

TD3-Based Current Control for Grid-Tied Inverters: Markov-State Design and Per-Unit Normalization

Licong Wu¹, Ruitie Li¹, Yu Chen², Binbin Li¹, Wei Cao¹, Chao Qi¹, Dianguo Xu¹

¹ Harbin Institute of Technology, Harbin, China

² Shandong University of Science and Technology, Qingdao, China

CONTROLLER DESIGN

Physics-derived state construction makes the TD3 observation physically interpretable, while physically based per-unit normalization removes severe numerical-scale imbalance and enables stable learning in the studied case.

1 CONTROL PROBLEM

- Grid-tied inverter current control determines dynamic power shaping, harmonic performance and disturbance rejection.
- Conventional PI tuning is tied to nominal filter and grid parameters.
- DRL provides continuous voltage control, but state sufficiency and heterogeneous feature scaling must be explicitly designed.

2 TD3 CURRENT-CONTROL STRUCTURE

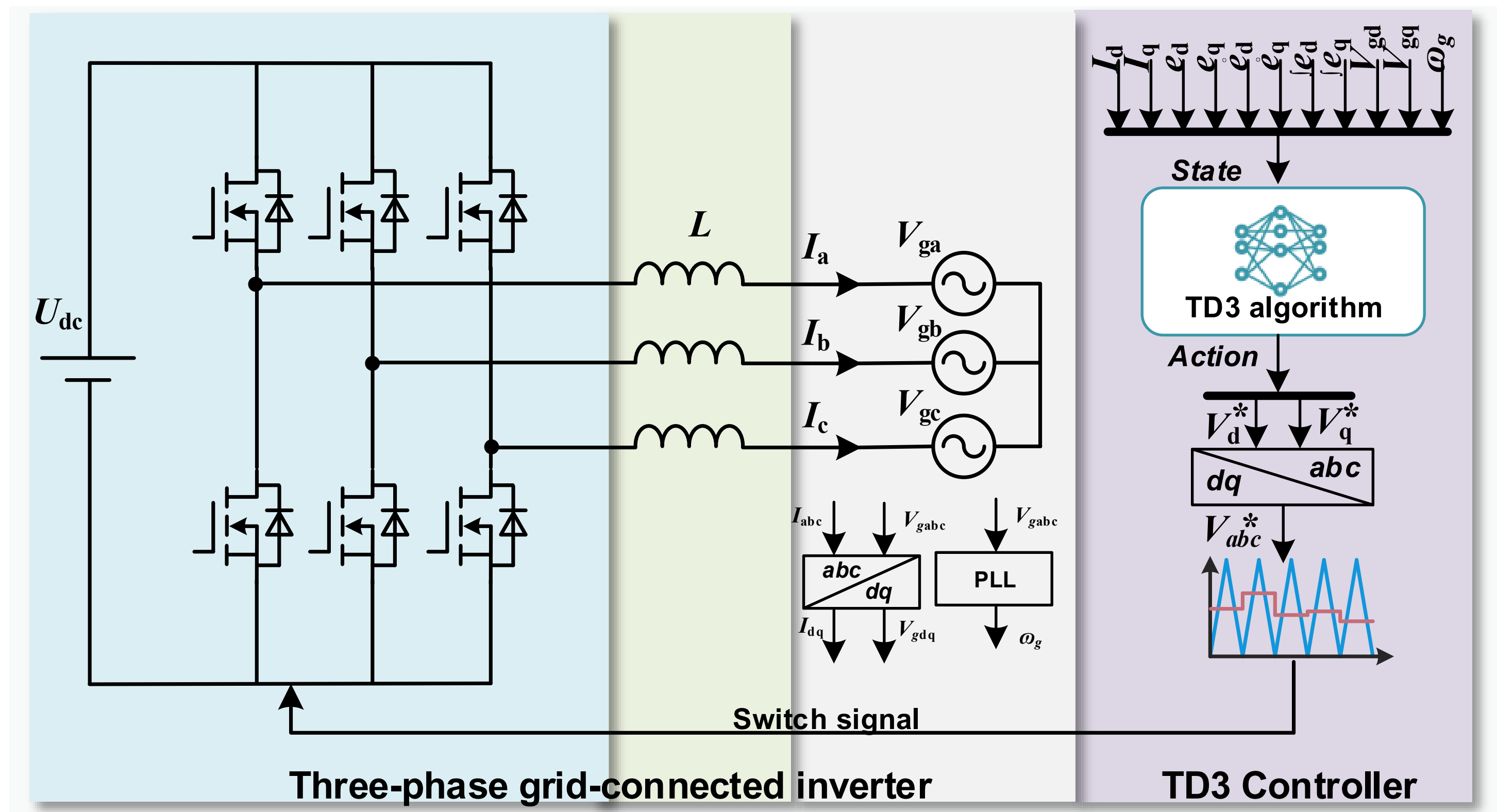
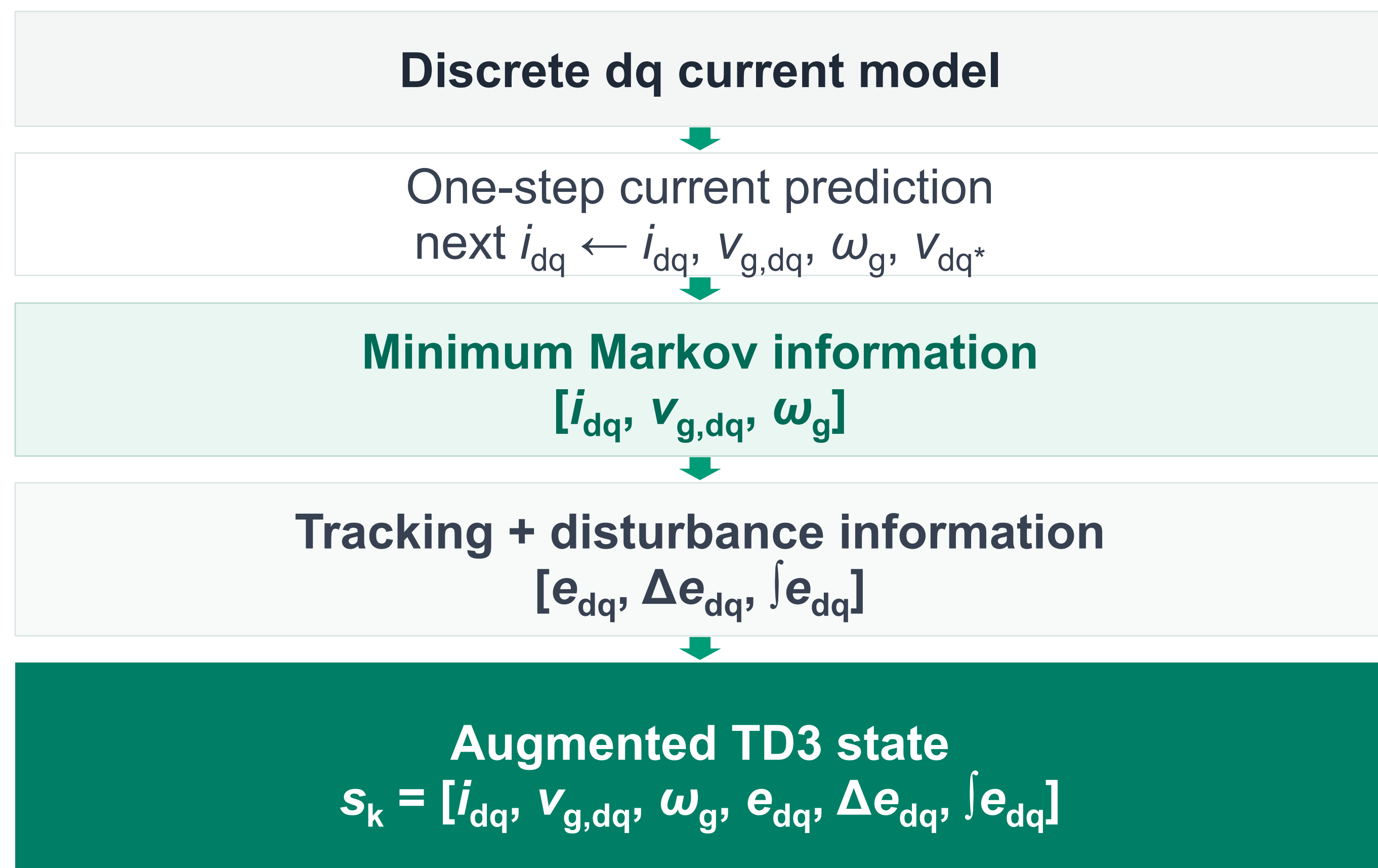


