Safe Reinforcement Learning for Grid-forming Inverter Based Frequency Regulation with Stability Guarantee

Fangxing(Fran) Li, The University of Tennessee Knoxville
Buxin She, Pacific Northwest National Laboratory

07/25/2024

Introduction

Background and Objective

Background:

Future Inverter-based Power Grid

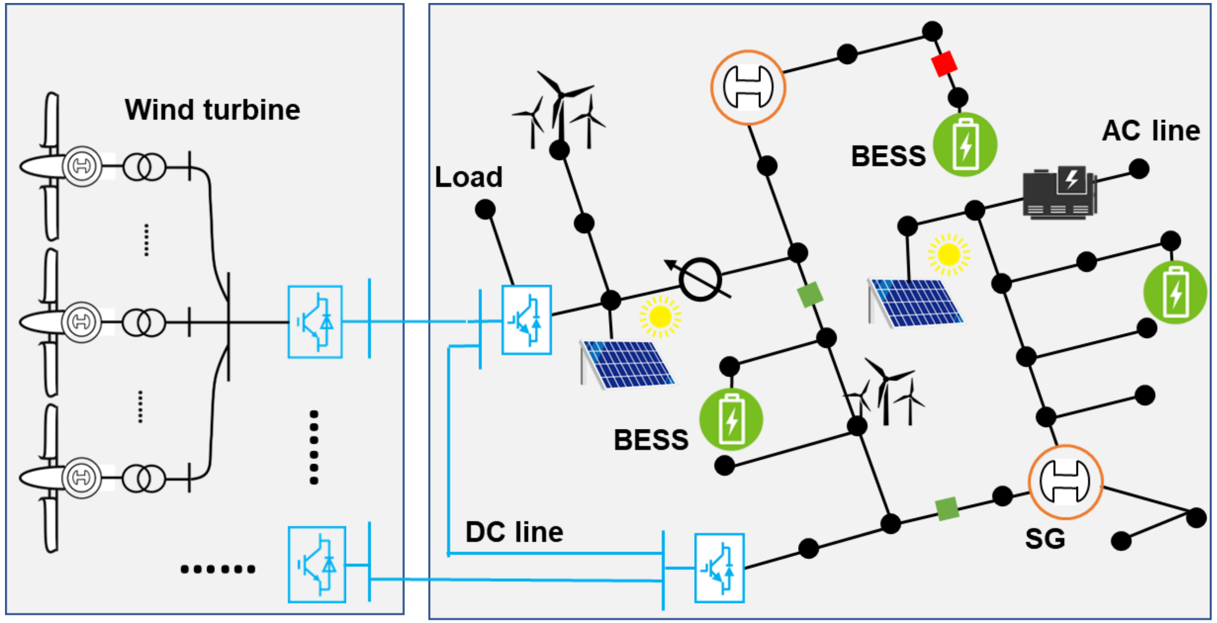


Figure 1. Diagram future inverter-dominant power grid

- Higher uncertainty
- Faster dynamics of inverter-based resources (IBRs)
- Elements that are difficult to model
- Model and parameter accessibility/Privacy

Background:

