Convergence of AI, Physics, Computing, and Control for Intelligent Power System Control

Qiuhua Huang
Associate Professor
Colorado School of Mines
(qiuhuahuang@mines.edu)

PSERC Webinar
February 15, 2023

Acknowledgement

Collaborators

- PNNL (including former colleagues): Yousu Chen, Renke Huang, Tianzhixi (Tim) Yin, Sohom Datta, Xueqing Sun, Bruce Palmer, Yuan Liu, Z. Jason Hou, Long Vu, Sayak Mukherjee, Kaveri Mahapatra, Xinda Ke, Shaobu Wang, Xinya Li, Yan Du, Yujiao Chen, Ramij Raja Hossain
- V&R Energy: Mariana Vaiman, Michael Vaiman
- Google: Jie Tan, Wenhao Yu
- EPRI: Adrian Kelly
- RTE-France: Antoine Marot, Benjamin Donnot, Karim Chao
- Song Wang (now with PGE, previously with PacifiCorp), Rui Fan (DU), Wei Qiao (UNL), Jochen Cremer (DTU),

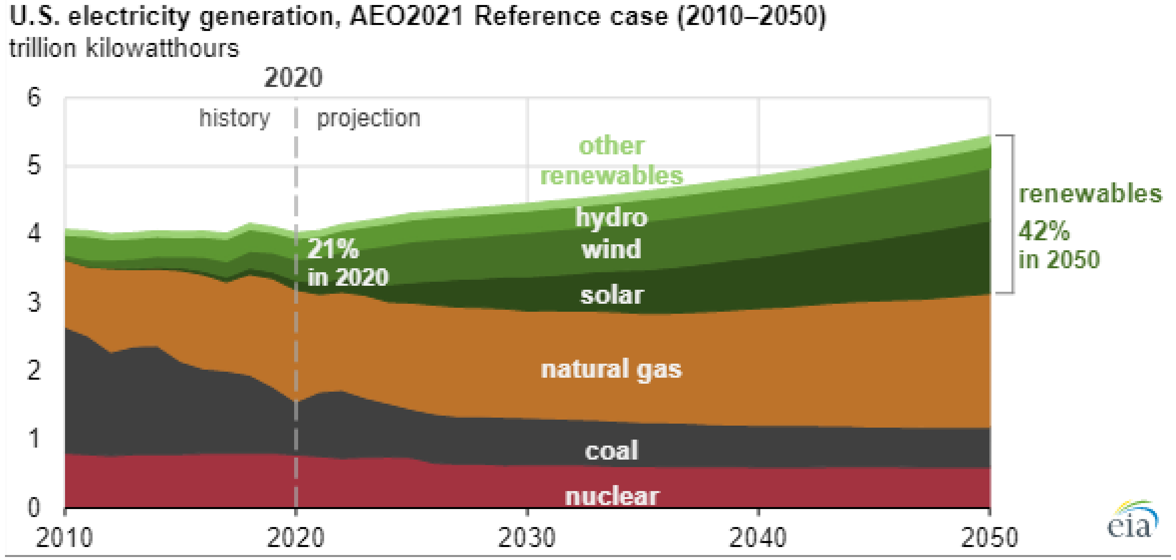
Funding support: ARPA-E OPEN 2018 Program, DOE OE AGM, PNNL DeepScience LDRD

Presentation Outline

- **Motivation and background**
- **Introduction to deep reinforcement learning**
- **Convergence of AI, physics, computing and control for intelligent power system control**
  - **Integrated framework**
  - **Open-source environments for DRL-based grid control**
  - **Advanced DRL algorithms for grid control**
  - **Test results, demo and applications**
- **Future Work and Potential Applications**

Increasing Renewables and Rapidly Changing Operation Conditions

EIA projects renewables share of U.S. electricity generation mix will double by 2050



California ISO net load "duck curve"

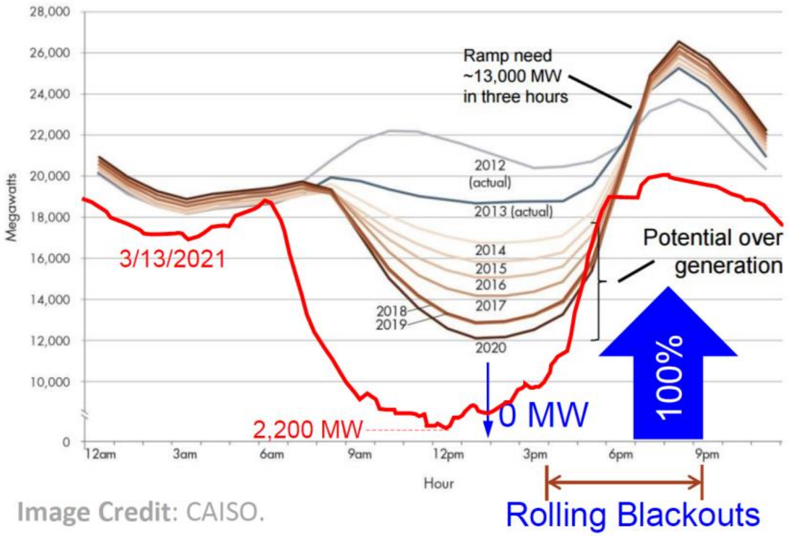
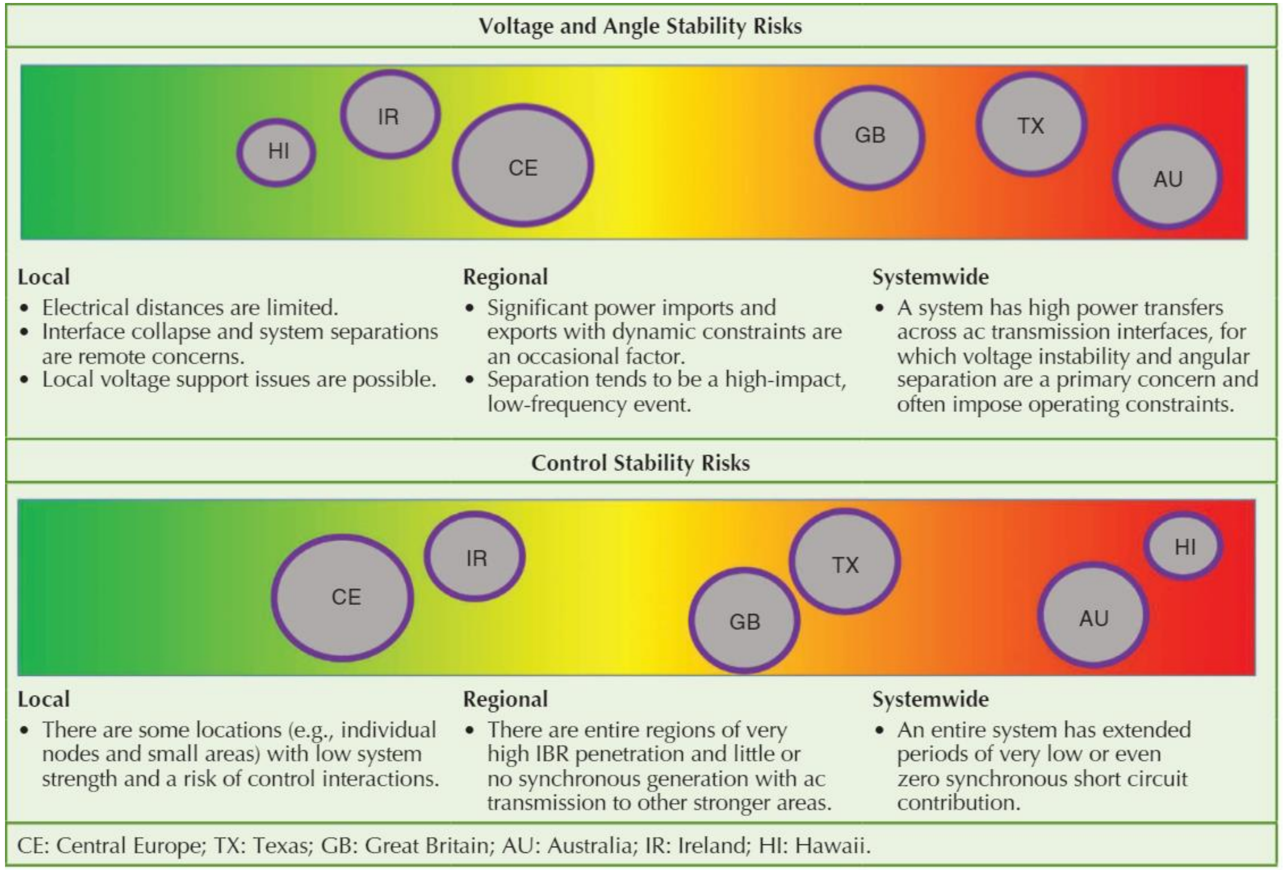


Image Credit: CAISO.

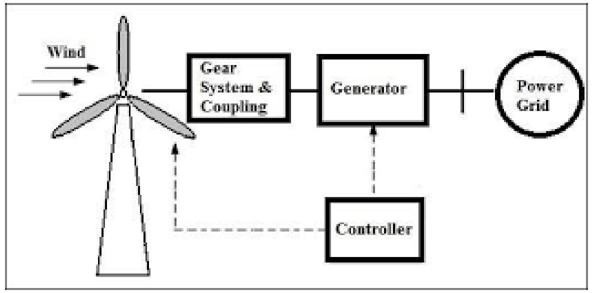
Sources: EIA

Power System Stability Risks Are Becoming A Global Issue

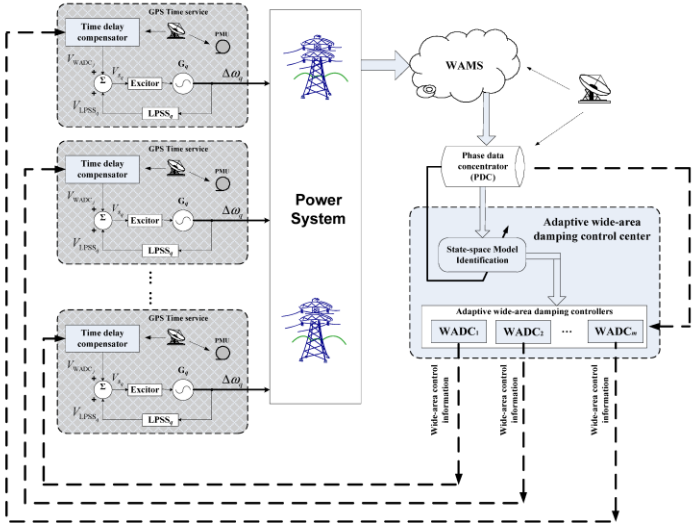


Source: J. Matevosyan, et al. “A future with inverter-based resources”, IEEE Power and Energy Magazine, Nov/dec, 2021

Power System Stability Control and Operation for Keeping the Lights On



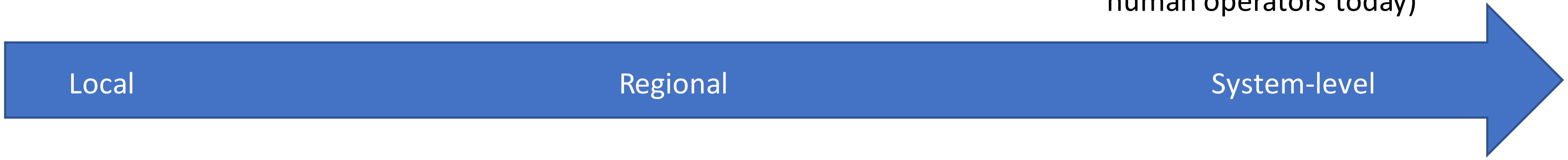
Local control



Wide-area control[1]

Imagine credit: ERCOT

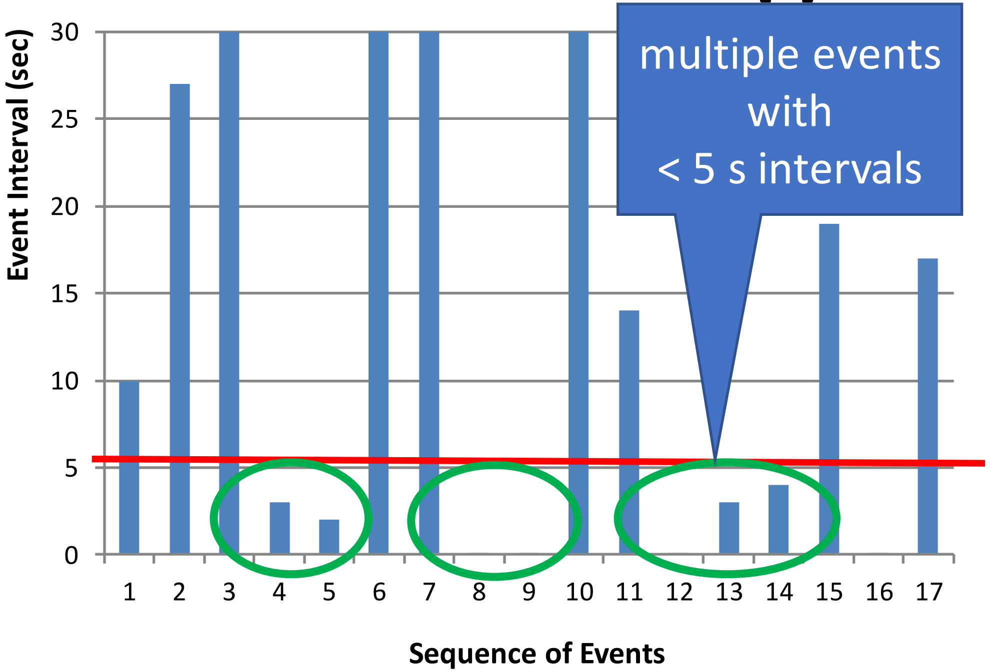
System control and operation (primarily by human operators today)



