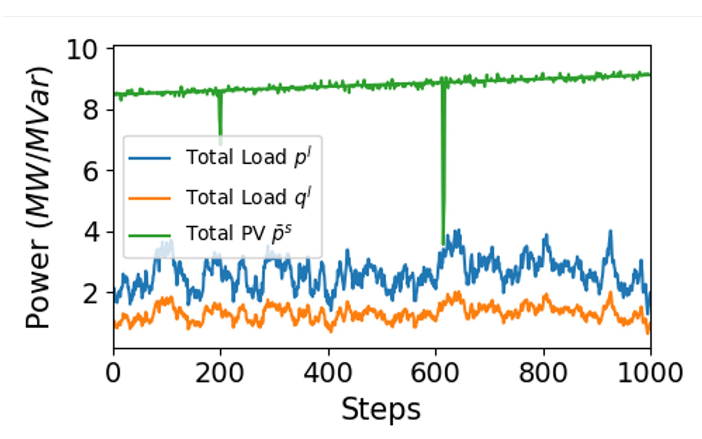
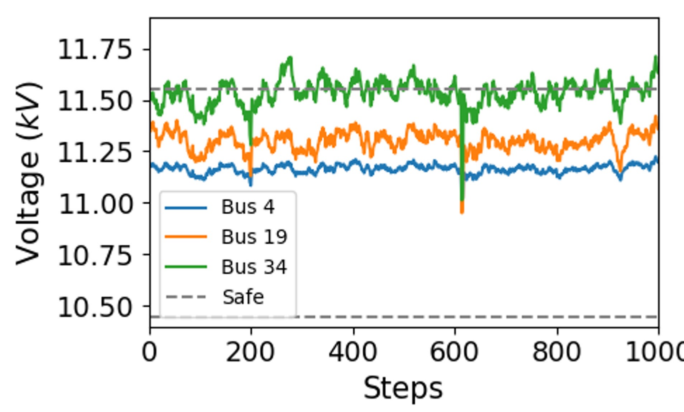
Online Voltage Regulation of Distribution Systems with Disturbance-Action Controllers

STUDENT: Peng Zhang, ADVISOR: Baosen Zhang

Overvoltage? Control Installed Inverters



(a) Total load and PV generation profiles



(b) Voltages with no control implemented

Challenges

Distribution Systems

- Inaccurate estimation on the line parameters and topology
- Fast-changing operation conditions on loads and solar

Inverter Controllers

- Communication delay
- Inaccessibility to load information

Formulation: Optimal Power Flow

$$\underset{u_t \in \mathbb{R}^{2n}}{\text{minimize}} \quad C_t(x_{t+1}, u_t)$$

$$\text{subject to} \quad x_{t+1} = B(u_t - u_t^l) \;\Rightarrow\; x_{t+1} = Bu_t + w_t, \quad u_t = f_t(x_t, \dots, x_1; u_{t-1}, \dots, u_0)$$

$$\underline{u}_t \le u_t \le \bar{u}_t$$

$$\underline{x}_{t+1} \le x_{t+1} \le \bar{x}_{t+1}$$

x_t: Voltage deviation, u_t: Control input, u_t^l: Uncontrollable loads,

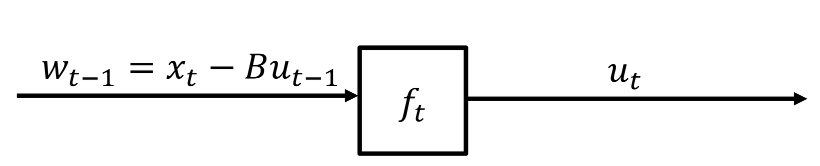
C_t: Loss function, B: Linearized model dynamics,

ω_t: Voltage drop from uncontrollable loads

Time-varying Optimization	⇒	Optimal Control
Value-based		Policy-based

Goal: Find reliable policy f_t to approach the time-varying optima

Design: Disturbance Action Controller



- Correlated disturbance (system-level approach)
- Online updating (physical grid solvers)

