Microgrid PQ Control with Guaranteed Trajectory: Model-Based Analysis, Physics-informed Learning and Hardware Experiment

Buxin She[1], Fangxing Li[1], Hantao Cui[2], Hang Shuai[1], Oroghene Oboreh-Snapps[2], Rui Bo[2], Nattapat Praisuwanna[1], Jingxin Wang[1], Leon M. Tolbert[1]

[1] The University of Tennessee, Knoxville [2] Oklahoma State University [3]Missouri University of Sci. & Tech

❑ Introduction

To enhance the controllability and flexibility of the IBRs, this poster presents an adaptive PQ control method with a **guaranteed response trajectory**, combining model-based analysis, physics-informed reinforcement learning, and power hardware-in-the-loop (HIL) experiment. With the model-based derivation, the learning space of the RL agent is narrowed down from a function space to a real space, which reduces the training complexity significantly

Original control (Fixed gain)

- Devices in microgrid systems are controlled separately based on **preconfigured controller gains.**

VS

MFAC (Adaptive gain-tuning)

- In the event of a disturbance, keep the actual response following the desired trajectory by **adaptively adjusting the control gains.**

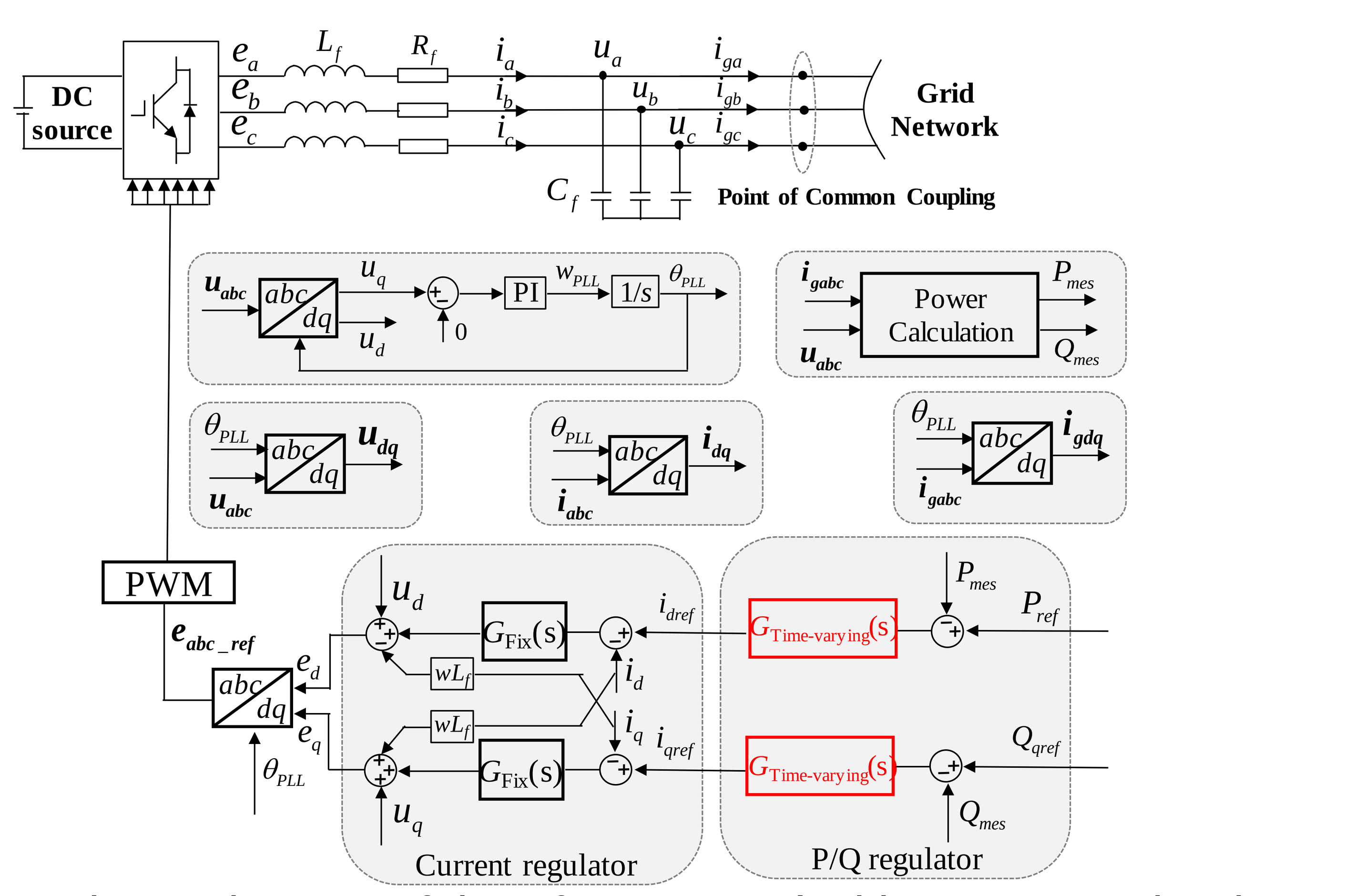