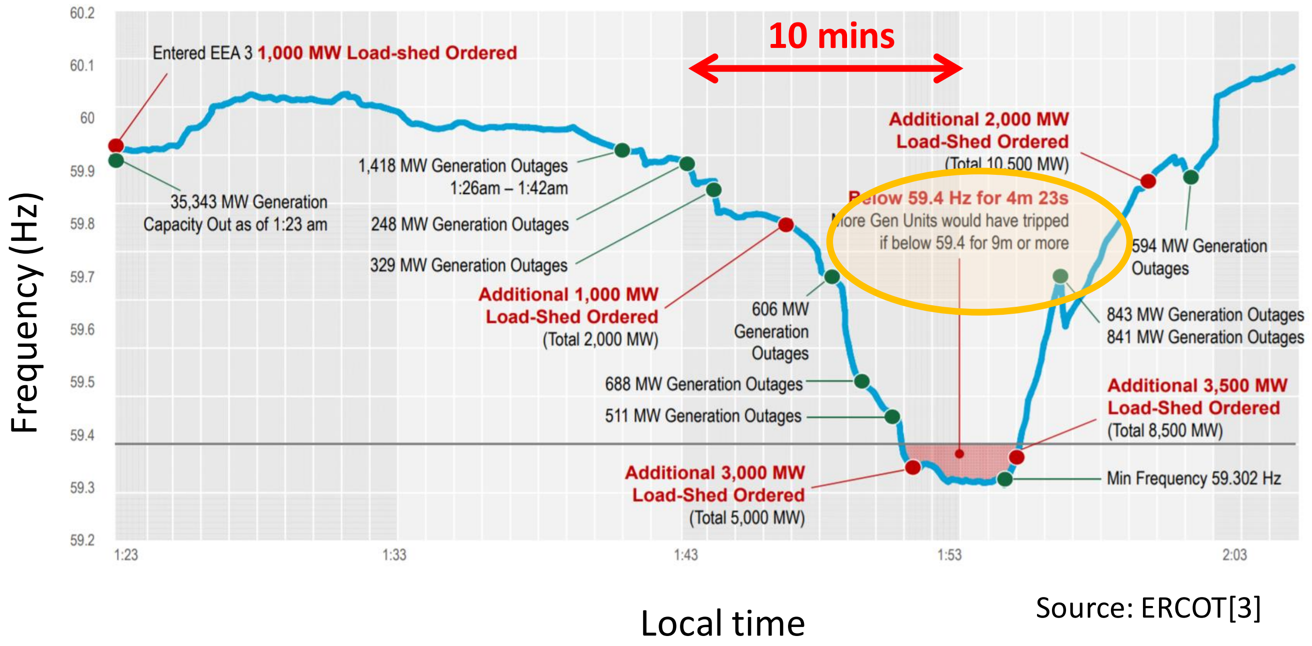
[1] Cai, G.; Yang, D.; Liu, C. Adaptive Wide-Area Damping Control Scheme for Smart Grids with Consideration of Signal Time Delay. *Energies* **2013**, *6*, 4841-4858.

Big Challenges in Grid Operation and Control

September 8, 2011 Pacific Southwest Blackout in U.S. [1]



Texas was "seconds and minutes" away from catastrophic months long blackouts[2]



Source: ERCOT[3]

[1] https://www.nerc.com/pa/rrm/ea/Pages/September-2011-Southwest-Blackout-Event.aspx#:~:text=On%20the%20afternoon%20of%20September,%2C%20and%20Baja%20California%2C%20Mexico.
[2] https://www.texastribune.org/2021/02/18/texas-power-outages-ercot/
[3] http://www.ercot.com/content/wcm/key_documents_lists/225373/Urgent_Board_of_Directors_Meeting_2-24-2021.pdf

Two Types of System-level Control for Dealing with Contingencies

- Preventive control (prior to contingency)
  - Minimize the chance of emergencies
  - Often used conservatively to ensure sufficient security margin against creditable contingency events (e.g., N-1) all the time
  - Typical methods: generation redispatch, var compensation, demand response/load curtailment
  - Emerging methods: transmission switching
- Emergency control (post-contingency/disturbance)
  - Minimize impacts of large disturbances; serve as a safety net for the system
  - Usually used against severe or low probability-high impact contingencies
  - Typical control methods: load shedding, generator tripping, controlled islanding
  - Emerging methods: emergent demand response, HVDC redispatch

Focus of the talk today

The Grand Challenge of Achieving Intelligent Emergency Control

- Power system post-event emergency control has **strong requirements:**
  - **Scalability:** >20,000 buses (with 1000s of control devices now → millions in the near future)
  - **Solution time:** < 5 seconds
  - **Security and adaptability** (to fast-changing conditions)

- Existing control methods and their issues:
  - Rule-based control (not adaptive, time-consuming to develop and update them)
  - Model-predictive control (scalability and solution time issues)
  - Learning-based (or data-driven) control (scalability, security and adaptability issues)

Can We Transfer AI Successes in Games and Robotics to Complex Grid Control?

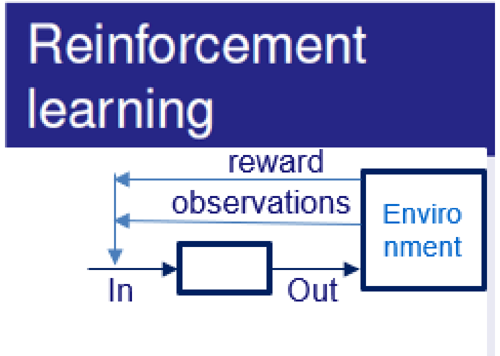
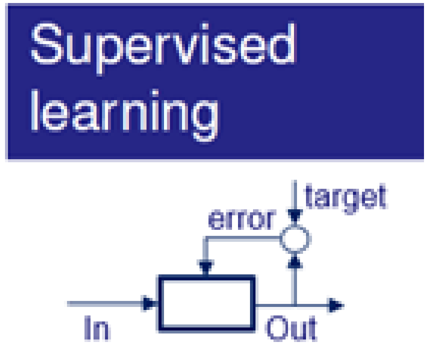
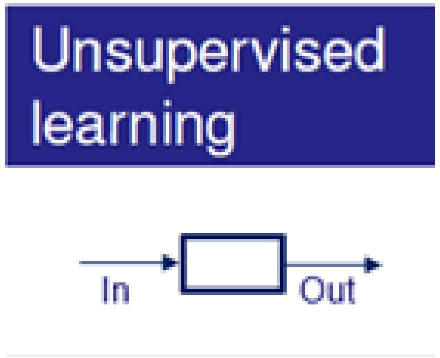
Credit: Nature

Credit: OpenAI

https://www.wecc.org/epubs/StateOfTheInterconnection/Pages/Western-Interconnection.aspx

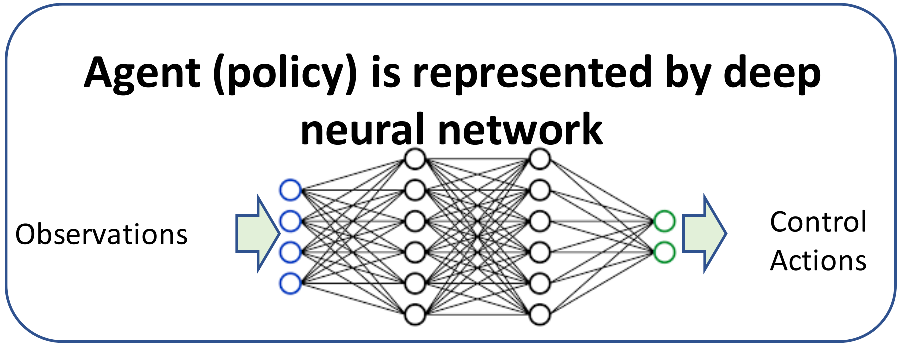
Introduction to Deep Reinforcement Learning

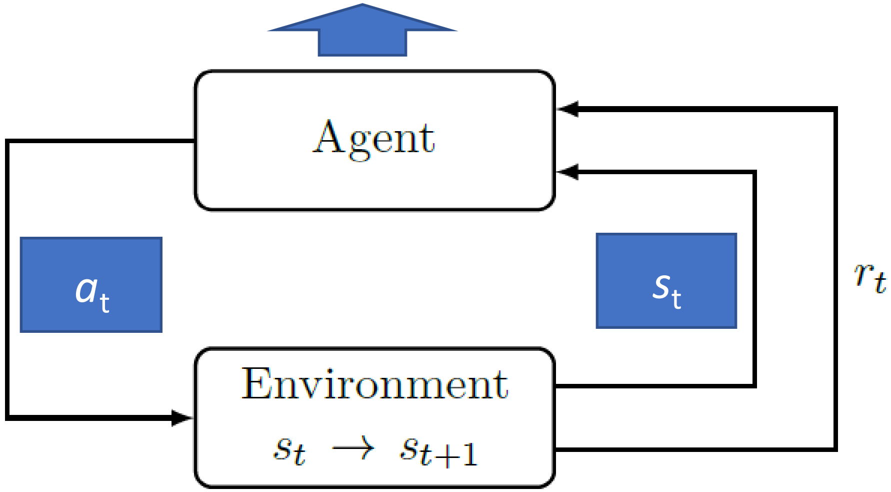
Reinforcement Learning vs Other ML Methods

Machine learning categories			
Type of data	Data for training include observations and rewards from the environment. They are usually not predefined. Data is generated on-the-fly by interacting with the environment.	Data for training include pairs of input and the desired output (Needs a large amount of data and labels)	Data for training do not include desired outputs
Example methods	•Q-Learning •Deep reinforcement learning	•Neural networks •Decision trees	•K-mean clustering •Autoencoders
Specific tasks	•Control •Scheduling •Sequential decision-making	•Classification •Regression (prediction)	•Discover clusters •Identify factors/structures

Reinforcement Learning in a Nutshell

- (Why) Reinforcement learning (RL) is designed for solving **sequential decision-making problems** in a stochastic environment.
- (What) The goal is to learn a control policy to **maximizes expected accumulative reward** over time.
- (How) The agent learns a control policy **iteratively** through interacting with the environment via **trial-and-errors guided by the reward signal.**
- **Deep** reinforcement learning = **deep learning** + reinforcement learning
  - Not just deep neural network, but more about deep learning techniques and ecosystem.





- Power Grid

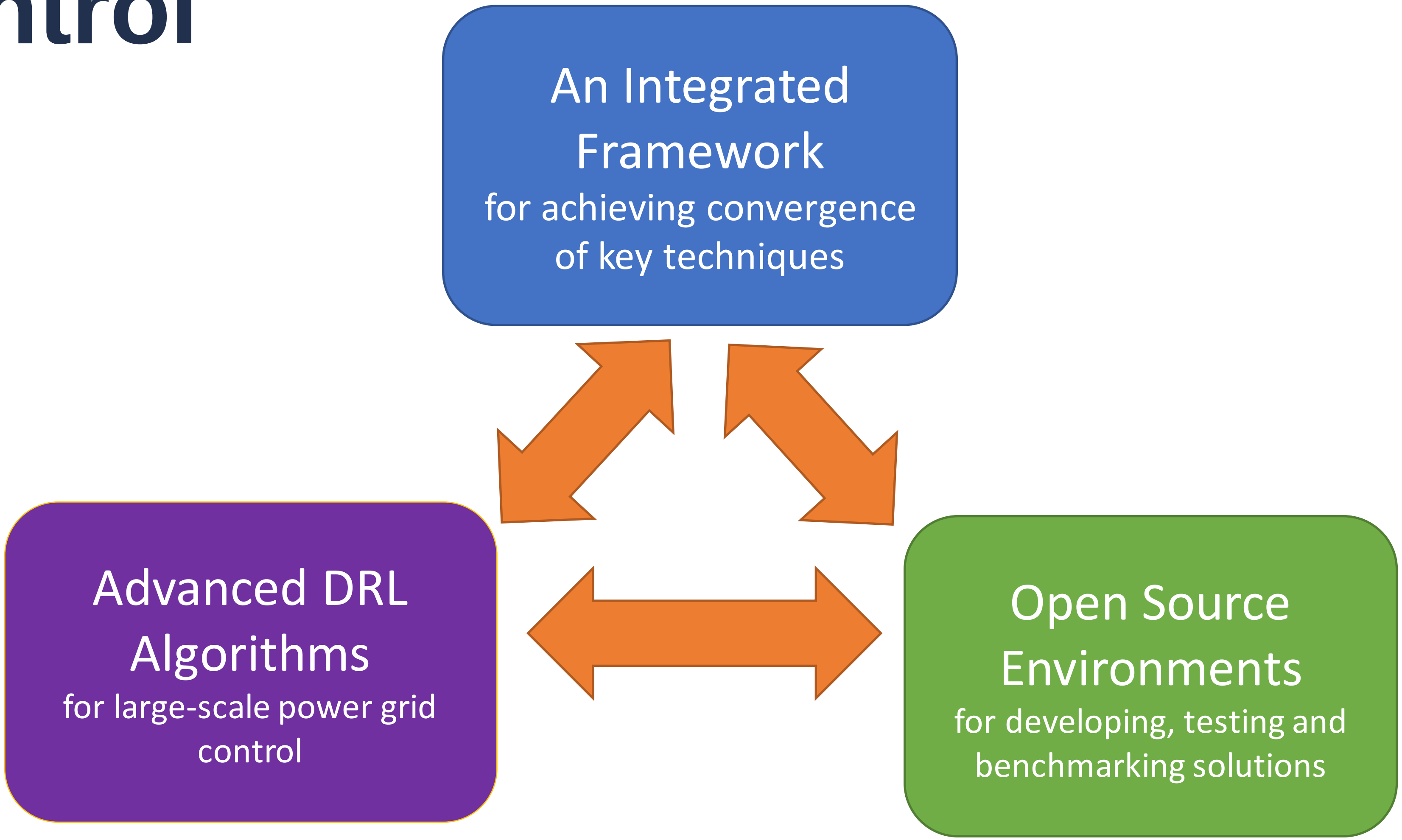
Key Challenges in DRL for Grid Control

Challenge	Note
Scalability	large system (observation space), large number of control points (action space) and scenarios (exploration space)
Adaptability	Adapt to changing operation conditions and network topology
Security	High security requirement (Insecure or outage is almost prohibitive)
Trustworthiness	Control actions or decisions should be trusted by operators and engineers
Training and test environments	• Grid environments are essential for DRL training and test; should be OpenAI gym-compatible • Grid simulator ≠ environment for DRL training and test
Reproducibility	• The complexity of DRL for grid control makes them hard to reproduce • Open-source is a proven, good practice for reproducibility