Reinforcement Learning in Power Grid Control

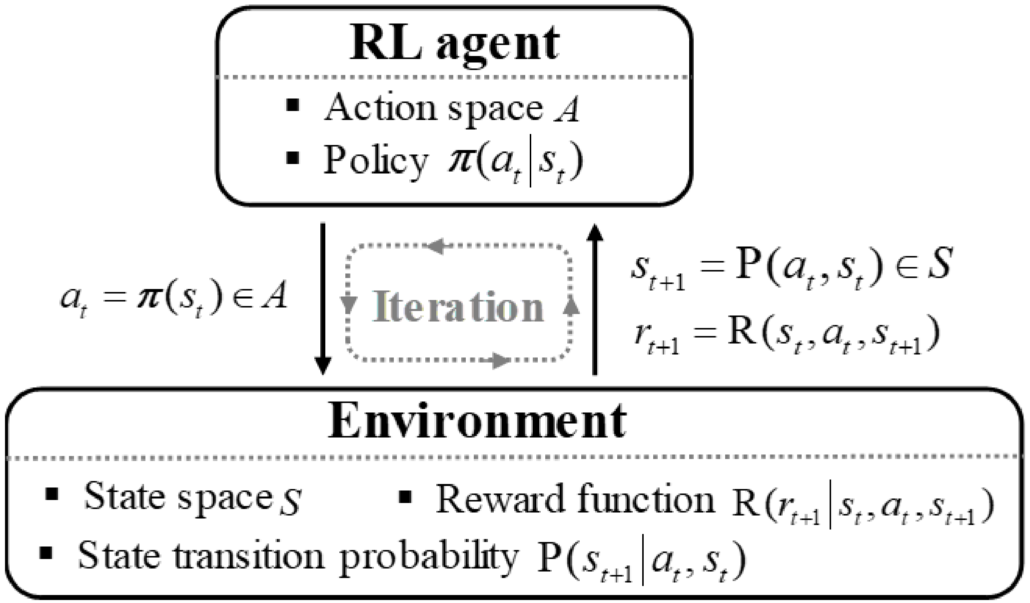


Figure 2. Diagram Reinforcement Learning

Application Domain

- Frequency regulation
- Voltage regulation
- System operation and planning

Application Challenges

- Environment
- Explainability
- Generalization
- Static Security
- Dynamic Stability

Objective:

Safe RL for Frequency Regulation with uncertain system parameters [1]

- Frequency regulation based on grid-forming inverters
- Stability guarantee
- Parameter uncertainty – dynamic (virtual) system inertia and damping

[1] H. Shuai, B. She, J. Wang and F. Li, "Safe Reinforcement Learning for Grid-Forming Inverter Based Frequency Regulation with Stability Guarantee," in *Journal of Modern Power Systems and Clean Energy*, doi: 10.35833/MPCE.2023.000882.

Methodology

Safe RL for Inverter Control and Frequency Regulation

Overview:

Optimal Frequency Regulation/Control Policy

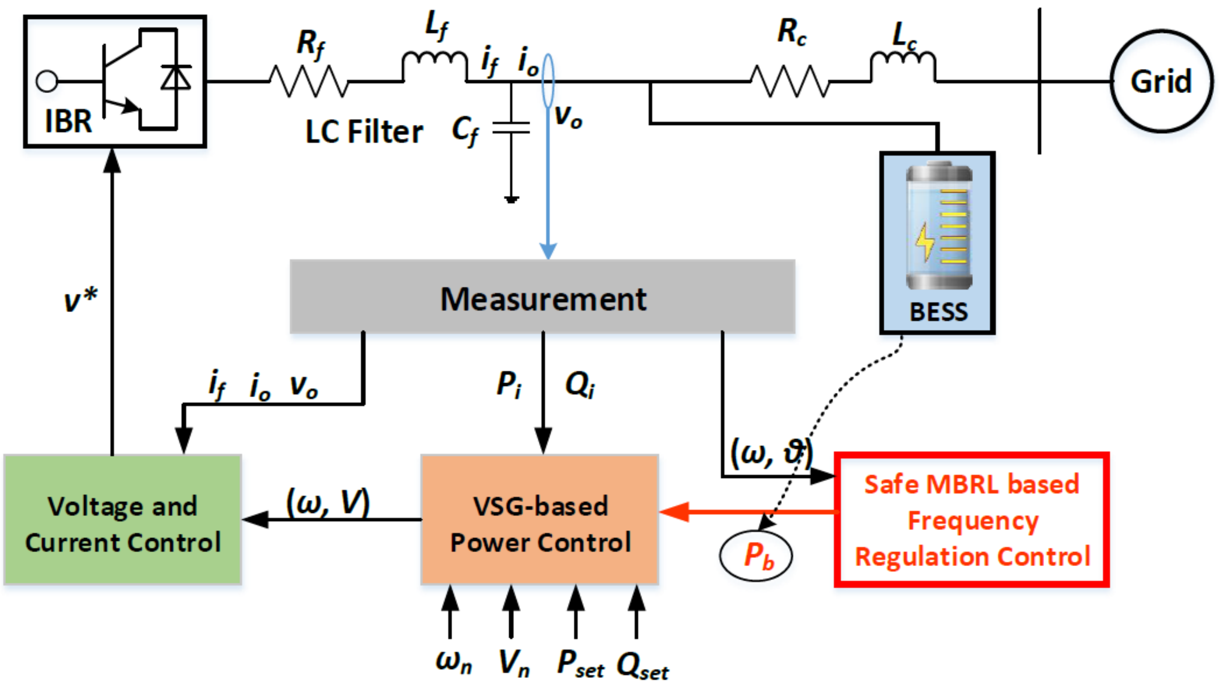


Figure 3. Diagram of frequency GFM inverter based primary frequency control

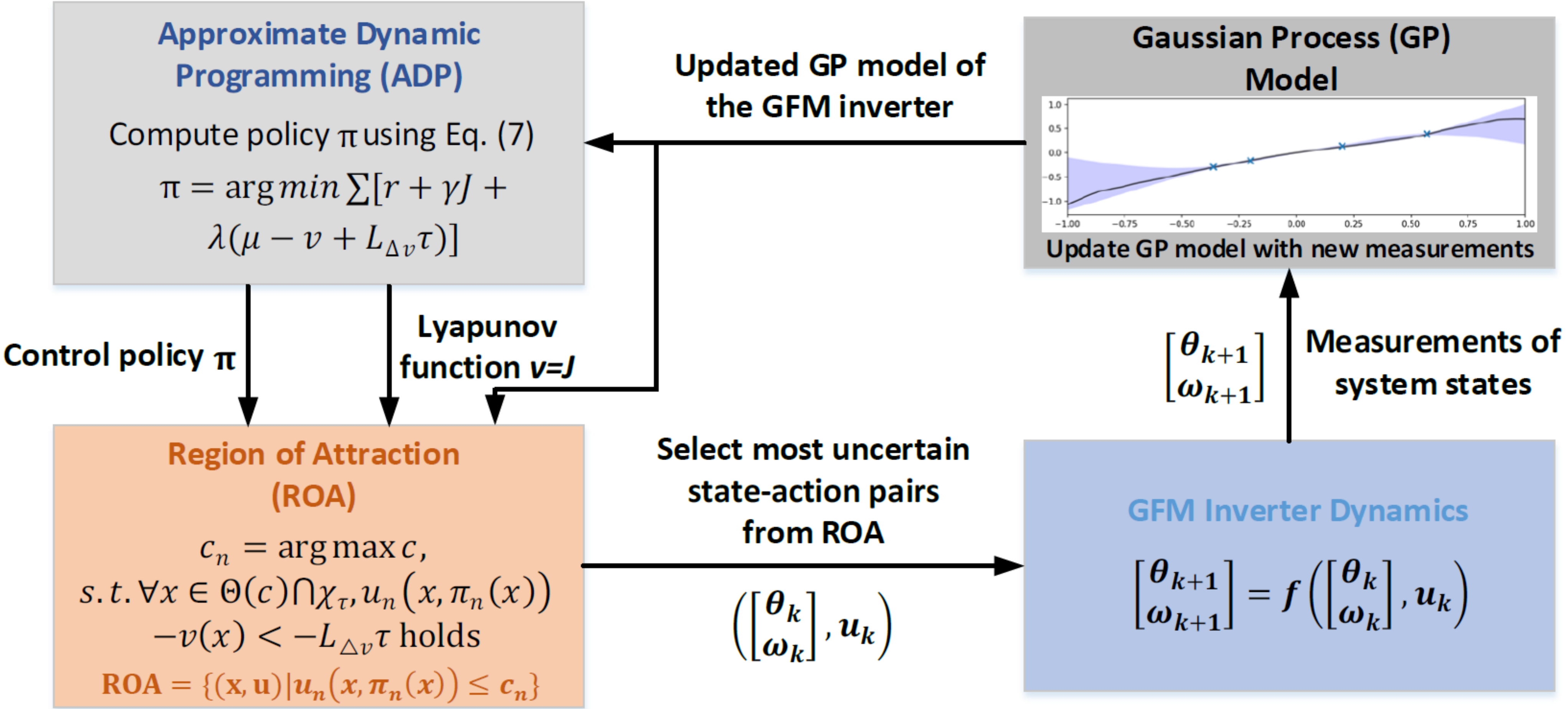
$$\min_u u^T R u + x^T Q x$$

$$\text{s.t.}\begin{cases} \dfrac{d\theta}{dt} = \omega \\ M\dfrac{d\omega}{dt} = P_{set} - P_i - D\omega - u(\theta,\omega) \\ P_i = \sum_{j\in\{i,g\}} V_i V_j \left[B_{ij}\sin(\theta_i-\theta_j) + G_{ij}\cos(\theta_i-\theta_j)\right] \\ \underline{u} \le u = f(x) \le \overline{u} \\ \boxed{x \in \Theta} \longrightarrow \text{Guaranteed stability} \end{cases}$$

Where $u(\cdot)$ is the optimal control policy and Θ is the region of attraction (ROA).

Overview:

Proposed Safe RL Strategy embedding GP Model



Key Component 1:

Gaussian Process Model to Quantify System Uncertainty

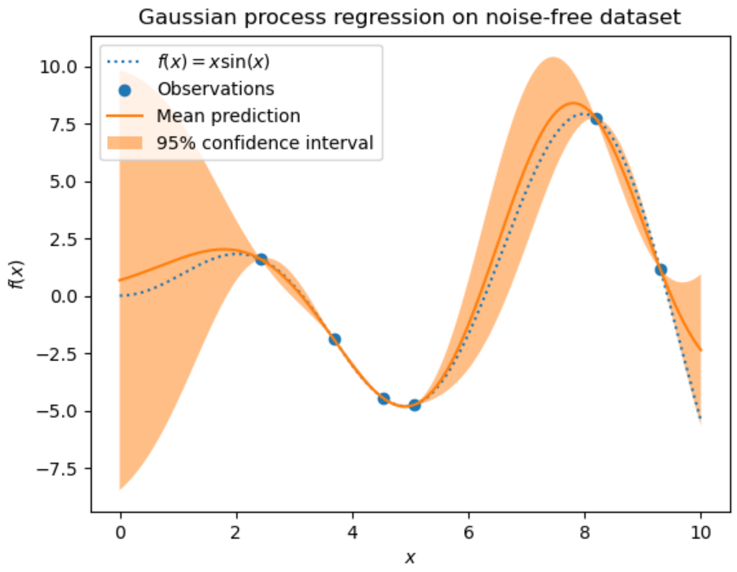


Figure 5. Diagram of Gaussian Process [1]

GPs model can approximate and quantify the uncertainty in the dynamic system model under a predefined confidence level.

- **Linear kernel**

$$k_L(x,x') = x^T x'$$

- **Mat´ern kernel**

$$k_M(x,x') = \frac{1}{\Gamma(v)2^{v-1}}\left(\frac{\sqrt{2v}}{l}d(x,x')\right)^v K_v\left(\frac{\sqrt{2v}}{l}d(x,x')\right)$$

[1] Downloaded from: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy_targets.html

Key Component 2:

Stability Guarantee based on Lyapunov Function

Lyapunov function *V*

$$\begin{cases} V(0) = 0 \\ V(\Delta x) > 0, \forall x \in D \backslash \{0\} \\ \dot{V}(\Delta x) = \nabla V \dot{\Delta x} \leq 0, \forall \Delta x \in D \backslash \{0\} \end{cases}$$

➡ Certified a subset of ROA

$$\mathcal{S} = \left\{ \Delta x \in \mathbb{R}^n \mid V(\Delta x) \leq \alpha \right\} \quad \text{with} \quad \mathcal{S} \subset DOA$$