Fig. 1 Diagram of the of PQ control with guaranteed trajectory

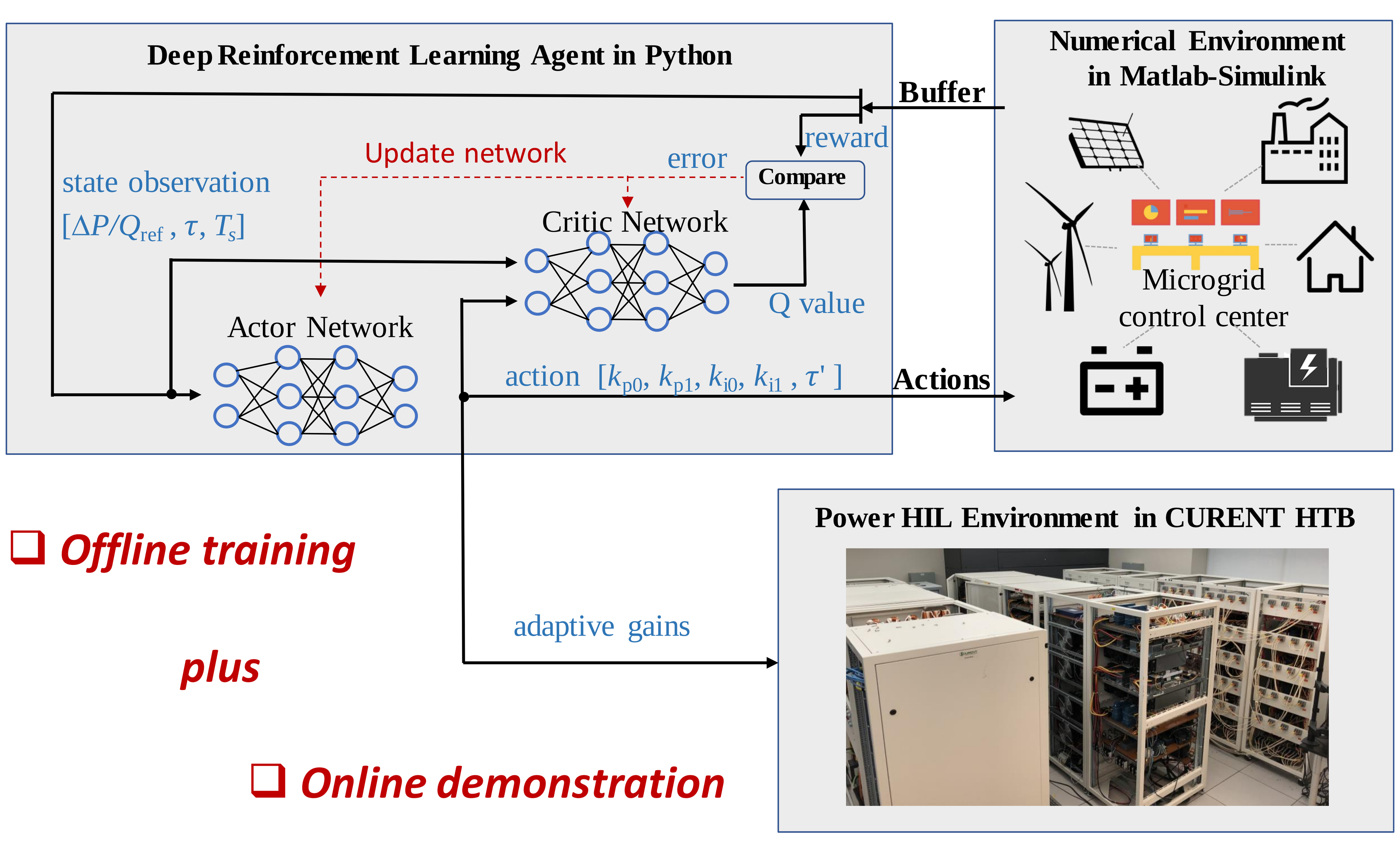

Fig. 2 Diagram of the proposed control framework

❑ Model-based analysis

- The time-varying gains that can guarantee an exponential P/Q trajectory consist of **a constant factor** and **an exponentially decaying factor**.
- The four constant coefficients k_{p0}, k_{p1}, k_{i0} and k_{i0} as well as the decaying time constant are determined by system parameters.

Time-varying Gains:

$$\begin{cases} k_p(t) = k_{p0} + k_{p1}e^{-t/\tau'} \\ k_i(t) = k_{i0} + k_{i1}e^{-t/\tau'} \end{cases}$$

where

Constant Coefficients:

$$\begin{cases} k_{p0} = \dfrac{L_f(1-1.5T_s/\tau)}{\tau K_{PWM}(k_{i2}/k_{p2}-1/\tau)} \\ k_{p1} = \dfrac{L_f}{\tau K_{PWM}}\left(1.5T_s + \dfrac{1.5T_s/\tau - 1}{k_{i2}/k_{p2} - 1/\tau}\right) \\ k_{i0} = 0 \\ k_{i1} = k_{p1}/\tau \\ \tau' = k_{i2}/k_{p2} \end{cases}$$

❑ Physics-informed Reinforcement learning

Algorithm 1: Physical informed TD3 training

Select T, N, $\boldsymbol{b}$, $\boldsymbol{\sigma}$, $\boldsymbol{\eta}$, α
Initialize θ_a and θc; Initialize physics function f based on (15)
Initialize replay buffer $\boldsymbol{B}$
for $t \leftarrow$ to T **do**
  $\boldsymbol{S} \leftarrow \boldsymbol{S}'$ [Update state]