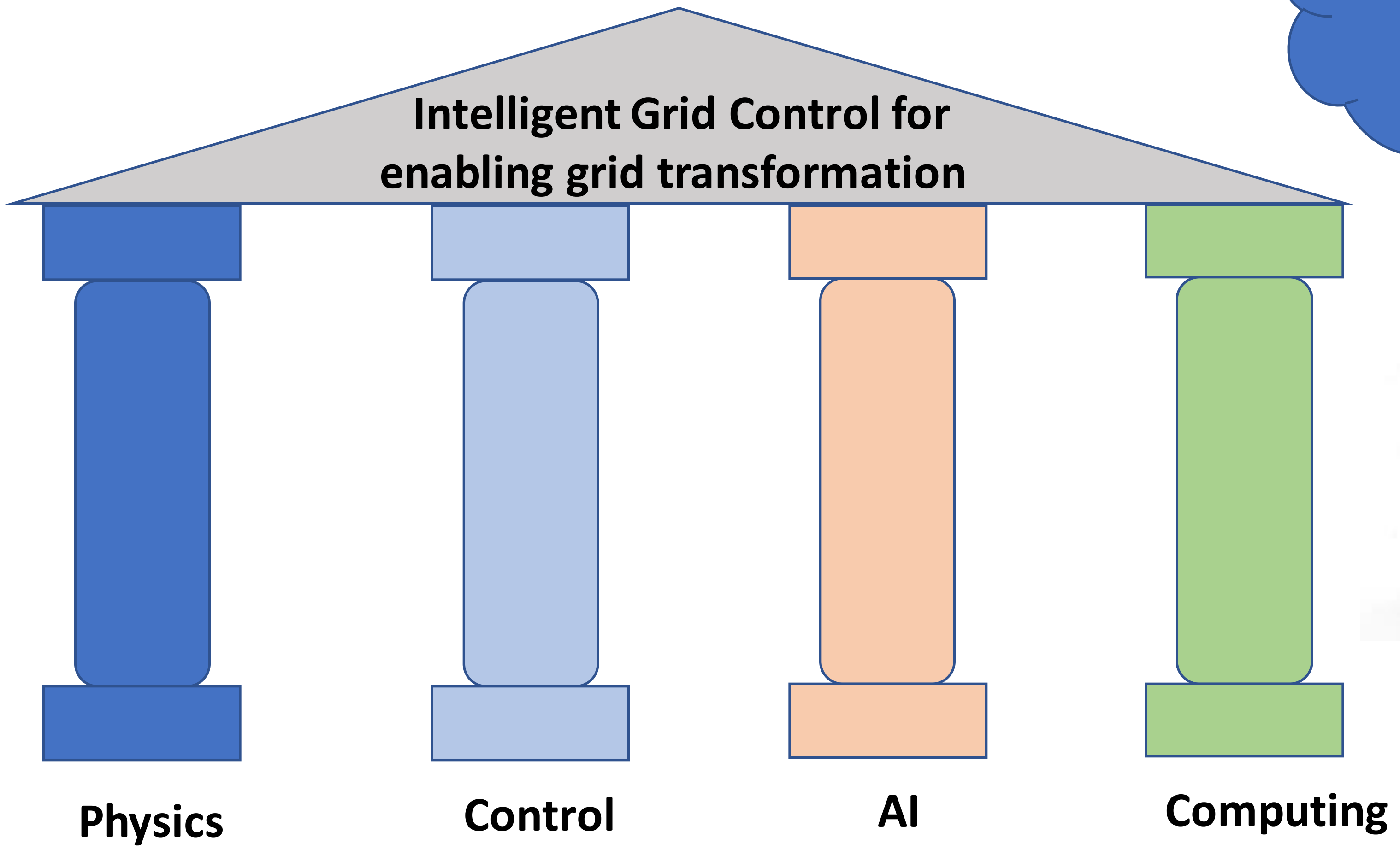
A single domain breakthrough is unlikely to address these challenges! → Need holistic solutions

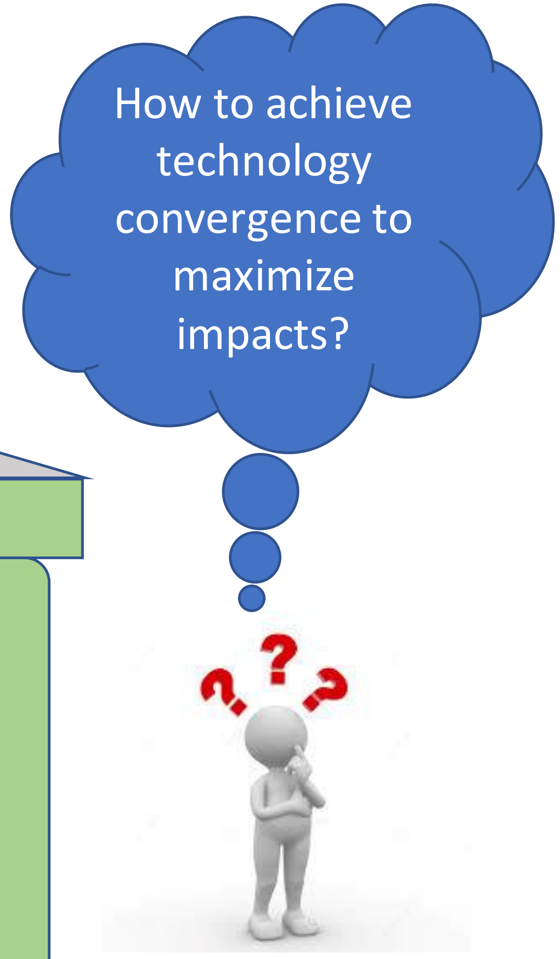
Convergence of AI, physics, computing and control for intelligent power system control

A Holistic Approach to Intelligent Grid Control

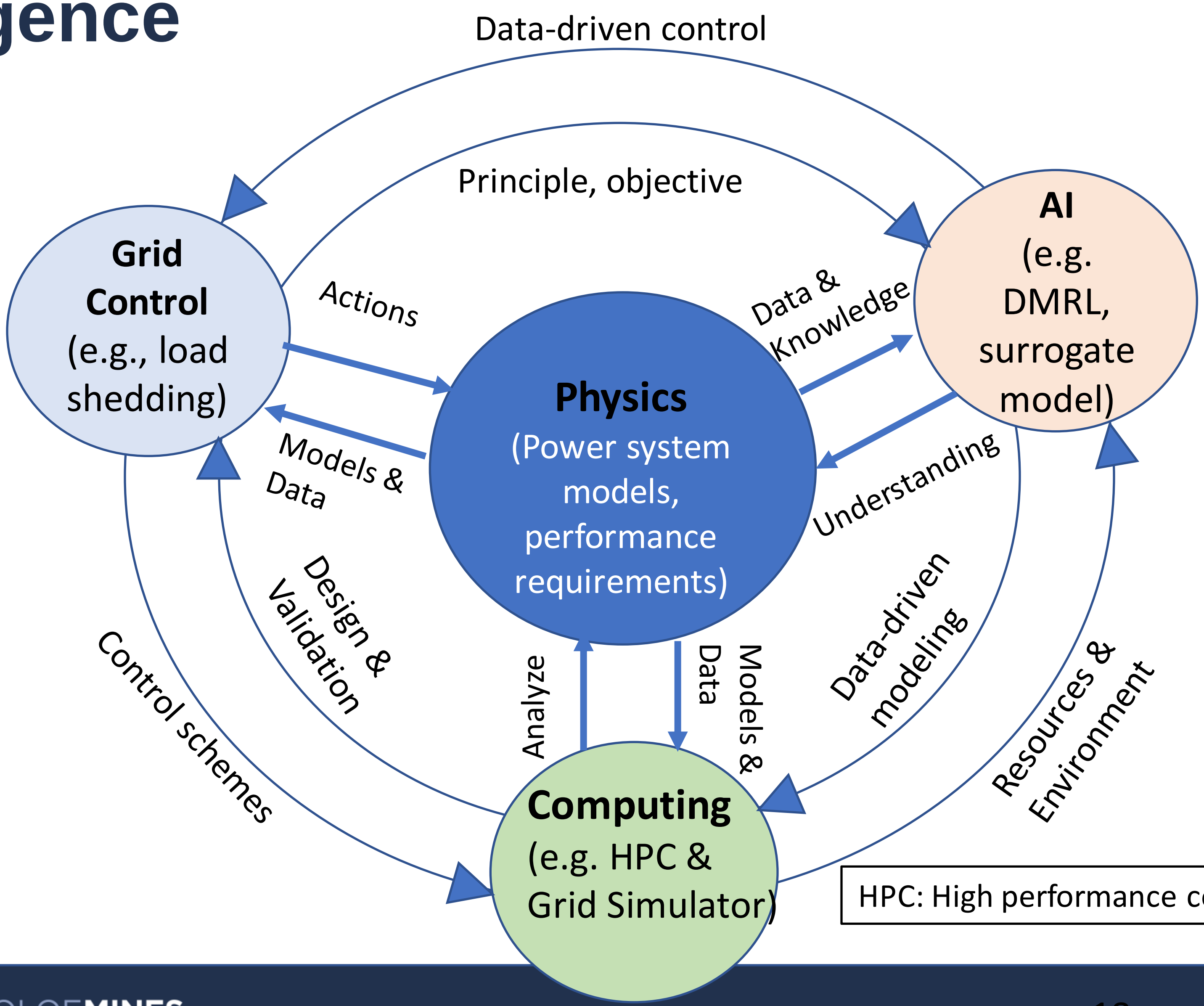


Four Technology Pillars





An Integrated Framework for Technology Convergence

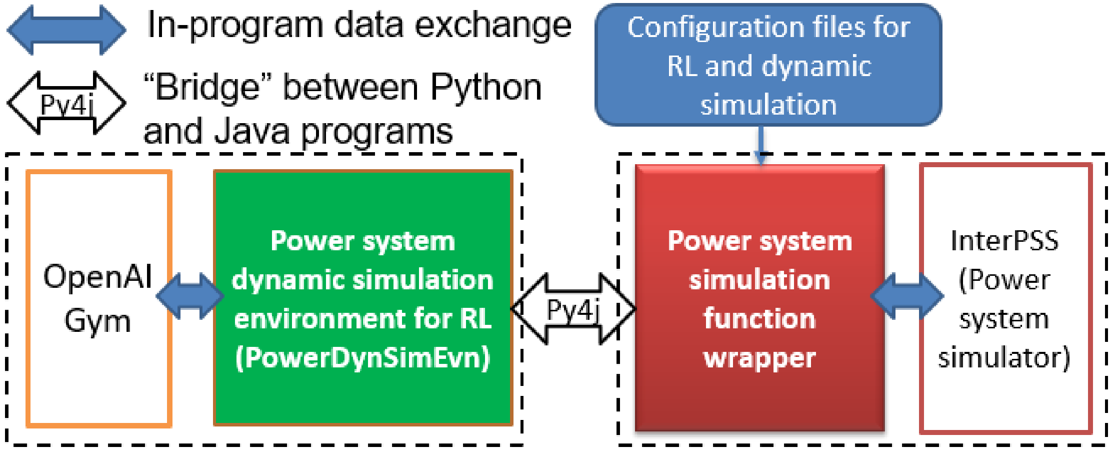


DMRL: Deep Meta-Reinforcement Learning

HPC: High performance computing

RLGC: A Lightweight, Open-source Environment for RL for Grid Dynamic Control

- The first open-source grid environment for developing, testing and benchmarking RL algorithms for grid dynamic control[1]
- Applications: load shedding for voltage control, generator tripping, dynamic breaking, controlled islanding (to be released), T&D coordination (to be released)



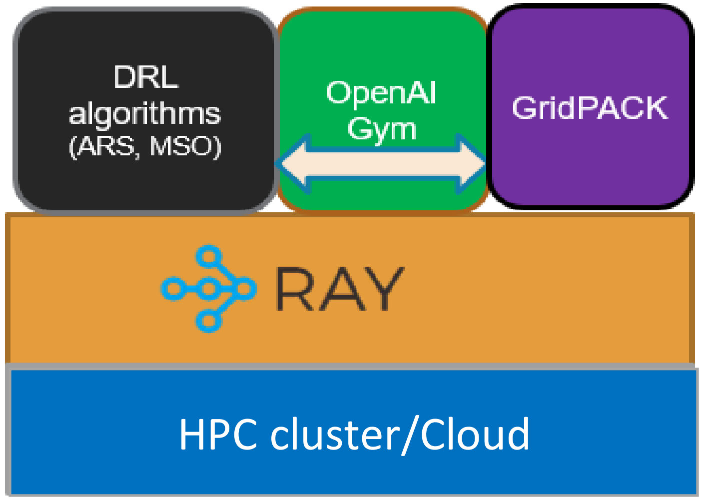
The Architecture of RLGC platform

Open-sourced on GitHub: https://github.com/RLGC-Project/RLGC

[1] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, Z. Huang. "Adaptive Power System Emergency Control Using Deep Reinforcement Learning," IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1171-1182, March 2020

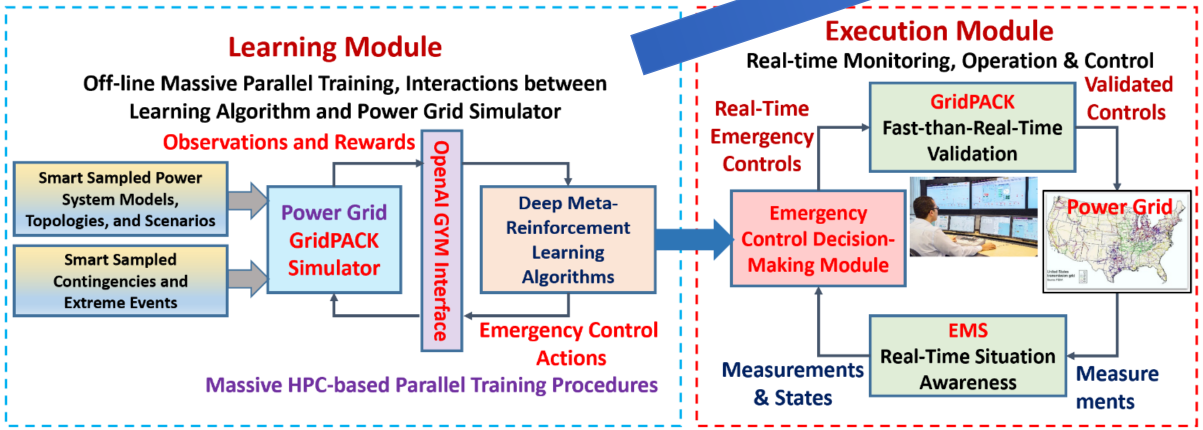
An HPC-based Grid Environment

- Scalable from laptop to HPC clusters/clouds with the RAY platform [1].
- GridPACK: HPC-based advanced grid dynamic simulator [2]
- The environment and DRL algorithms are open-sourced: https://github.com/pnnl/HADREC/