➡ Positive invariant domain in a ROA

$$V = J(x) = r\left(x, \pi_W(x)\right) + \gamma J\left(u_{n-1}\left(x, \pi_W(x)\right)\right)$$

Find policy $\pi_W(x)$ such that *V* is decreasing

Key Component 3:

Policy Update through Adaptive Dynamic Programming

$$\pi_n = \underset{\pi_W \in \Pi_P}{arg\,min} \sum_{x \in X_r} \boxed{r\left(x, \pi_W(x)\right) + \gamma J_{\pi_W}\left(u_{n-1}\left(x, \pi_W(x)\right)\right)} + \boxed{\lambda\left(u_n\left(x, \pi_W(x)\right) - v(x) + L\Delta v \tau\right)}$$

Piecewise Linear Bellman Equation

Maximize the accumulated rewards

Lagrange Relaxation to guarantee the third criteria (stability)

Guarantee stability

Safely improved by approximate dynamic programming (ADP) to find the optimal control policy with guaranteed stability

Validation

Case Study (Numerical Simulation)

Training Results

Reward curve and Exploration

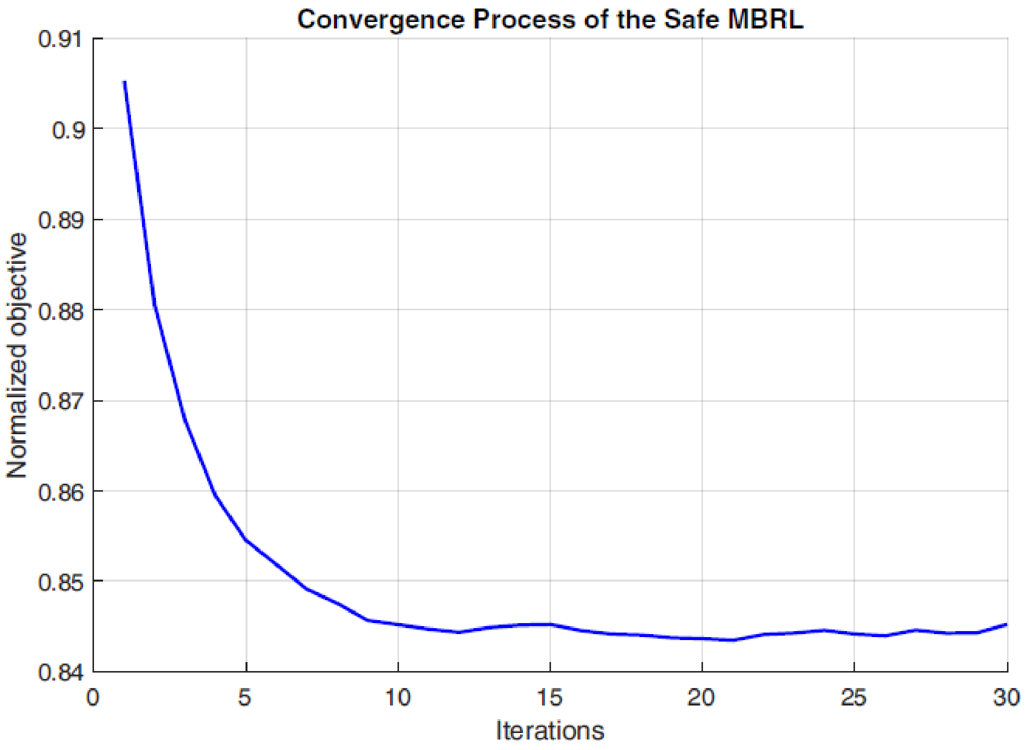


Figure 1. Converged Reward Curve

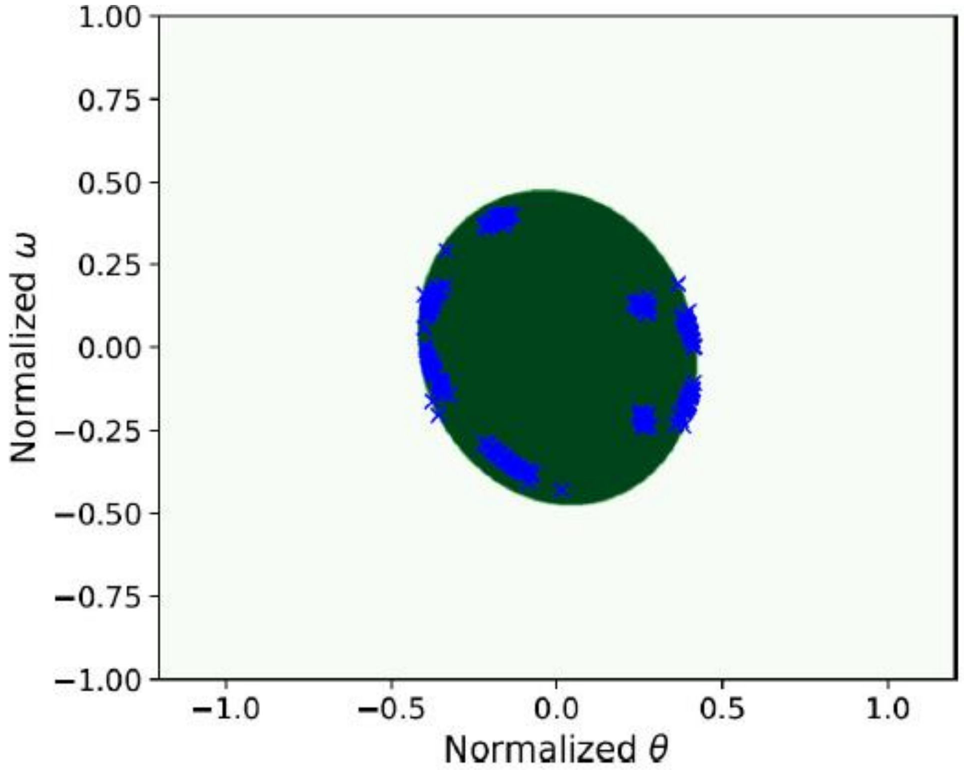


Figure 2. Relationship between ROA and Exploration Process

Agent performance

Frequency Deviation under Various Control Policy

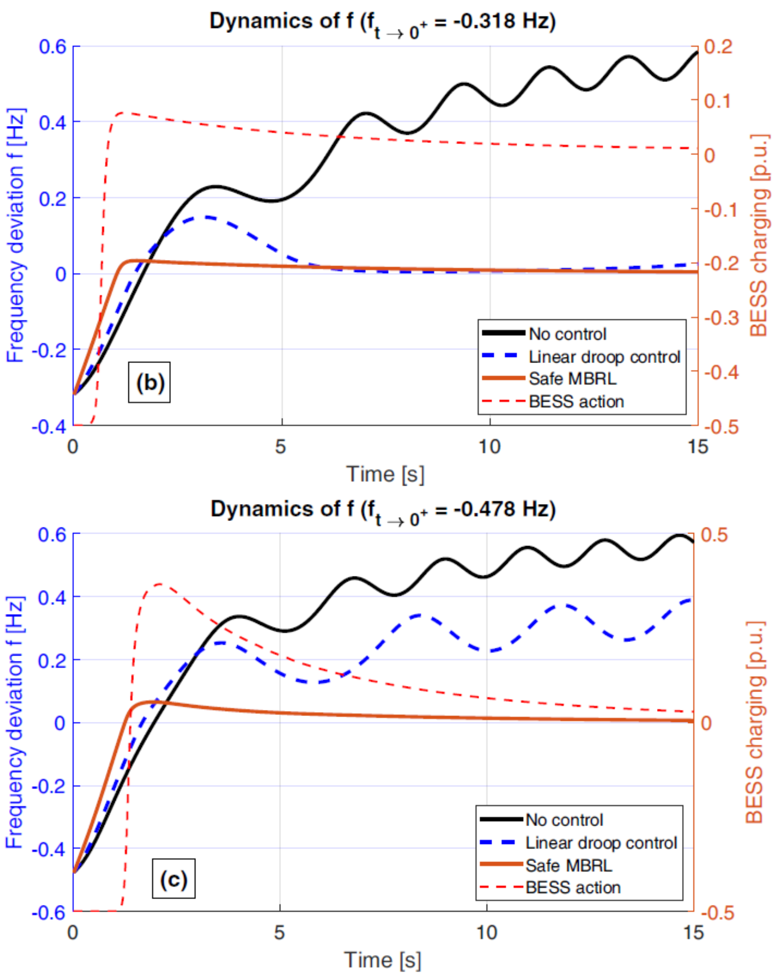