  $\boldsymbol{a} = \pi_\theta(\boldsymbol{S})+\mathcal{E}$, where $\mathcal{E}\sim N(0, \boldsymbol{\sigma})$ [Select action]
  $k_p, k_i \leftarrow f(\mathbf{a})$ [Physics Transformation]
  $\boldsymbol{B} \leftarrow$ Append ($\boldsymbol{S}$, $\boldsymbol{a}$, r, $\boldsymbol{S}'$) [Store transition tuple]
  $\boldsymbol{B}_M \leftarrow \boldsymbol{B}'_M$ [Sample mini-batch tuples]
  $\boldsymbol{a}' = \pi_{\theta a}(\boldsymbol{S})+\mathcal{E}'$, where $\mathcal{E}'$=clip($\mathcal{E}$, -$\boldsymbol{b}$, $\boldsymbol{b}$)
  $y \leftarrow r + \alpha \min(Q_{\theta c1}(\boldsymbol{S}', \tilde{\mathbf{a}}), Q_{\theta c2}(\boldsymbol{S}', \tilde{\mathbf{a}}))$
  $\theta_c \leftarrow \mathrm{argmin}_{\theta c}\, \mathbf{E}\sum[y - Q_\theta(S,a)]^2$ [Update critics]
  if t mode d **then**
    $\nabla J(\theta) = E\nabla_a Q_{\pi\theta}(s,a)|_{a=\pi_\theta(s)} \nabla_\theta \pi_\theta(s)\nabla f_a$ [deterministic policy gradient]
    $\theta_a \leftarrow \eta\theta_a + (1-\eta)\theta'_a$ [Soft update for target actor networks]
    $\theta_c \leftarrow \eta\theta_c + (1-\eta)\theta'_c$ [Soft update for target critic networks]
  end if
end for
Output well-trained parameterized policy $\pi(\boldsymbol{\theta}_a)$