Architecture of the platform for training and testing

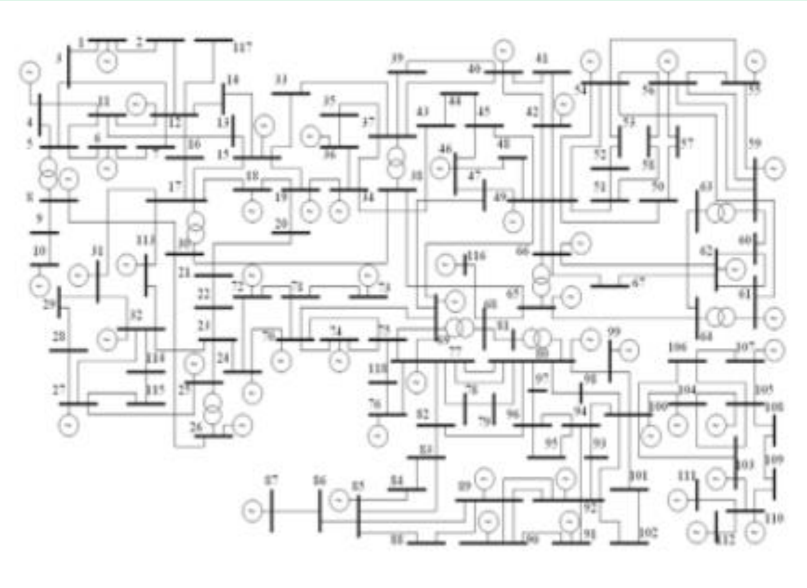
[1] https://www.ray.io/
[2] https://www.gridpack.org/wiki/index.php/Main_Page



HPC-based learning and execution modules for intelligent grid control

An Example: Load Shedding for Emergency Voltage Control

Dynamic simulation engine



IEEE 300-bus system model

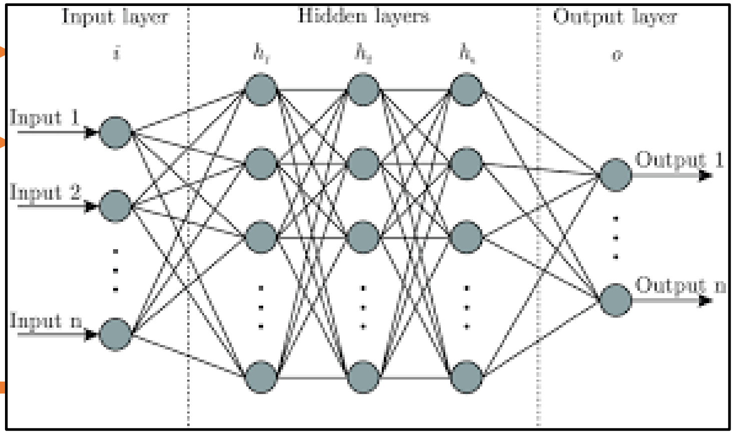
o_t: Observations

154 bus voltage magnitudes and 46 bus load levels

o_t

r_t (for training)

a_t



(for illustration only)

A Neural Network for representing agent's policy

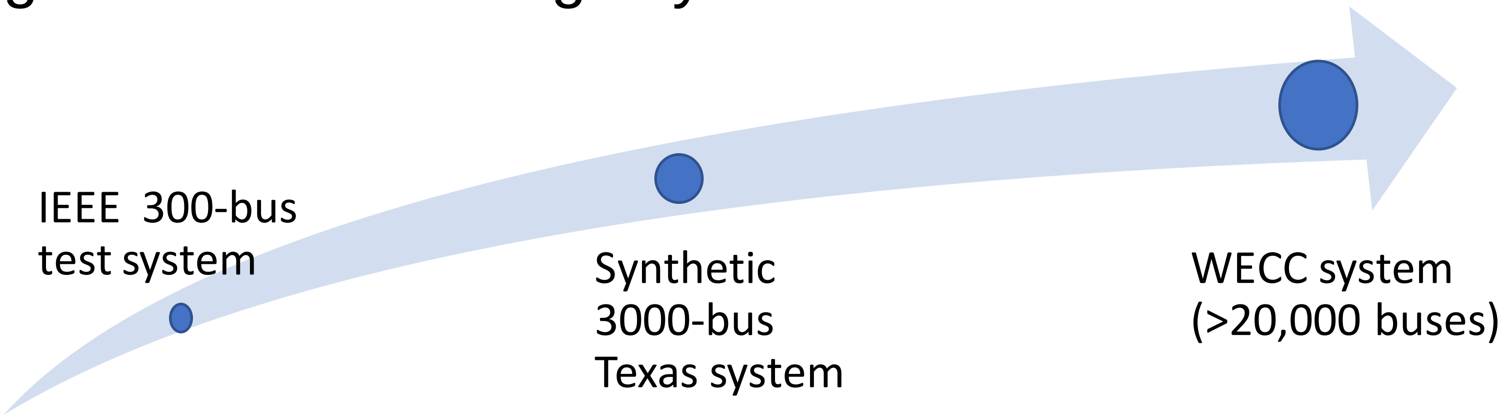
a_t: Actions

- 46 load substations could shed load.
- Each area, for each training time step, the load could be shed between 0% and 20% .
- The action space is 46.

R. Huang, Y. Chen, T. Yin, Q. Huang, J. Tan, W. Yu, X. Li, A. Li, Y. Du, "Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning," IEEE Transactions on Power Systems, 2022

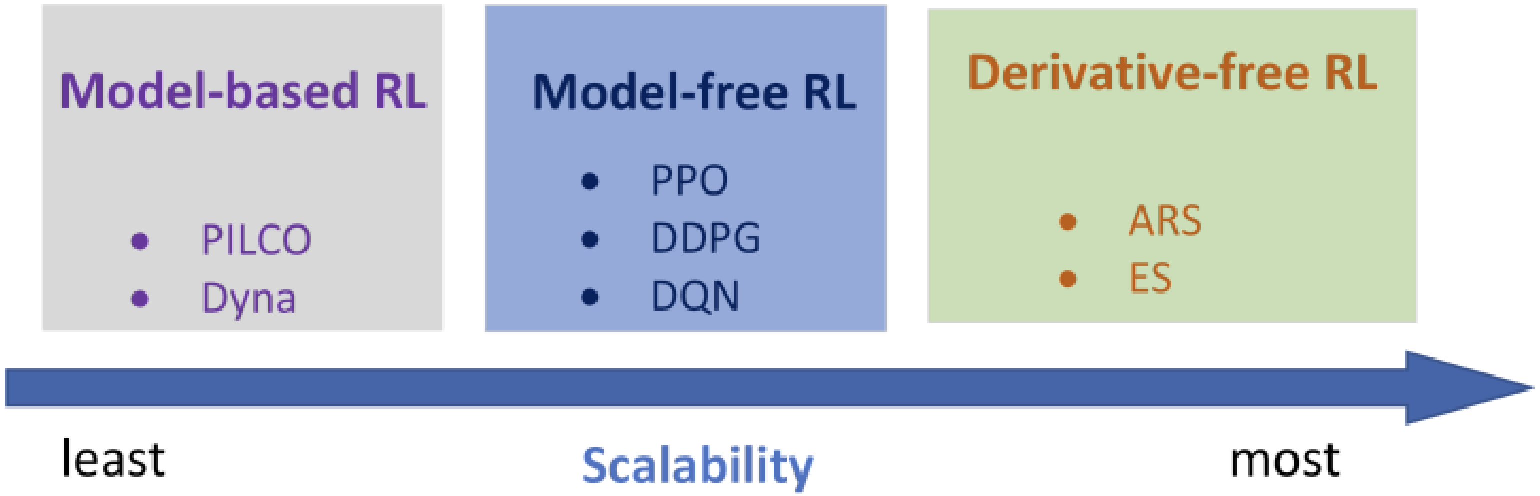
We Need Scalable RL Methods

- Training a RL agent for IEEE 39-bus system using the **DQN** algorithm on a laptop took about **2 days** [1]
  - Existing methods are very difficult, if not impossible, to scale up for controlling large power systems (e.g., >2000 buses)
- Rich Sutton: "The biggest lesson that can be read from 70 years of AI research is that **general methods that leverage computation are ultimately the most effective, and by a large margin**" [2]
- We target at real-world large systems→ **Scalable RL methods**



[1] Q. Huang, et al "Adaptive Power System Emergency Control using Deep Reinforcement Learning," in *IEEE Transactions on Smart Grid*. 2019
[2] http://incompleteideas.net/IncIdeas/BitterLesson.html

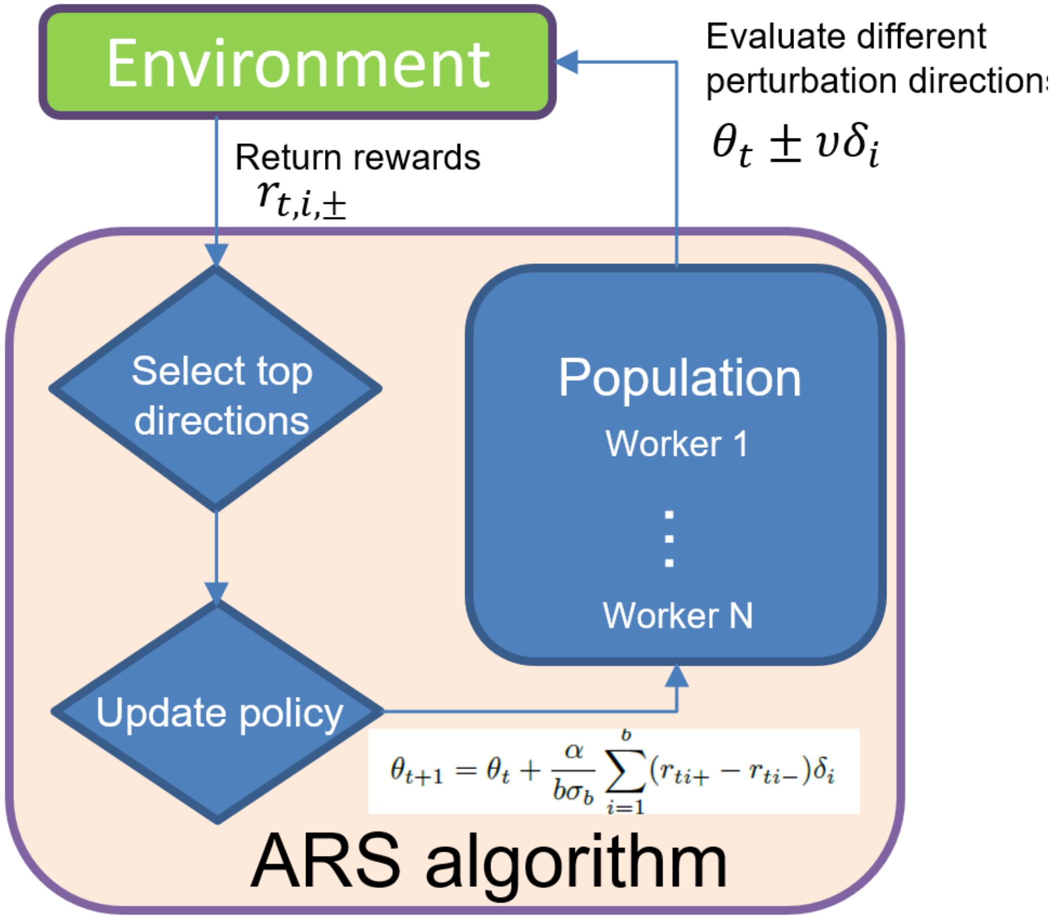
Three Types of RL Methods



R. Huang, Y. Chen, T. Yin, X. Li, A. Li, J. Tan, W. Yu, Y. Liu, Q. Huang. "Accelerated Derivative-free Deep Reinforcement Learning Based Load Shedding for Emergency Voltage Control," *in IEEE Transactions on Power Systems,* vol. 37, no. 1, pp. 14-25, Jan. 2022

Parallel Augmented Random Search Algorithm

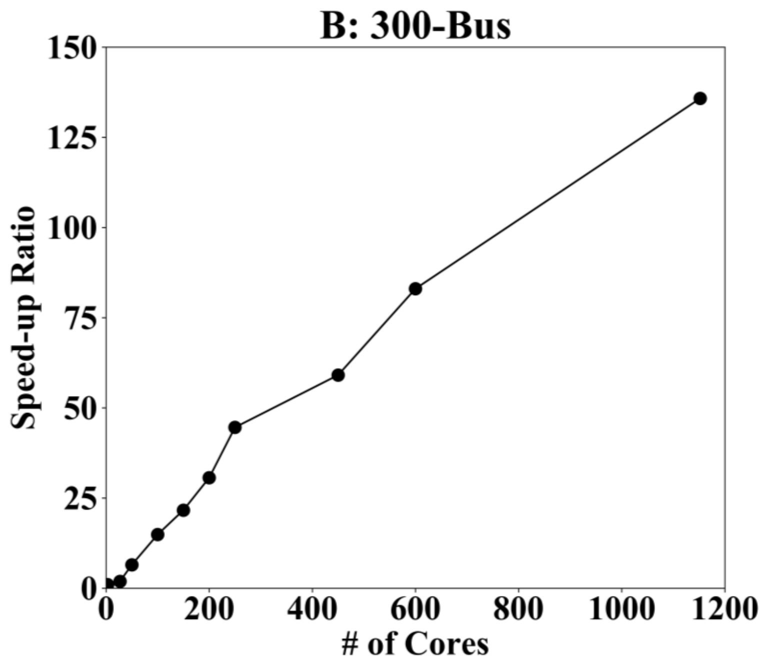
- Based on evolutionary strategy
- Unique features/advantages
  - Easy to **scale** and parallel (via workers)
  - Easy to **tune** (only 5 main hyper-parameters vs 20+ in existing DRL algorithms)
  - More **robust** for training (much less sensitive to random seeds and hyper-parameters)
  - Support **larger learning rates** to achieve faster training