Main Results

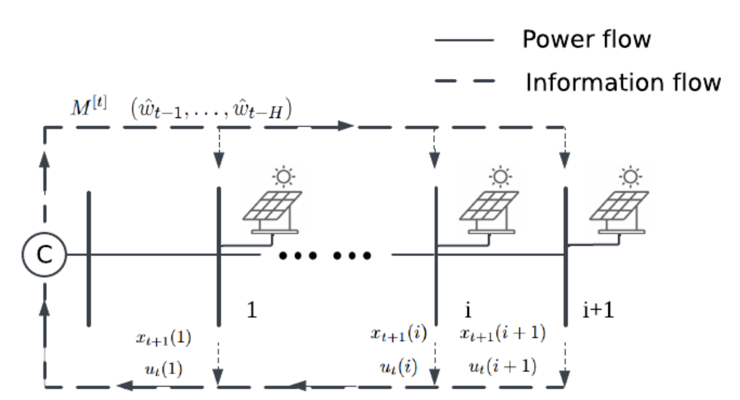


Figure: Algorithm architecture for online voltage control

Implementing

$$u_t = \left[\tilde{u}_t + \sum_{i=1}^{H} M_i^{[t]} \hat{w}_{t-i}\right]_{\underline{u}_t}^{\bar{u}_t}$$

Updating

$$M^{[t+1]} = M^{[t]} - \eta \nabla_M C_t(\hat{x}_{t+1}(M), u_t(M))$$

Stability condition

Theorem 7. Under *Assumptions 1-4*, it is sufficient to achieve stability on the state and input variables with model estimation error as $\|B - \hat{B}\| \le \epsilon_B$, by choosing initialization of controller as $\|M^{[0]}\| \le \frac{2\tilde{U}}{\epsilon_B \tilde{U} + W}$ and learning rate as

$$\eta \le \frac{2\tilde{U}}{LDd(1+\kappa_B)(\epsilon_B \tilde{U} + W)^2}. \tag{12}$$

Performance degradation to model inaccuracy

Theorem 8. Suppose that the disturbance-action controller is implemented with the stability condition on η and $\|M^{[0]}\|$, and the estimation error is bounded by $\epsilon_B \le \frac{W}{\tilde{U}}$. Then, it holds true that $\|\Delta u_t\| \le \bar{Y}_t$ and $\|\Delta x_{t+1}\| \le \bar{X}_{t+1}$, where $\bar{Y}_t$ and $\bar{X}_{t+1}$ are defined as follows,

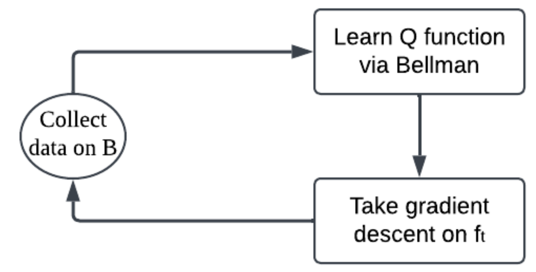
$$\bar{Y}_t \triangleq \begin{cases} (\bar{M}(\kappa_B + \epsilon_B))^{t-1} \|\Delta u_1\| & \text{if } \bar{M}(\kappa_B + \epsilon_B) \le 1, \\ \min\{(\bar{M}(\kappa_B + \epsilon_B))^{t-1} \|\Delta u_1\|, \tilde{U}\} & \text{if } \bar{M}(\kappa_B + \epsilon_B) > 1, \end{cases} \tag{13}$$

$$\bar{X}_{t+1} \triangleq \kappa_B \bar{Y}_t, \tag{14}$$

with the base case calculated as $\|\Delta \hat{w}_0\| \le \tilde{U}\epsilon_B$ and $\|\Delta u_1\| \le \bar{M}\tilde{U}\epsilon_B$.