❑ Power HIL Experiment

A power HIL experiment is involved in to further demonstrate the proposed control method after the reward curve converges in numerical simulator. The HIL environment will be emulated **through CURENT HTB**, which uses identical commercial-grade power electronics inverters to emulate real microgrids.

❑ Case study

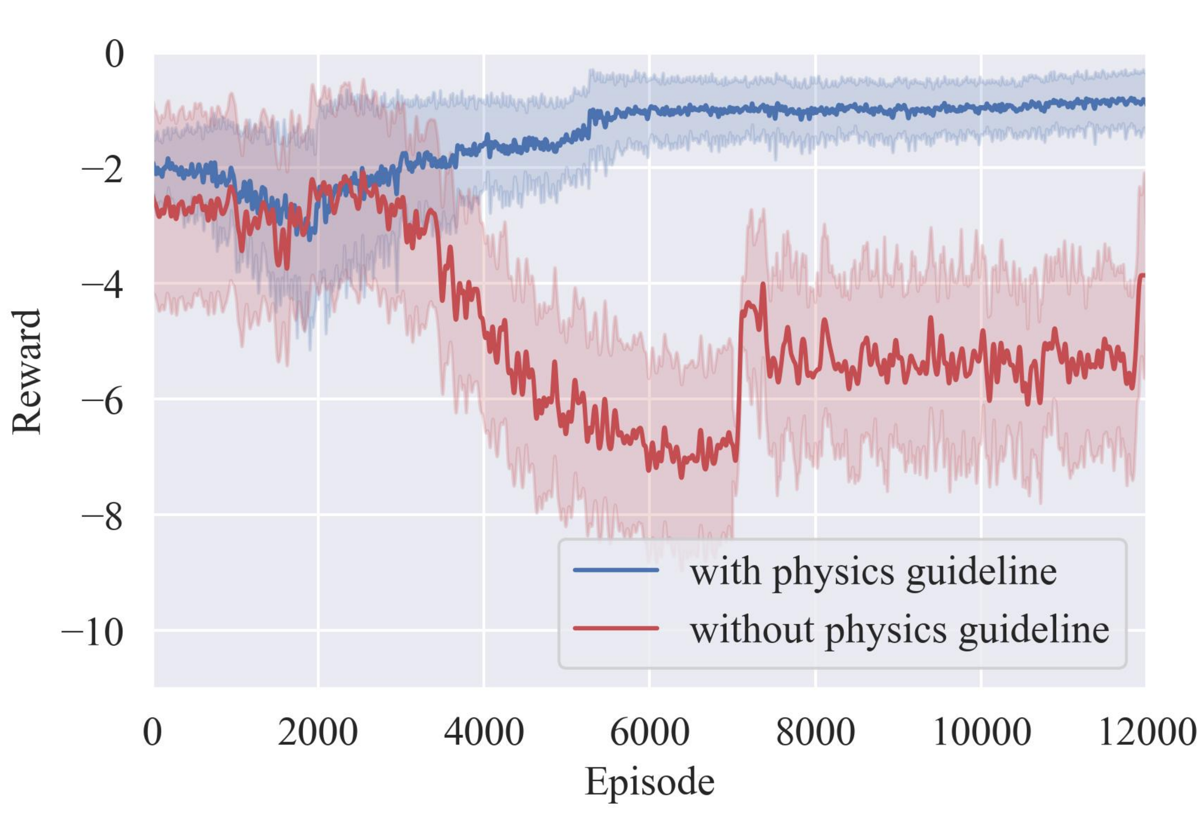

Fig. 4 Reward curve with and without the model-based derivation