Fig. 1 TD3-based dq current-control structure

Twin critics · target-policy smoothing · delayed actor updates

3 PHYSICS-DERIVED MARKOV-STATE DESIGN



STATE IS DERIVED FROM THE PLANT DYNAMICS,
NOT ARBITRARILY SELECTED

4 WHY NORMALIZATION MATTERS

40× → >10⁶×

RAW FEATURE-INDUCED UPDATE DISPARITY

Analysis for the studied inverter/controller parameters

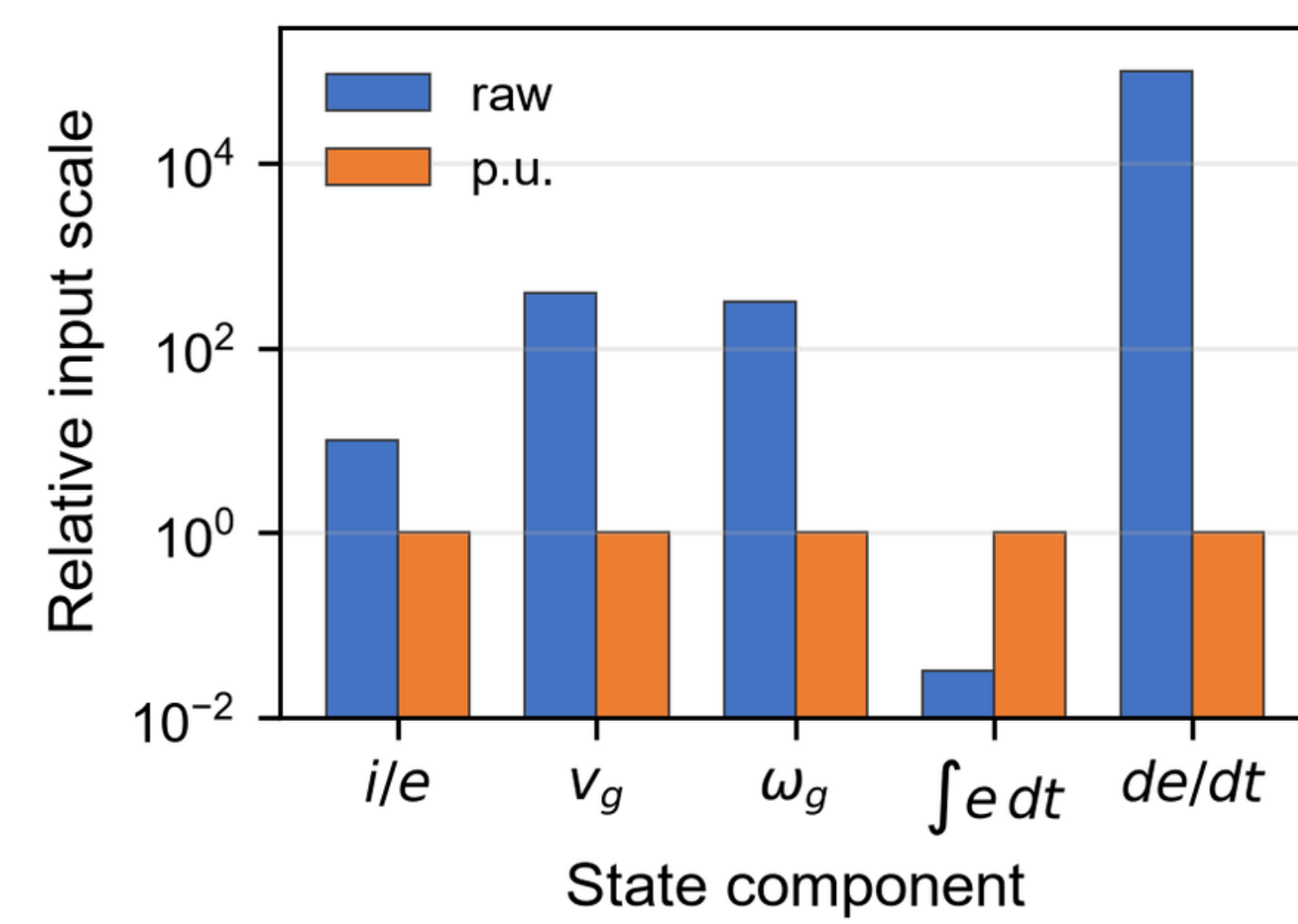


Fig. 2 Estimated input-scale imbalance

A / V / rad·s⁻¹ / integral / derivative → heterogeneous magnitudes
→ unequal neural-network updates → poor learning conditioning

PHYSICAL BASES

Current / error → $I_b = 10$ A
Grid voltage → $V_b = 400$ V
Grid frequency → $\omega_b = 2\pi \cdot 50$ rad/s
Error slope → $I_b/T_s = 1.0 \times 10^5$ A/s
Error integral → $I_b/\omega_b = 3.18 \times 10^{-2}$ A·s