Neural-Disturbance Action Controller

Neural Network-based Control Law
Model-based Policy Gradient



- Random initialization with NO pre-training
- Correlated perturbation as non-Gaussian variables

Twin Delayed Deep Deterministic policy gradient (TD3)

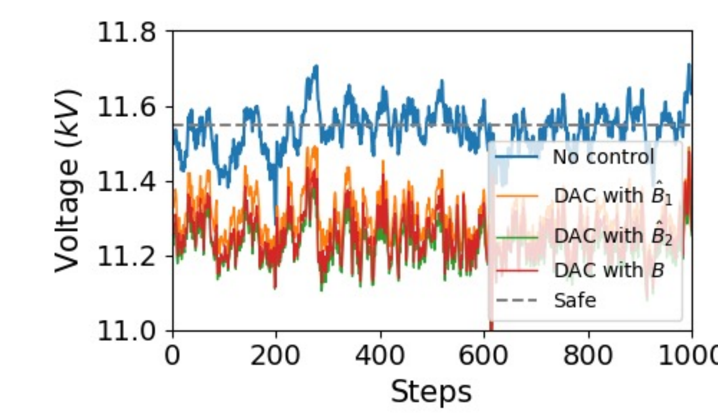
- Critic: $\phi_i \leftarrow argmin_{\phi_i} N^{-1} \sum (y - Q_{\phi_i}(x,u))^2$
- Actor: $\nabla_\theta J(\theta) = N^{-1} \sum \nabla_u Q_{\phi_1}(x,u)|_{u=f_\theta(w)} \nabla_\theta f_\theta(w)$

Difficulties

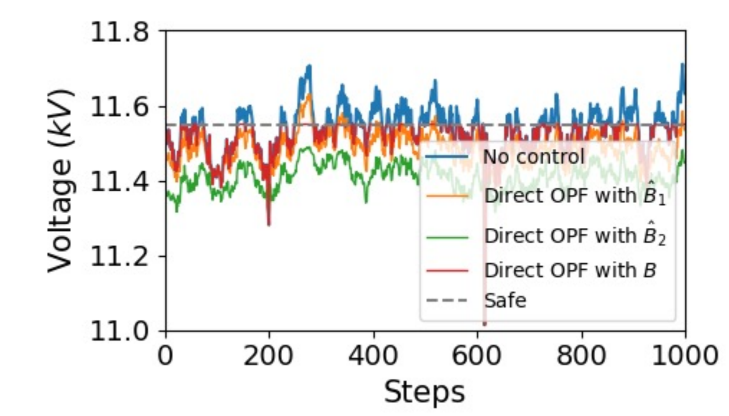
- Over-estimation bias: Double Q-learning and clipped
- High variance: target networks and delayed updates