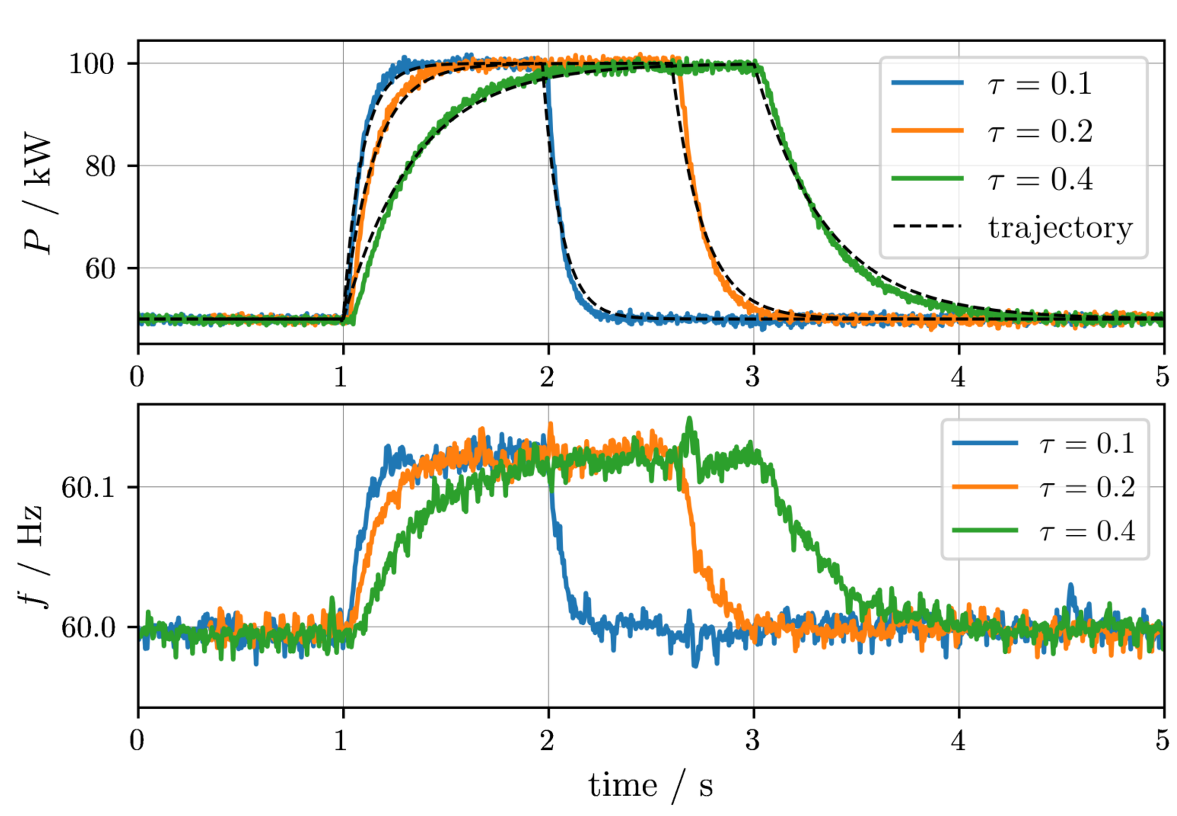

Fig. 5 Power HIL test results

❑ Conclusions

- The time-varying gains that can guarantee an exponential P/Q trajectory of inverters consist of **a constant factor** and **an exponentially decaying factor**.
- The physics-informed implementation reduces the learning space of RL agent from a **function space** to a **real space**, thus reducing the training complexity significantly.