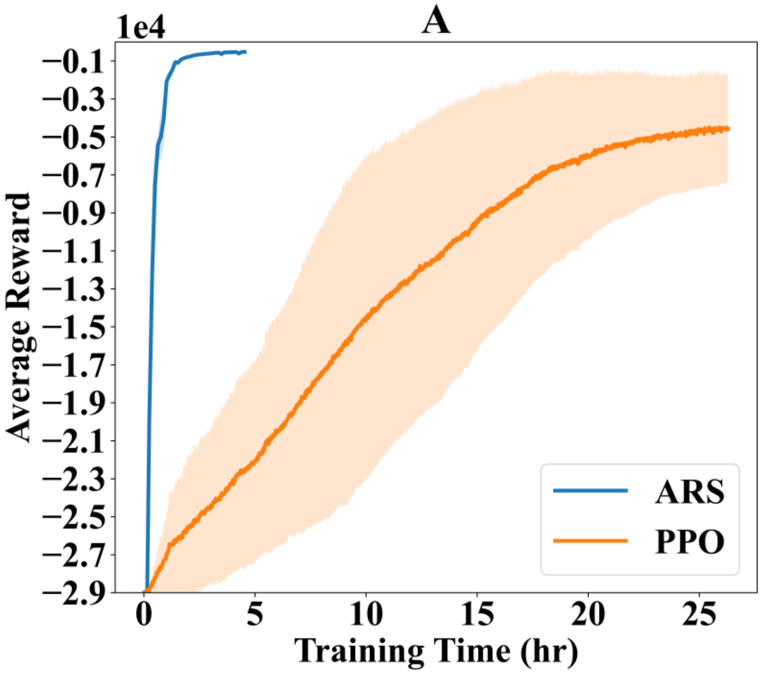


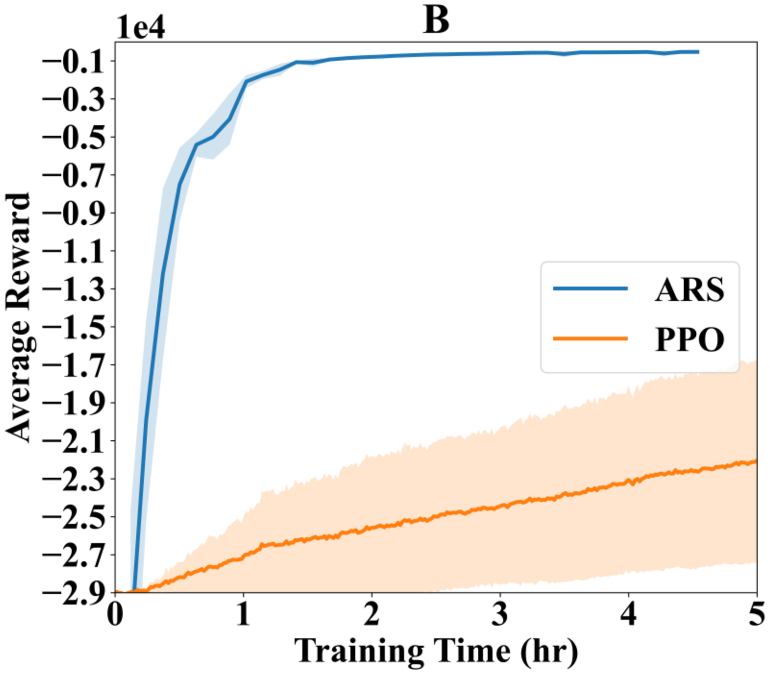
R. Huang, Y. Chen, T. Yin, X. Li, A. Li, J. Tan, W. Yu, Y. Liu, Q. Huang. "Accelerated Derivative-free Deep Reinforcement Learning Based Load Shedding for Emergency Voltage Control," *in IEEE Transactions on Power Systems,* vol. 37, no. 1, pp. 14-25, Jan. 2022

Parallel ARS Algorithm Test Results

- Emergency voltage control on the IEEE 300-bus system







PPO: Proximal Policy Optimization

High scalability of Parallel ARS

Much faster and more robust training with larger average rewards using Parallel ARS

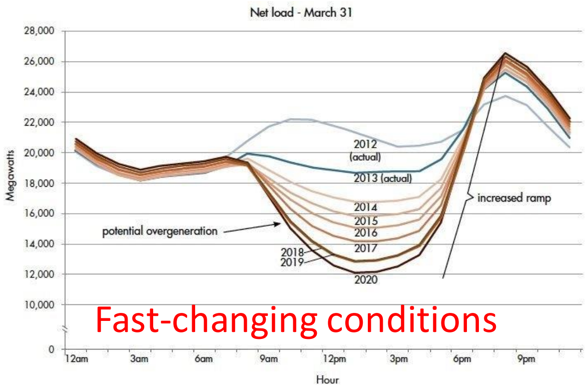
R. Huang, Y. Chen, T. Yin, X. Li, A. Li, J. Tan, W. Yu, Y. Liu, Q. Huang. "Accelerated Derivative-Free Deep Reinforcement Learning Based Load Shedding for Emergency Voltage Control," in IEEE Transactions on Power Systems, vol. 37, no. 1, pp. 14-25, Jan. 2022, doi: 10.1109/TPWRS.2021.3095179.

Learning a Context to Enhance Control Adaptiveness and Effectiveness

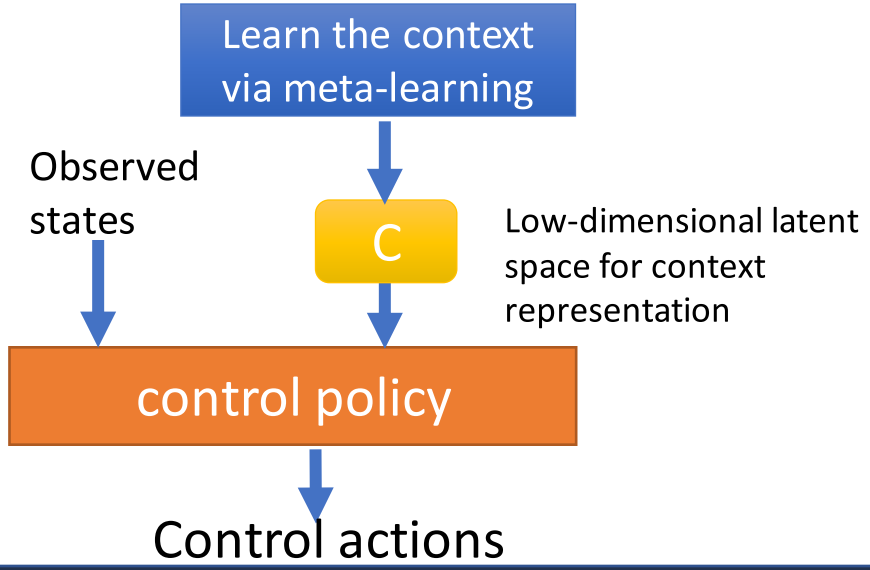
- We can enhance the control adaptiveness and effectiveness if we could efficiently construct a context for the control agent.
- !!Challenge: What should be the proper information for representing the context for different grid controls?
- Learn the context automatically from data and learning experiences via **meta-learning**
- Mathematical formulation:

$$\theta^*, c(\mathcal{E})^* = \arg\max_{\theta, c(\mathcal{E})} \mathbb{E}_{\mathcal{E} \in \boldsymbol{P}(E_i)} \left[J\left(c(\mathcal{E}), \theta\right)\right]$$

θ – neural network weights for control policy
ε – one training environment
$c(\varepsilon)$ – context representation for environment ε
$\boldsymbol{P}(E_i)$ – a distribution of training environments



Source: CAISO

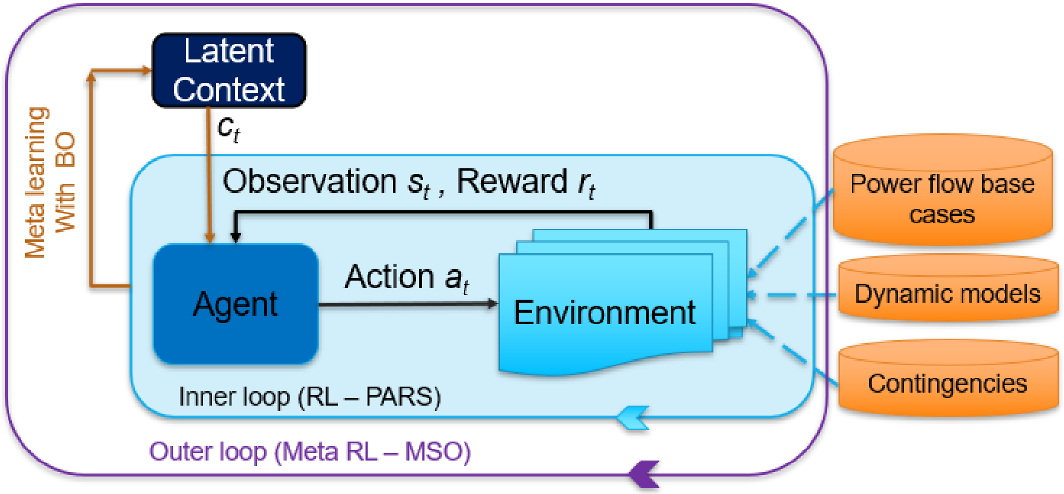


Realization of Deep Meta-Reinforcement Learning

- Reformulate (1) as (2) and (3) and solve them in an iterative manner (similar to the coordinate descent method):
  - **RL:** Learn the control policy θ while leveraging the latest context $c(\varepsilon)$ by solving equation (2)
  - **Meta-RL:** Learn the <u>low dimensional</u> latent context c using Bayesian Optimization to solve (3)
- Fast adaptation by only solving (3)

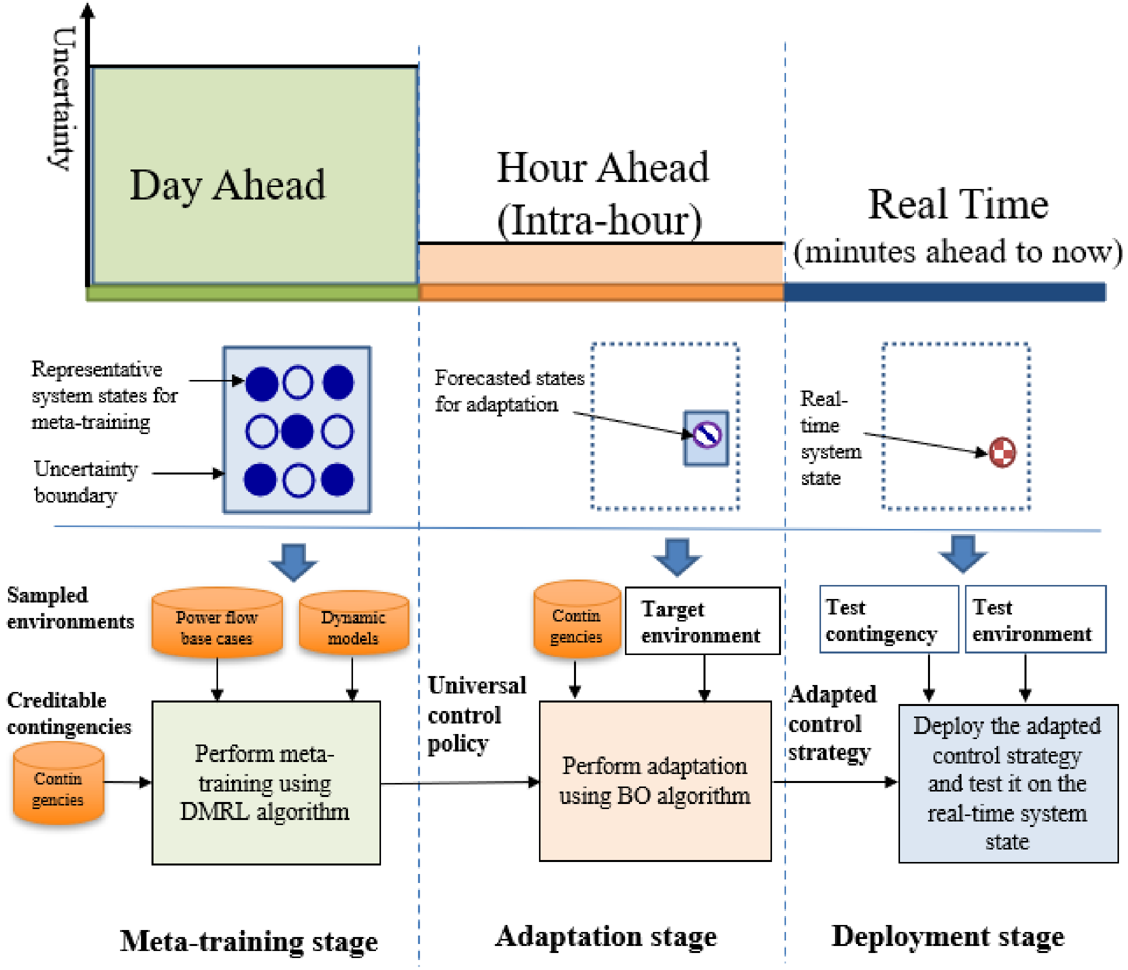
$$\boxed{\theta^*, c(\mathcal{E})^* = \arg\max_{\theta, c(\mathcal{E})} \mathbb{E}_{\mathcal{E}\in \boldsymbol{P}(E_i)}\left[J\left(c(\mathcal{E}), \theta\right)\right] \qquad (1)}$$

$$\begin{cases} \theta_{k+1} = \arg\max_{\theta} \mathbb{E}_{\mathcal{E}\in \boldsymbol{P}(E_i)}\left[J\left(c(\mathcal{E})_k, \theta\right)\right] & (2) \\ c(\mathcal{E})_{k+1} = \arg\max_{c} J\left(c, \theta_k\right), \forall \mathcal{E} \in \boldsymbol{P}(E_i) & (3) \end{cases}$$