Simulation Results

Robustness against model inaccuracy

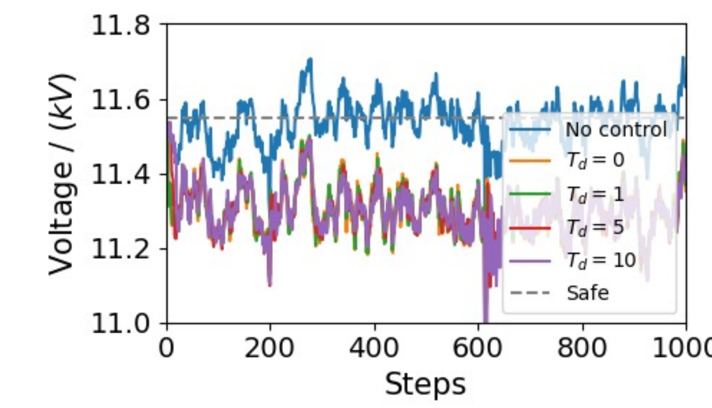


(a) Disturbance-action controller

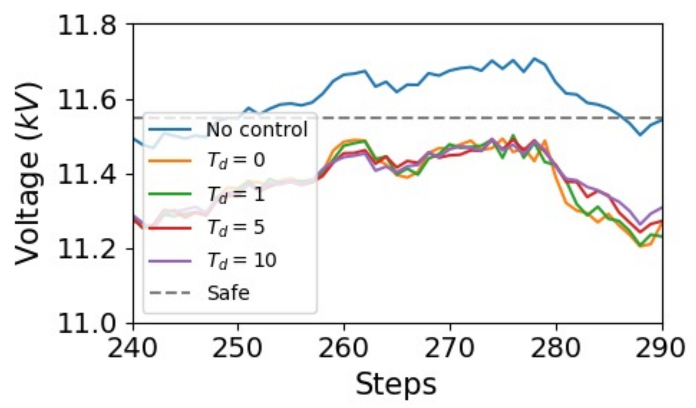


(b) Direct optimization method

Robustness against latency

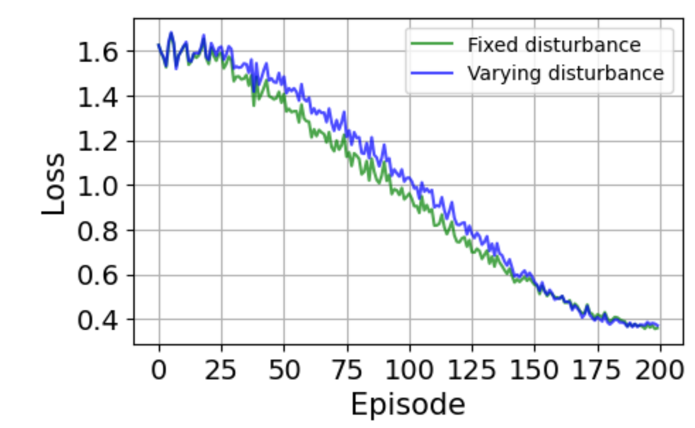


(a) Overall voltage profile at Node 34

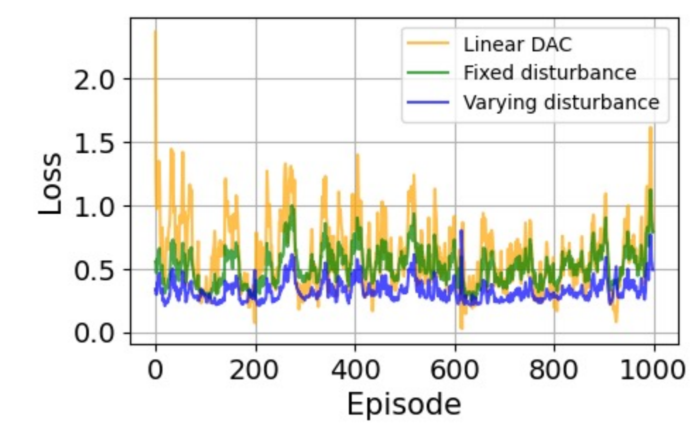


(b) Local voltage profile at Node 34

Improvement on optima tracking



(a) Loss during offline training



(b) Loss during online adapting

Conclusion & Future Work

Conclusion

- Policy-based control law with robustness to model inaccuracy and latency
- Stability condition on controller design
- Synthesis of offline training and online adapting through Neuro-Network policy

Future Work

Incorporate linearization error

$$x_{t+1} = f(x_t, u_t) \approx Bu_t + w_t$$

$$\Rightarrow \hat{w}_t' = x_{t+1} - \hat{B}u_t = \hat{w}_t + \delta_t$$

Decentralize controller design

$$\hat{w}_t = x_{t+1} - \hat{B}u_t$$

$$\hat{w}_{t,i} = x_{t+1,i} - \sum_{j=1}^{N} \hat{B}_{ij} u_{t,j} \approx x_{t+1,i} - \sum_{j=1}^{N} \hat{B}_{ij} u_{t,i}$$

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

References
[1] P. Zhang and B. Zhang, Online voltage regulation of distribution systems with disturbance-action controllers, arXiv preprint arXiv:2412.00629, 2024.
[2] Scott Fujimoto, Herke Hoof, and David Meger. Addressing function approximation error in actor-critic methods. In International conference on machine learning, pages 1587–1596. PMLR, 2018.