Figure 1: Safe RL vs Linear Control

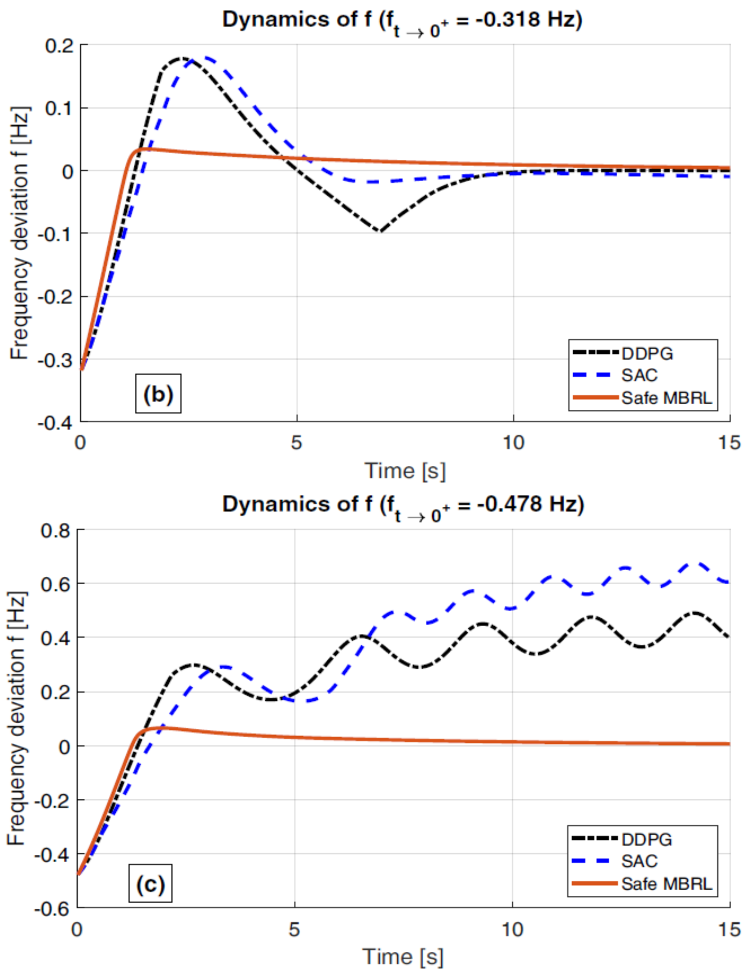
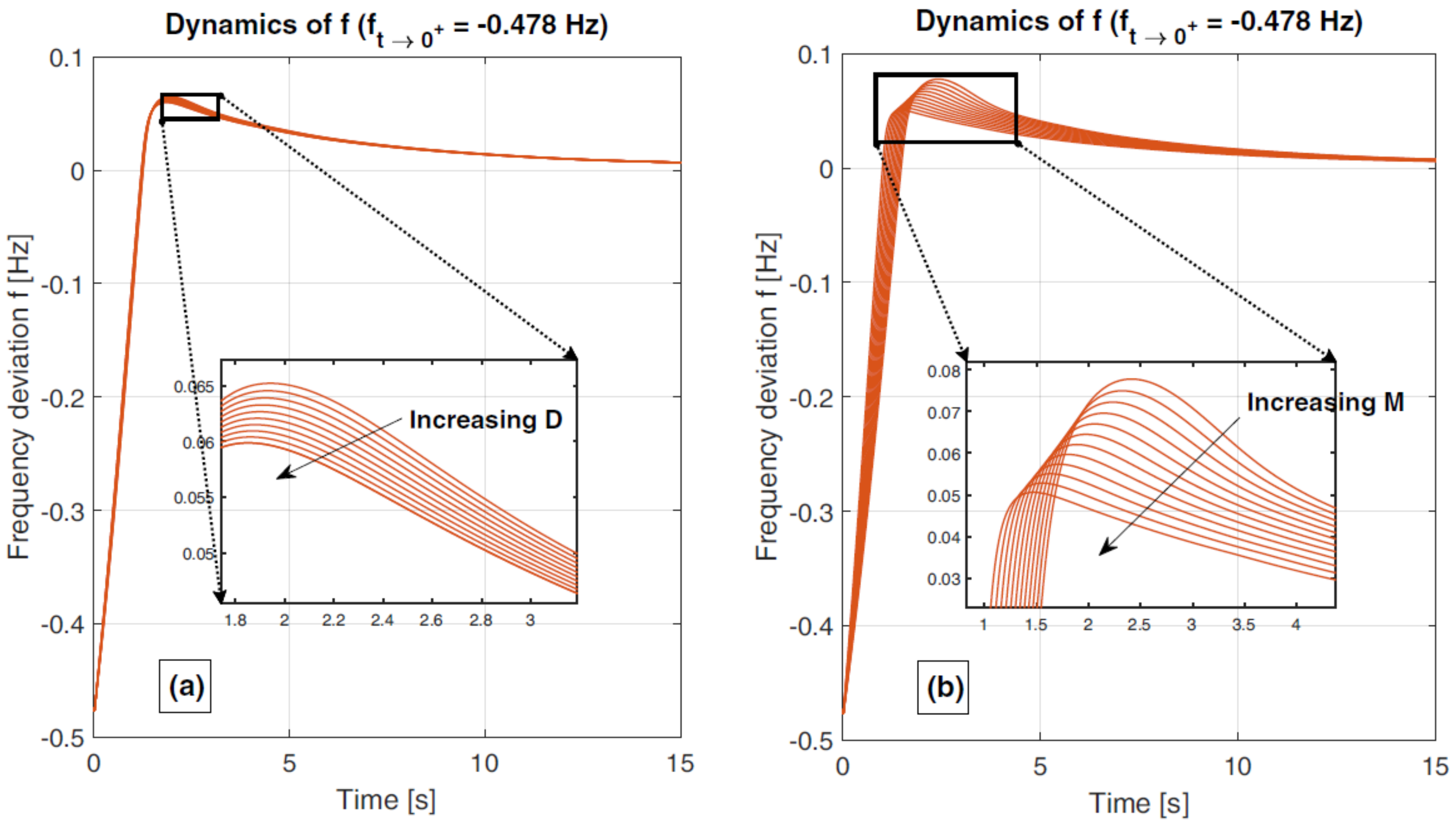


Figure 2: Safe RL vs Conventional RL

- Linear control and conventional RL cannot guarantee frequency stability under large disturbance, but safe RL can
- Safe RL have better frequency response performance

Agent performance

Safe RL under Varying System Parameters



- Safe RL is robust to virtual inertia and damping uncertainty.

Conclusion

Take Aways and Next Steps

Take aways

- Safe RL ensures dynamic stability under various disturbance. Conventional RL may loss stability under large disturbance.

- Safe RL is achieved by designing the Lyapunov function as the value function of the Bellman's equation.

- Safe RL is robust to system uncertainty.

Next Steps

- Scalability issue for dynamic stability guarantee (ROA quantification)
- Embed static security constraints
- Explore explainability for train RL-based policy

Acknowledgement

This work was supported by the CURENT research center and in part by the NSF grant ECCS-2033910.

Thanks Dr. Hang Shuai for leading this research and Jinning Wang's contribution to this work.

Thank you!

Q&A