R. Huang, Y. Chen, T. Yin, Q. Huang, J. Tan, W. Yu, X. Li, A. Li, Y. Du, “Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning,” IEEE Transactions on Power Systems, 2022

Our DMRL Algorithm Fits Well into Power System Operation Time Frames and Procedures



IEEE 300-Bus Power Flow Conditions for Training and Test

We use this small system to compare our algorithms with existing solutions that are difficult to scale

TABLE I
POWER FLOW CONDITIONS FOR TRAINING

Training power flow case	Generation	Load
1	Total 22929.5 MW (100%)	Total 22570.2 MW (100%)
2	120% for all generators	120% for all loads
3	135% for all generators	135% for all loads
4	115% for all generators	150% for loads in Zone 1

TABLE II
POWER FLOW CONDITIONS FOR ADAPTATION AND TESTING

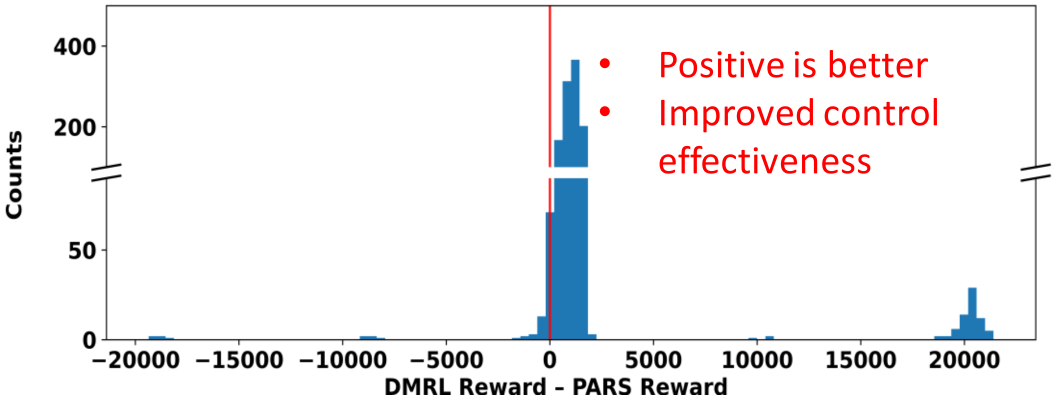
Adaptation/testing power flow case	Adaptation		Testing	
	Generation	Load	Generation	Load
1	90% for all generators	90% for all loads	92.4% for all generators	92.4% for all loads
2	110% for all generators	110% for all loads	107.7% for all generators	107.7% for all loads
3	115% for all generators	115% for all loads	117.2% for all generators	117.2% for all loads
4	125% for all generators	125% for all loads	122.5% for all generators	122.5% for all loads
5	140% for all generators	140% for all loads	142.1% for all generators	142.1% for all loads
6	95% for all generators	82.8% for loads in Zone 1	97.1% for all generators	85.32% for loads in Zone 1
7	107% for all generators	124.4% for loads in Zone 1	104.6% for all generators	121.5% for loads in Zone 1
8	110% for all generators	134.3% for loads in Zone 1	112.3% for all generators	137.2% for loads in Zone 1
9	119% for all generators	159.7% for loads in Zone 1	121.1% for all generators	162.6% for loads in Zone 1

R. Huang, Y. Chen, T. Yin, Q. Huang, J. Tan, W. Yu, X. Li, A. Li, Y. Du, "Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning," IEEE Transactions on Power Systems, accepted, 2022

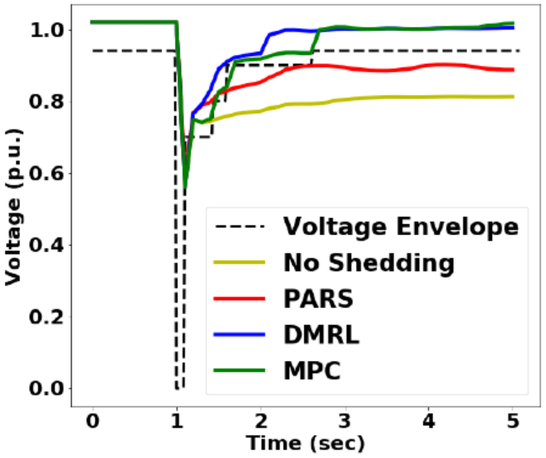
Deep Meta-Reinforcement Learning Test Results

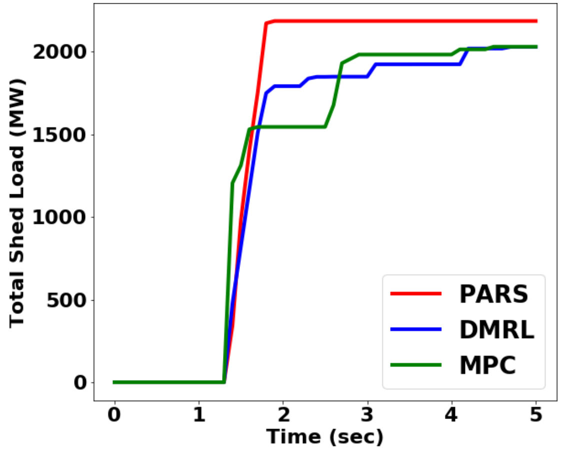
	DMRL	PARS	MPC*
Training	11.6 hours	9.5 hours	N/A
Adaptation	5.3 mins	N/A	N/A
Solution time	0.7 sec	0.7 sec	63.3 sec

*MPC: Model-predictive control



All test scenarios are unseen during training

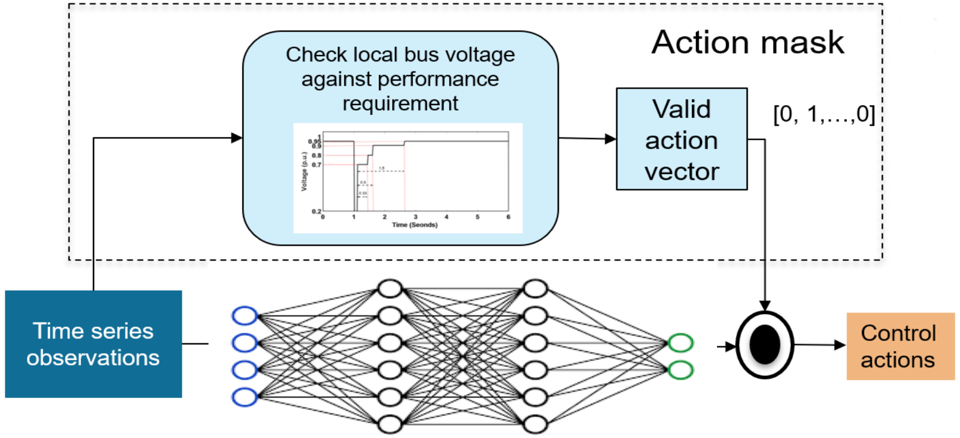




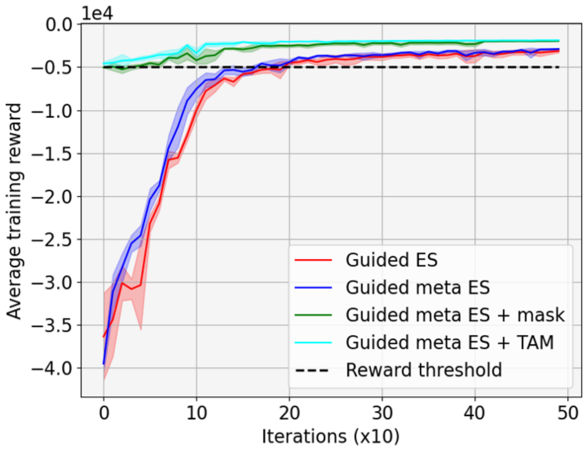
R. Huang, Y. Chen, T. Yin, Q. Huang, J. Tan, W. Yu, X. Li, A. Li, Y. Du, “Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning,” IEEE Transactions on Power Systems, 2022

Physics-informed DMRL Enhances Training Efficiency and Control Robustness

- Power system community have developed vast amount of domain knowledge in forms of physics laws, standards, rules, and performance requirements.
- Physics-informed DMRL: we incorporated system performance requirements as a trainable action mask (TAM) into the agent and significantly improved its sampling efficiency and robustness[1].



Incorporate prior knowledge into the agent with a fixed action mask [1]



Training results

Method	Average test reward	No. of failed cases
ARS	-1.27×10^4	72
Guided ES	-5.6×10^3	17
Guided meta ES	-4.3×10^3	12
Guided meta ES + mask	-2.8×10^3	8
Guided meta ES + TAM	-1.89×10^3	3
MPC	-1.82×10^3	3

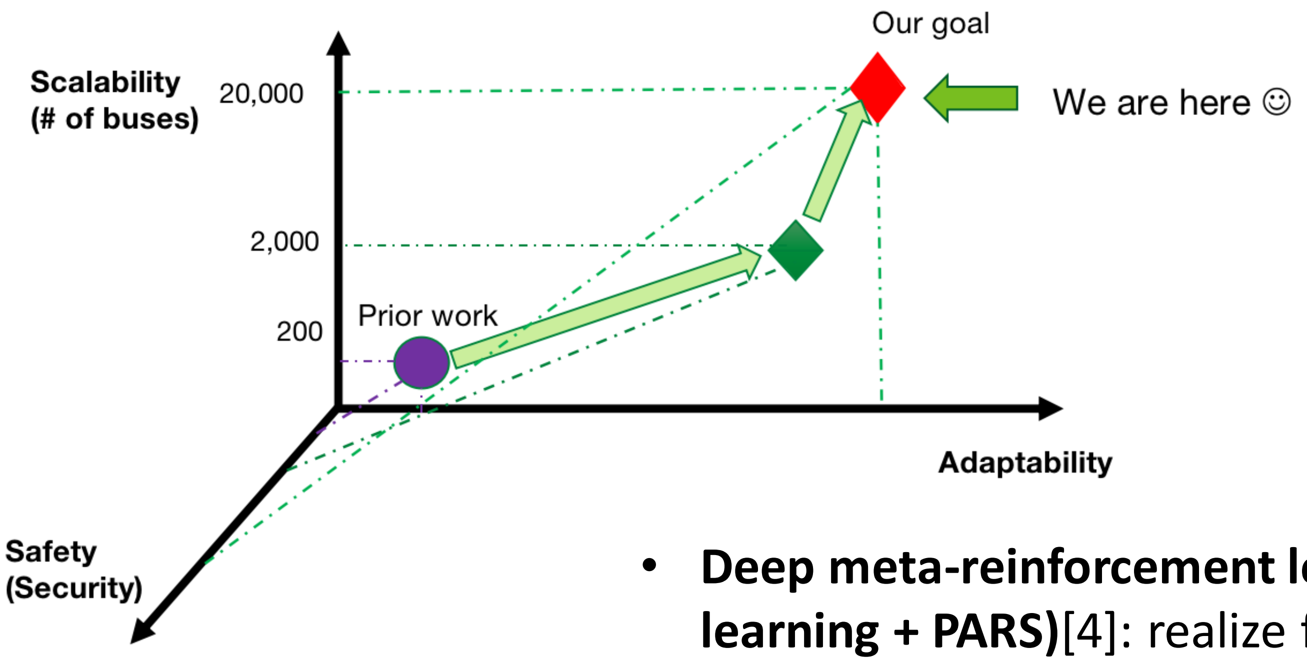
Test results

[1] Y. Du, **Q. Huang**, R. Huang; T. Yin; J. Tan; W.Yu; X. Li, "Physics-informed Evolutionary Strategy based Control for Mitigating Delayed Voltage Recovery," in IEEE Transactions onPower Systems, 2021

A Summary of Our Algorithms

- **Parallel Augmented Random Search (PARS) algorithm**[1],
- High performance power system simulation platform GridPACK

- **Physics-informed PARS** [2]: incorporate physics knowledge through a trainable action mask
- **Safe PARS** [3]: control barrier function + PARS



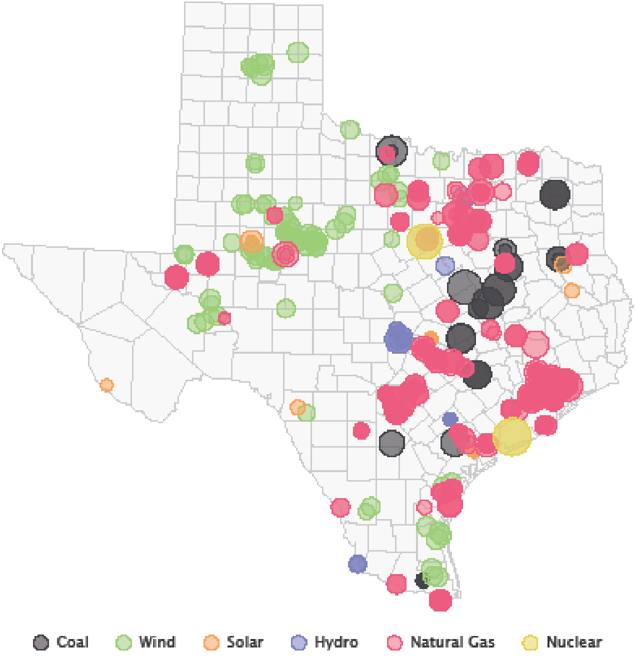
- **Deep meta-reinforcement learning (meta-learning + PARS)**[4]: realize fast adaptation (~5 mins) of control policies to changing grid conditions

[1] R. Huang, et al "Accelerated Derivative-free Deep Reinforcement Learning Based Load Shedding for Emergency Voltage Control," *IEEE Trans. on Power Systems*, 2021
[2] D. Yan, et al "Physics-informed Evolutionary Strategy based Control for Mitigating Delayed Voltage Recovery", *IEEE Trans. on Power Systems*, 2022
[3] T. Vu, et al. "Safe Reinforcement Learning for Emergency Load Shedding of Power Systems." In Proc of IEEE PES General Meeting 2021
[4] R. Huang, et al, "Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning, *IEEE Trans. on Power Systems*, 2022

