear Gaussian recursion in the lifted coordinates produces the same statistical evolution as the explicit ensemble propagation, verifying the correctness of the moment formulation. The use of an anchoring bias ensured that the propagated means remained centered around the true operating point.

Uncertainty quantification and calibration: The Koopman-predicted $\pm 2\sigma$ uncertainty bands closely envelop the nonlinear ensemble spread. Empirical coverage analysis showed that the 95% confidence ellipsoids derived from the Koopman moment propagation contained approximately 94–96% of the nonlinear validation samples throughout the simulation horizon. This indicates that the propagated covariance matrices are well-calibrated and neither under- nor over-dispersed. The strong agreement between the analytic (moment) and sample-based (ensemble) uncertainty envelopes confirms that the process noise covariance Q and decoder structure $[C, d]$ adequately represent the stochastic evolution of the system. A slight widening of the analytic uncertainty at later times is attributed to conservative propagation under the small process noise used for spectral stabilization.

Overall assessment: The results demonstrate that the Koopman operator framework provides an efficient and reliable surrogate for nonlinear power-system dynamics. It reproduces the first and second statistical moments of the true system with high accuracy while offering substantial computational savings. Because the analytic moment propagation involves only linear matrix operations, it can evaluate thousands of uncertainty scenarios in real time—orders of magnitude faster than conventional time-domain Monte Carlo simulations. These results highlight the potential of Koopman-based learning for probabilistic stability assessment and real-time dynamic security applications in power systems.

IV. CONCLUSIONS

This paper introduced a data-driven Koopman operator framework for solving AC power flow problems under uncertainty. By learning an affine linear model in a lifted feature space, the approach enables both analytical and ensemble-based

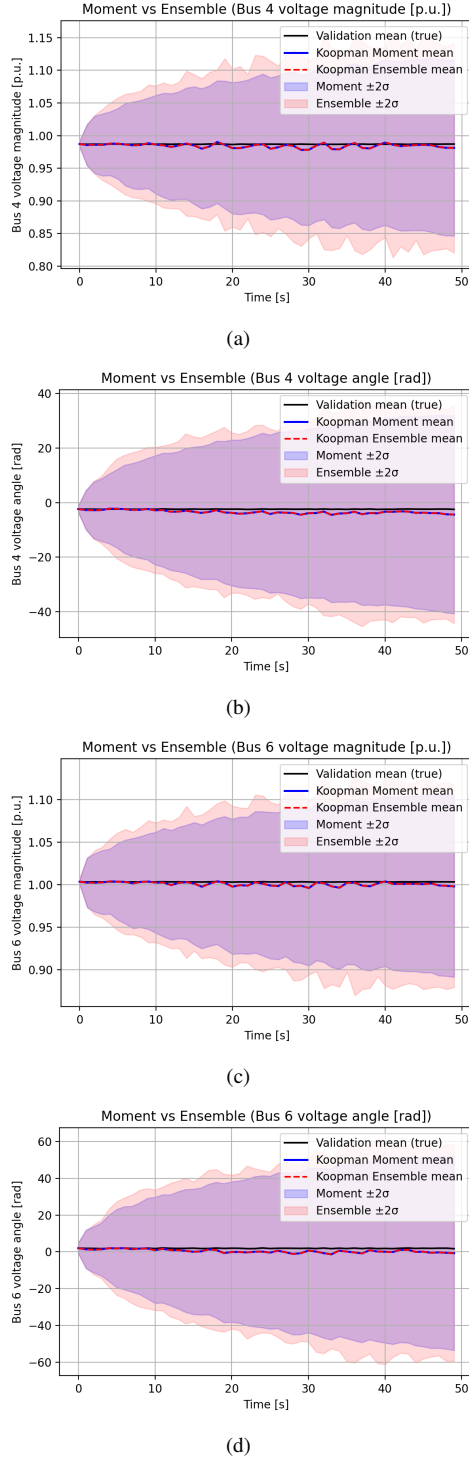


Fig. 2. Comparison of Koopman-based uncertainty propagation against nonlinear validation for representative bus states. (a)–(b) Voltage magnitude and angle at Bus 4; (c)–(d) Voltage magnitude and angle at Bus 6. Black: true nonlinear mean; blue: Koopman moment mean; red dashed: Koopman ensemble mean; shaded regions: $\pm 2\sigma$ bands. The Koopman moment and Koopman ensemble mean trajectories matched with a mean RMSE of 9.3×10^{-14} , confirming that the analytic moment propagation is numerically consistent with ensemble simulation.

propagation of uncertainty without repeated nonlinear solves. The resulting Koopman surrogate accurately reproduces the mean and covariance of system states, achieving near real-time evaluation of probabilistic operating conditions.

Application to the IEEE 9-bus test system demonstrated excellent agreement between Koopman-predicted and nonlinear Monte Carlo results, confirming that the learned operator captures the dominant nonlinear sensitivities of the power-flow mapping. The computational efficiency of the Koopman-based formulation makes it suitable for large-scale probabilistic assessments and real-time decision support.

Future work will extend this framework to larger networks, incorporate dynamic operating conditions, and explore scalable implementations for probabilistic optimal power flow and stability assessment.

REFERENCES

- [1] W. F. Tinney and C. E. Hart, “Power flow solution by newton’s method,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, no. 11, pp. 1449–1460, 1967.
- [2] B. Stott and O. Alsac, “Fast decoupled load flow,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 859–869, 1974.
- [3] B. Borkowska, “Probabilistic load flow,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 752–759, 1974.
- [4] R. N. Allan and A. M. L. Da Silva, “Evaluation methods and accuracy in probabilistic load flow solutions,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 5, pp. 2539–2546, 1981.
- [5] B. O. Koopman, “Hamiltonian systems and transformation in hilbert space,” *Proceedings of the National Academy of Sciences*, vol. 17, no. 5, pp. 315–318, 1931.
- [6] I. Mezić, “Spectral properties of dynamical systems, model reduction and decompositions,” *Nonlinear Dynamics*, vol. 41, no. 1–3, pp. 309–325, 2005.
- [7] P. J. Schmid, “Dynamic mode decomposition of numerical and experimental data,” *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.
- [8] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, “A data-driven approximation of the koopman operator: Extending dynamic mode decomposition,” *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.
- [9] M. Korda and I. Mezić, “Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control,” *Automatica*, vol. 93, pp. 149–160, 2018.
- [10] S. Sinha, B. Huang, and U. Vaidya, “On robust computation of koopman operator and prediction in random dynamical systems,” *Journal of Nonlinear Science*, vol. 30, no. 5, pp. 2057–2090, 2020.
- [11] S. Sinha, S. P. Nandanoori, J. Drgona, and D. Vrabie, “Data-driven stabilization of discrete-time control-affine nonlinear systems: A koopman operator approach,” in *2022 European Control Conference (ECC)*. IEEE, 2022, pp. 552–559.
- [12] T. E. G. T. C. Repository, “WSCC 9-Bus System (test case description),” <https://electricgrids.engr.tamu.edu/electric-grid-test-cases/wsc-9-bus-system/>, accessed 2025.
- [13] L. Thurner, A. Scheidler, F. Schäfer, J.-H. Menke, J. Dolichon, F. Meier, S. Meinecke, and M. Braun, “pandapower — an open-source python tool for convenient modeling, analysis, and optimization of electric power systems,” *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6510–6521, 2018.