PHYSICAL BASES
→ COMPARABLE LEVELS

5 TRAINABILITY EVIDENCE

WITH P.U.
max reward > 3.29×10^4

WITHOUT P.U.
negative after >3000 episodes

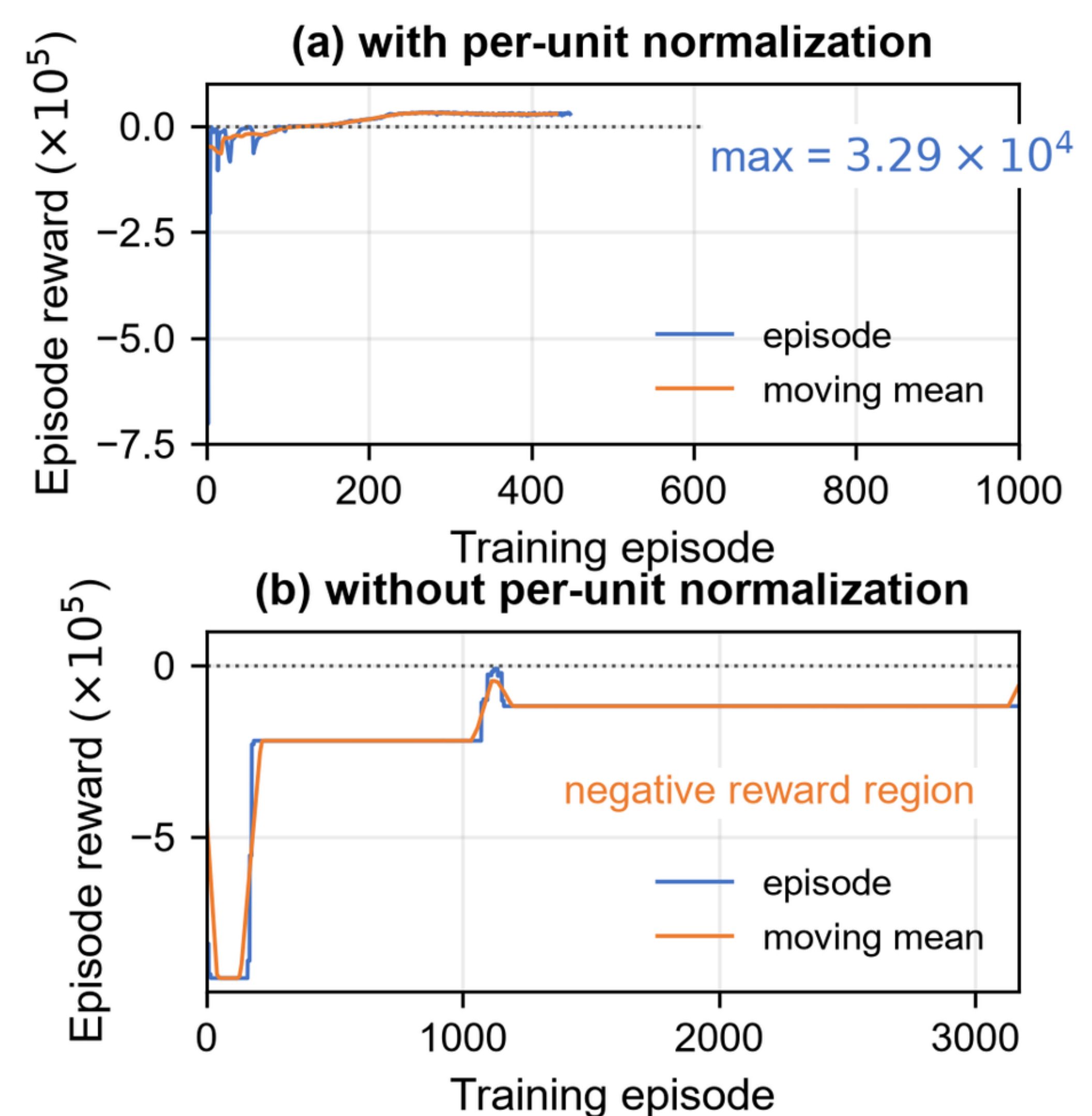


Fig. 3 Direct with/without evidence for per-unit normalization

- State comes from current-loop dynamics.
- Physical P.U. scaling enables stable learning in the studied case.

TD3-Based Current Control for Grid-Tied Inverters: Markov-State Design and Per-Unit Normalization

Licong Wu¹, Ruitie Li¹, Yu Chen², Binbin Li¹, Wei Cao¹, Chao Qi¹, Dianguo Xu¹

¹ Harbin Institute of Technology, Harbin, China

² Shandong University of Science and Technology, Qingdao, China

SIMULATION RESULTS

The TD3 actor achieves 1.59% current THD at 10 A, ≈ 1 ms settling for a 10 $\rightarrow$ 15 A step, and tracks 5–20 A current references without retraining in Simulink.