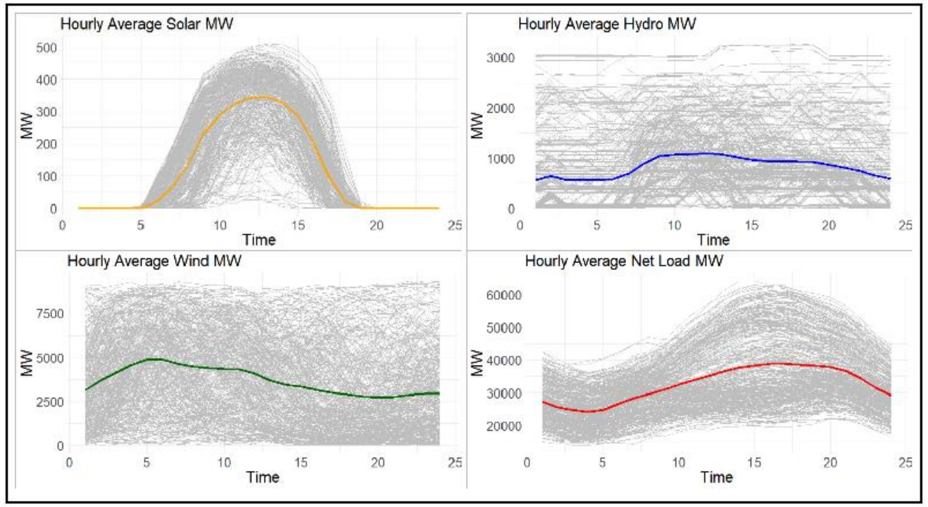
Test Results and Demo on Large Systems

A Synthetic 3000-bus Texas System

Single-line diagram



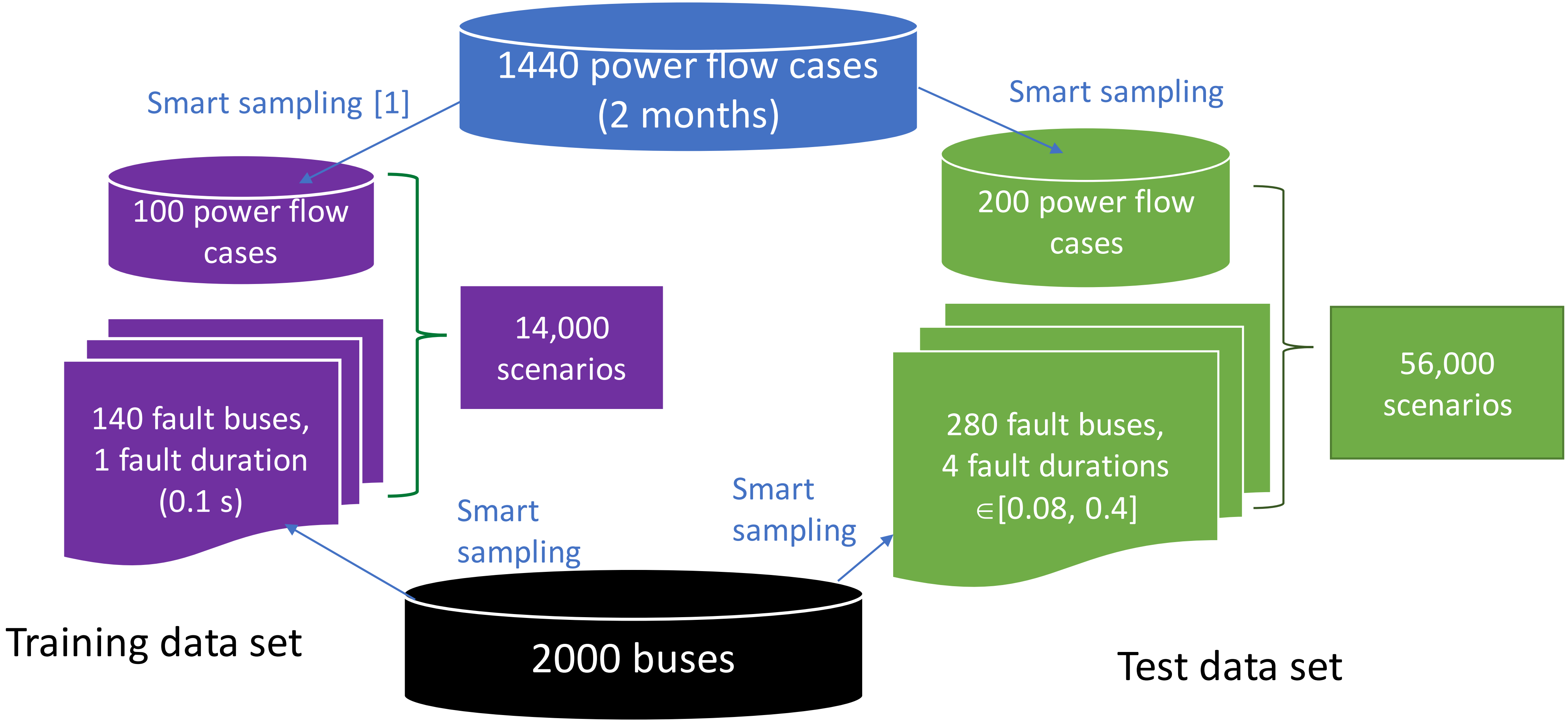
Generation mix



Hourly renewable outputs and net load demands

Source: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg2000/

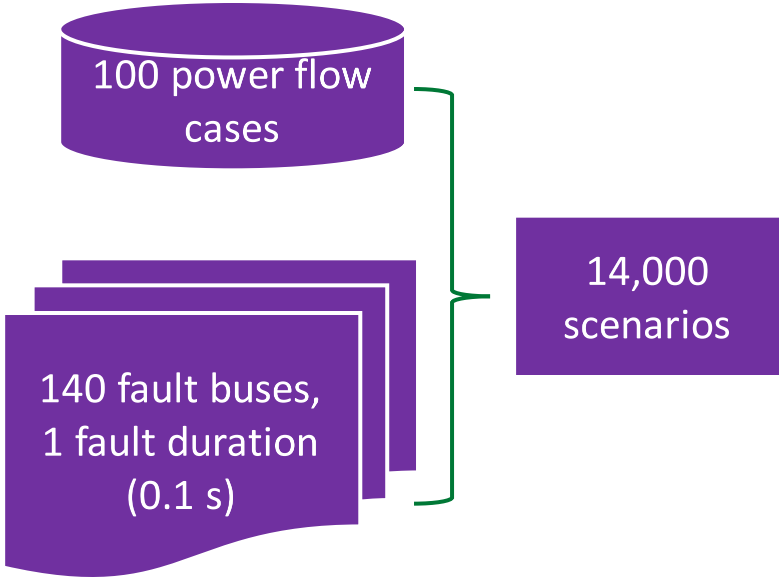
Texas 3000-bus System Dataset for Training and Testing

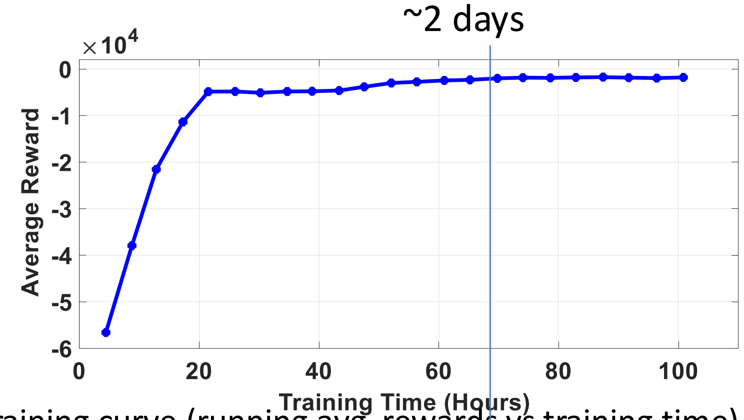


[1] X. Sun *et al.*, "Smart Sampling for Reduced and Representative Power System Scenario Selection," in *IEEE Open Access Journal of Power and Energy*, vol. 8, pp. 293-302, 2021,

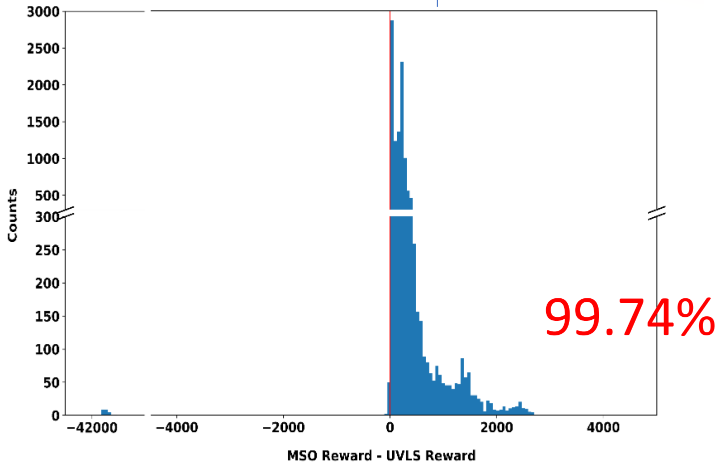
Texas 3000-bus System Off-line Training Results

- Action: load shedding fractions at 258 load buses, continuous, range [-0.2, 0]
- Observation: 726 bus voltage magnitudes + 258 substation load demand



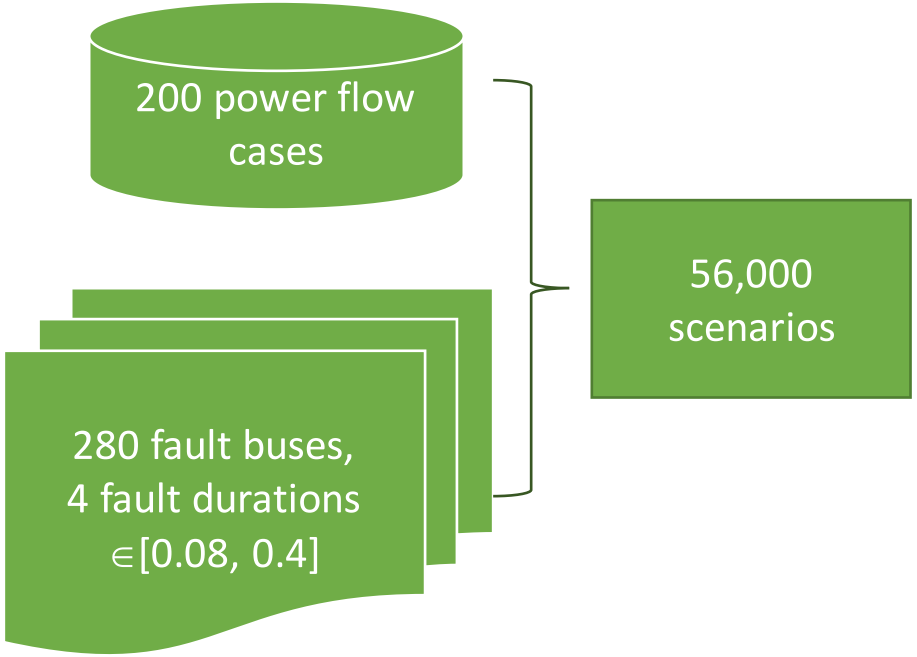


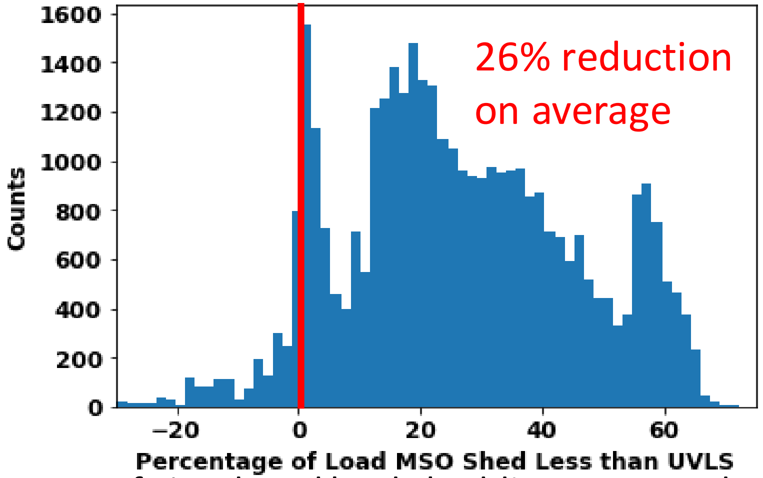
Training curve (running avg. rewards vs training time)



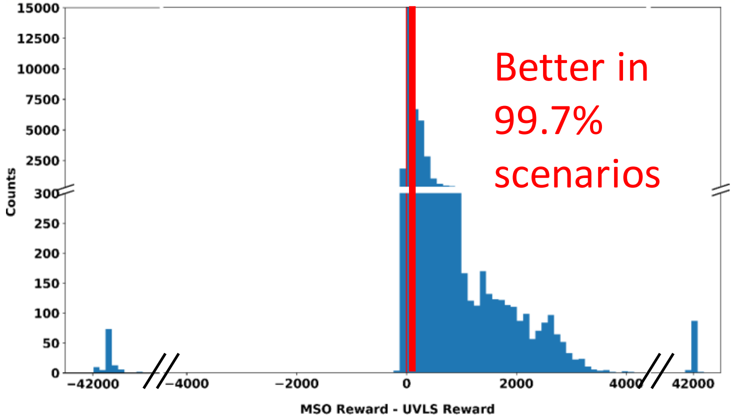
Objective value (total reward) differences (positive is better)

Off-line Testing Results





Histogram of % reduced load shedding compared with ruled-based UVLS (positive is better)