6 EVALUATION SETUP

800 V DC LINK	50 Hz GRID	100 μ s CONTROL T_s	20 kHz PWM	15.9 mH L_f
MATLAB/Simulink Evaluation				

7 SIMULATED CONTROL PERFORMANCE

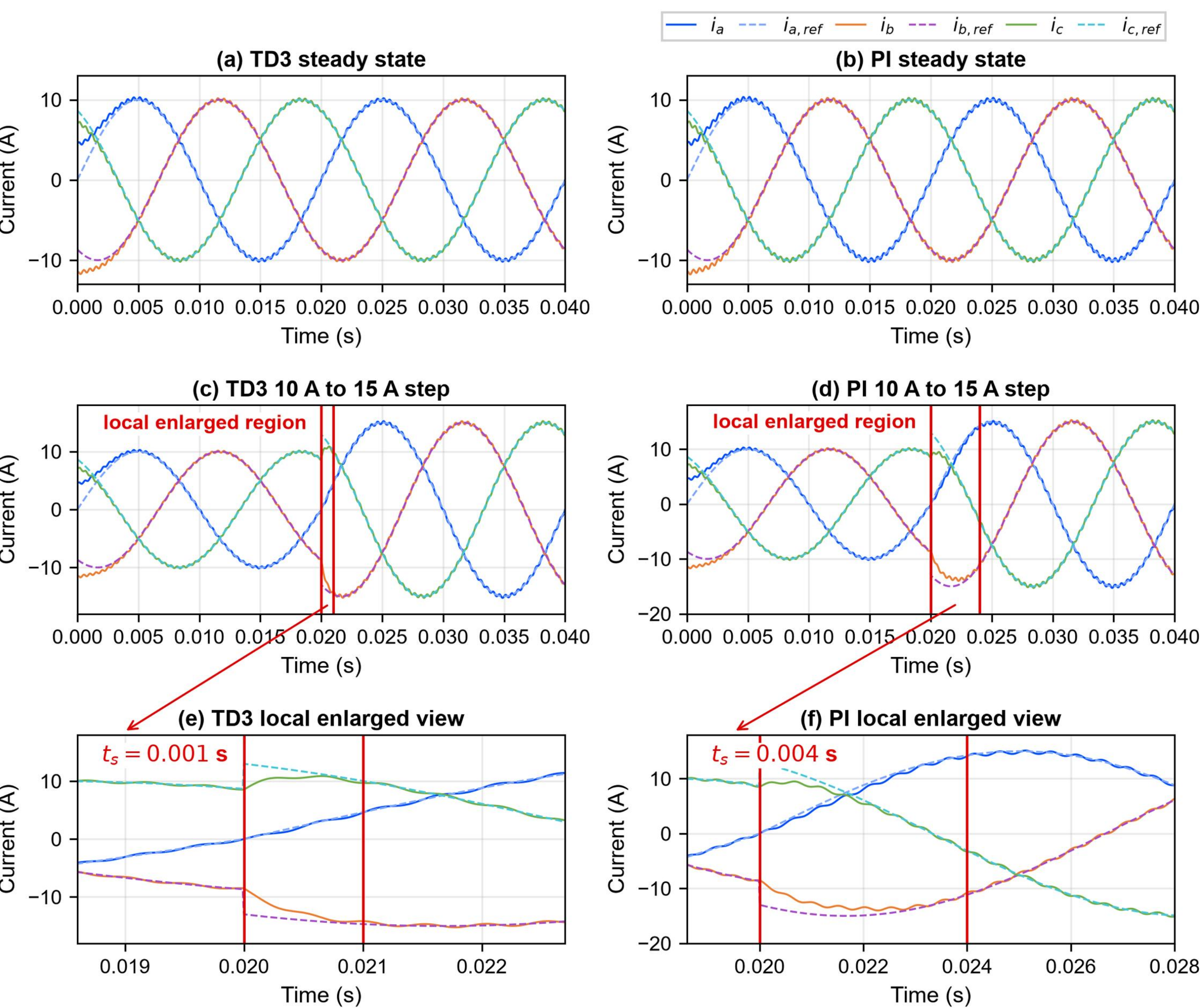


Fig. 4 Complete simulated TD3-PI comparison

8 PERFORMANCE SUMMARY

1.59% TD3 CURRENT THD @ 10 A	1.62% PI CURRENT THD @ 10 A	≈ 1 ms vs ≈ 4 ms TD3 vs PI · 10 A $\rightarrow$ 15 A
------------------------------------	-----------------------------------	---

$\approx 4\times$ shorter simulated settling time in the tested 10 A $\rightarrow$ 15 A step

Faster transient · comparable steady-state quality

9 REFERENCE ADAPTABILITY

FROZEN ACTOR · NO RETRAINING

Same trained actor; only the d-axis current reference changes.

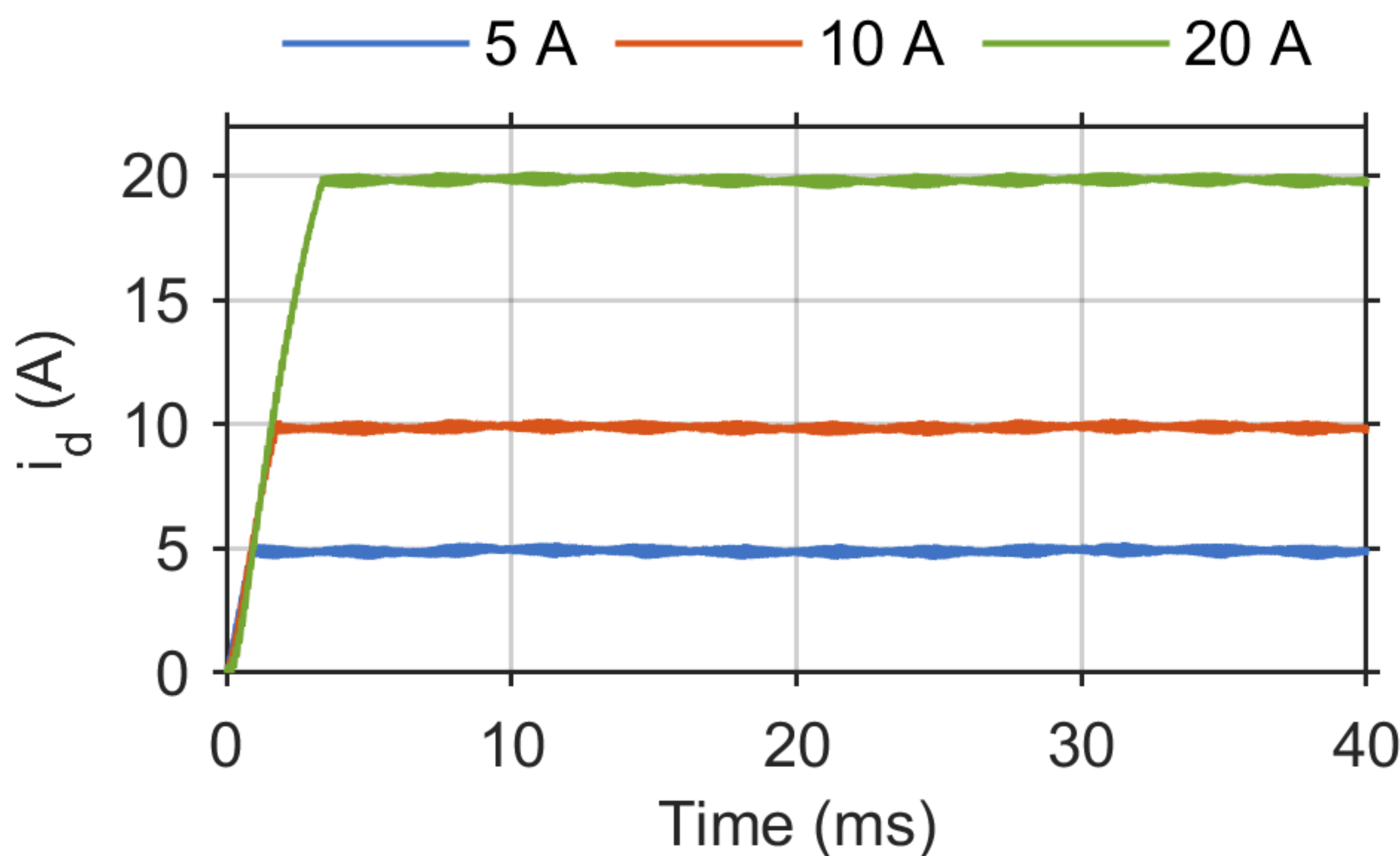


Fig. 5 5–20 A reference tracking without retraining in Simulink

5 A

mean 4.898 A
error 0.10 A
THD 3.22%

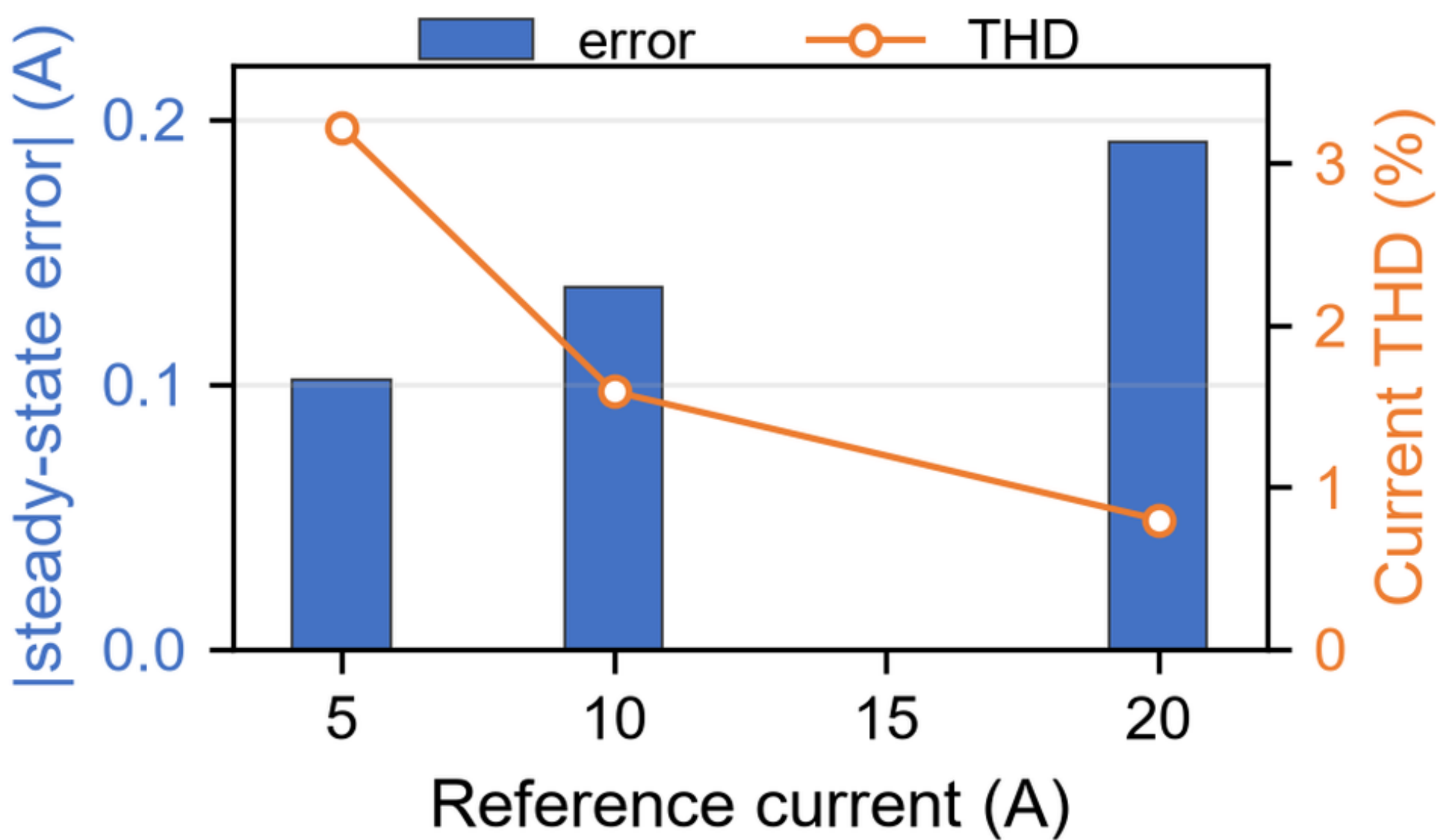
10 A

mean 9.863 A
error 0.14 A
THD 1.59%

20 A

mean 19.808 A
error 0.19 A
THD 0.80%

10 QUANTITATIVE REFERENCE SWEEP



MAIN TAKEAWAYS

- Similar low THD at 10 A.
- $\approx 4\times$ shorter simulated settling in the tested step.
- Frozen actor tracks 5–20 A without retraining.

SCOPE & NEXT STEP

Current: MATLAB/Simulink validation

Future (not yet demonstrated): HIL · weak-grid · parameter-drift robustness