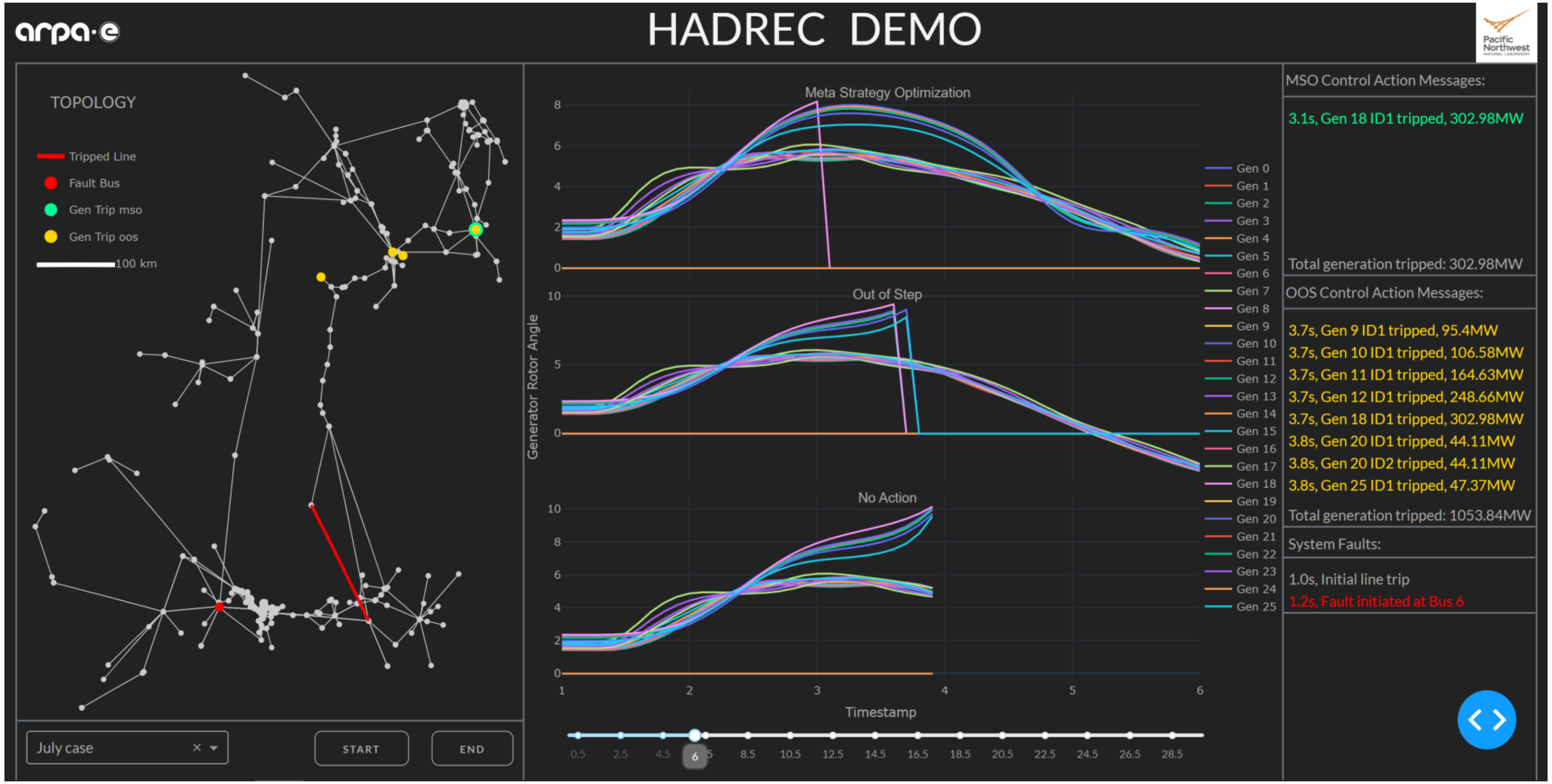


Objective value (total reward) differences (positive is better)

Demo on a Real Large System

- In the ARPA-E HADREC project, the platform and DMRL algorithm are integrated with V&R's real-time situational awareness tool (ROSE)
- Trained and tested on a large system based on real-world EMS snapshots provided by PacifiCorp
- We considered three emergency controls:
  - Load shedding for voltage control
  - Generator tripping for transient stability (next slide)
  - Controlled islanding for avoiding system-wide instability

Case study of a PacifiCorp system

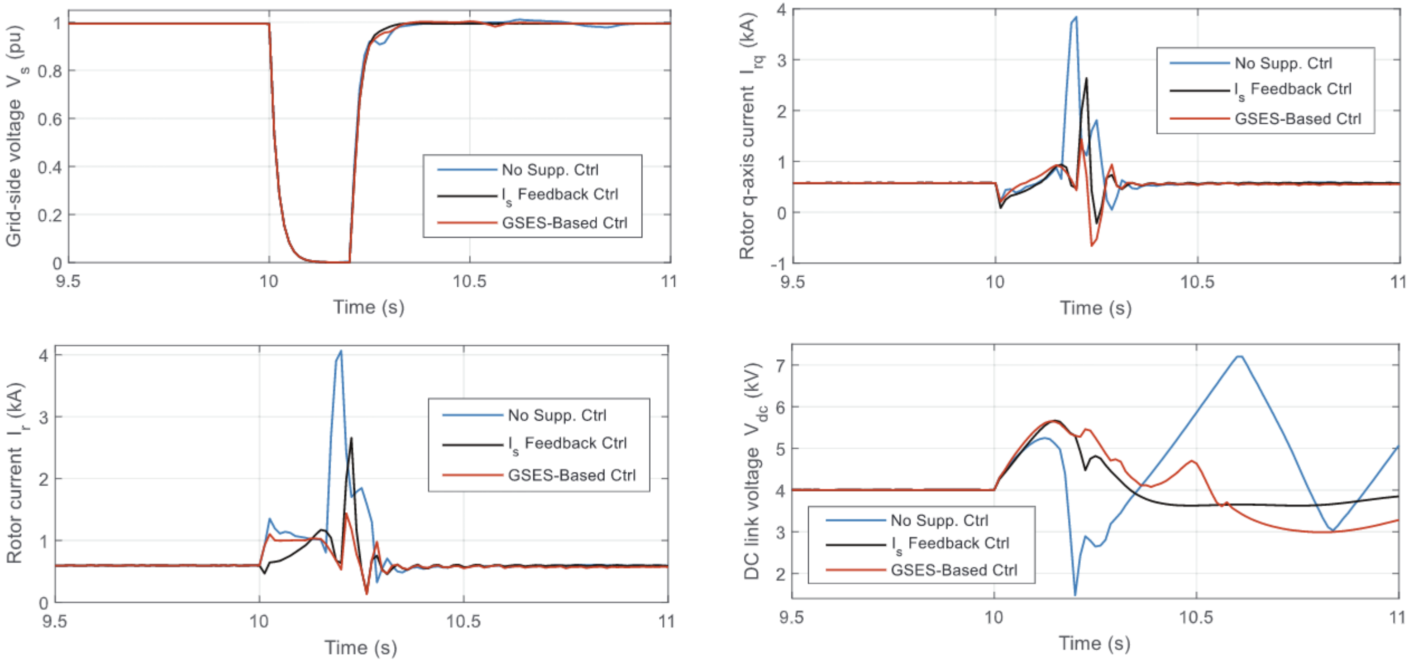


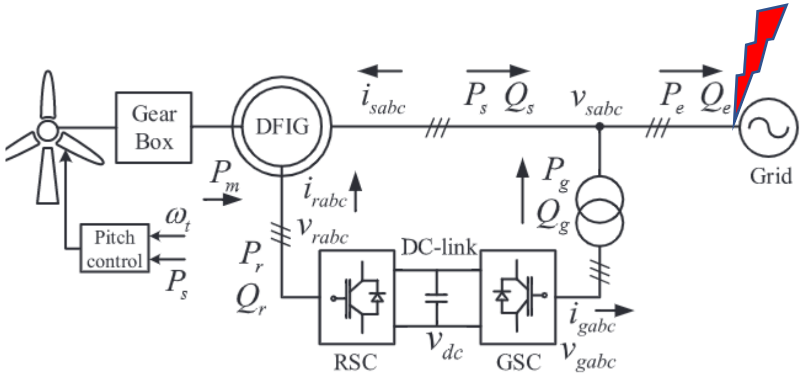
- HADREC: 1 generator tripped, 300MW total
- Out of Step (OOS): 8 generators tripped, 1 GW total
- ~20% improvement in responding time and 70% reduction in tripped generator output

Other Applications

Improve DFIG Performance in Grid Fault Scenarios

- Reduce the DFIG rotor over-current and DC-link over-voltage under different fault conditions

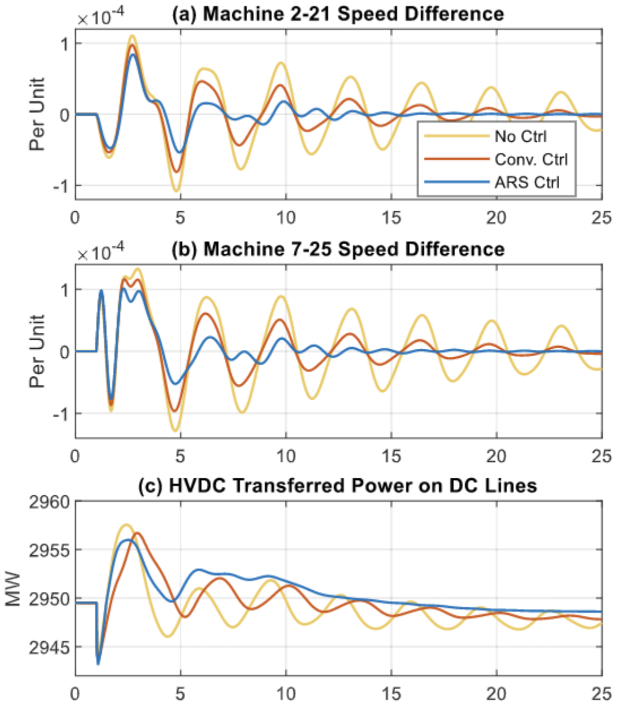




Gao, W., Fan, R., Huang, R., Huang, Q., Du, Y., Qiao, W., Wang, S., Gao, D.W.: Improving DFIG performance under fault scenarios through evolutionary reinforcement learning based control. *IET Gener. Transm. Distrib.* 16, 3825– 3836 (2022).

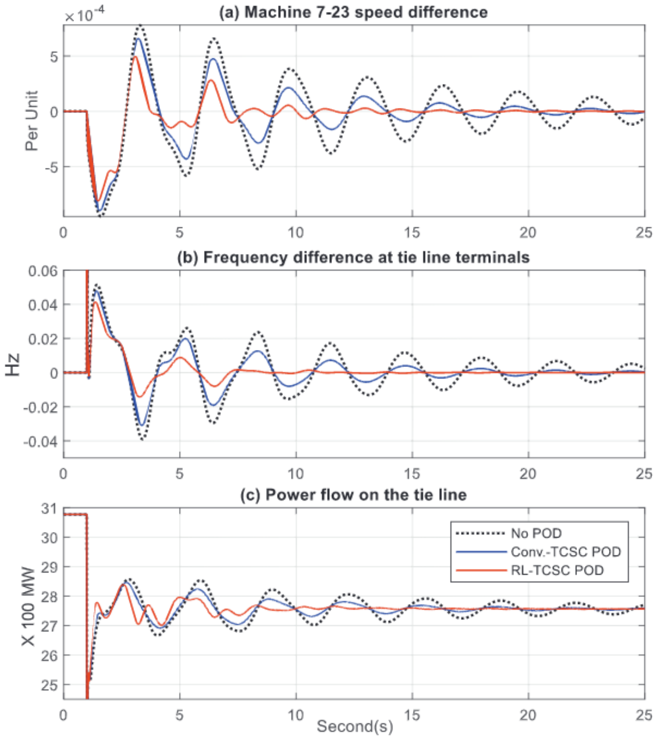
Improve Inter-area Damping Control through HVDC and FACTS

HVDC



Test results on the MinniWECC system from [1]

FACTS



Test results on the MinniWECC system from [2]

[1] W. Gao, R. Fan, R. Huang, Q. Huang, W. Gao, L. Du. “Augmented random search based inter-area oscillation damping using high voltage DC transmission. Electric Power Systems Research, 216, 109063, 2023
[2] R. Huang, W. Gao, R. Fan, Q. Huang. “Damping inter-area oscillation using reinforcement learning controlled TCSC.” IET Gener. Transm. Distrib. 16, 2265–2275 (2022).

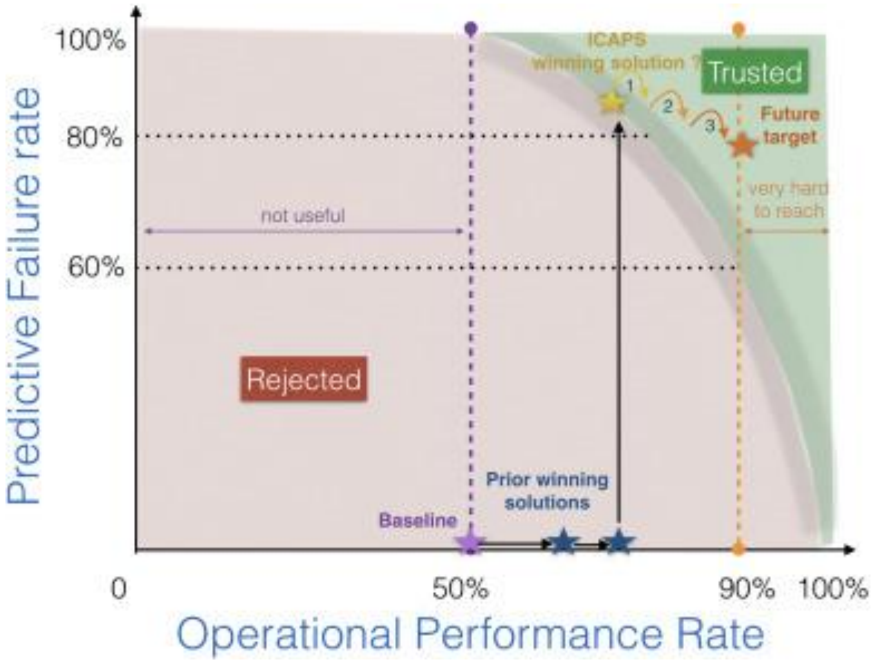
Key takeaways

- The latest AI show “generalist” capabilities – the same (very similar) NN model and DRL algorithm can be applied to different domains and applications
- This helps reduce extensive work in special designs based on strong domain expertise.

Future Work and Perspectives

Trustworthy DRL for Grid Control

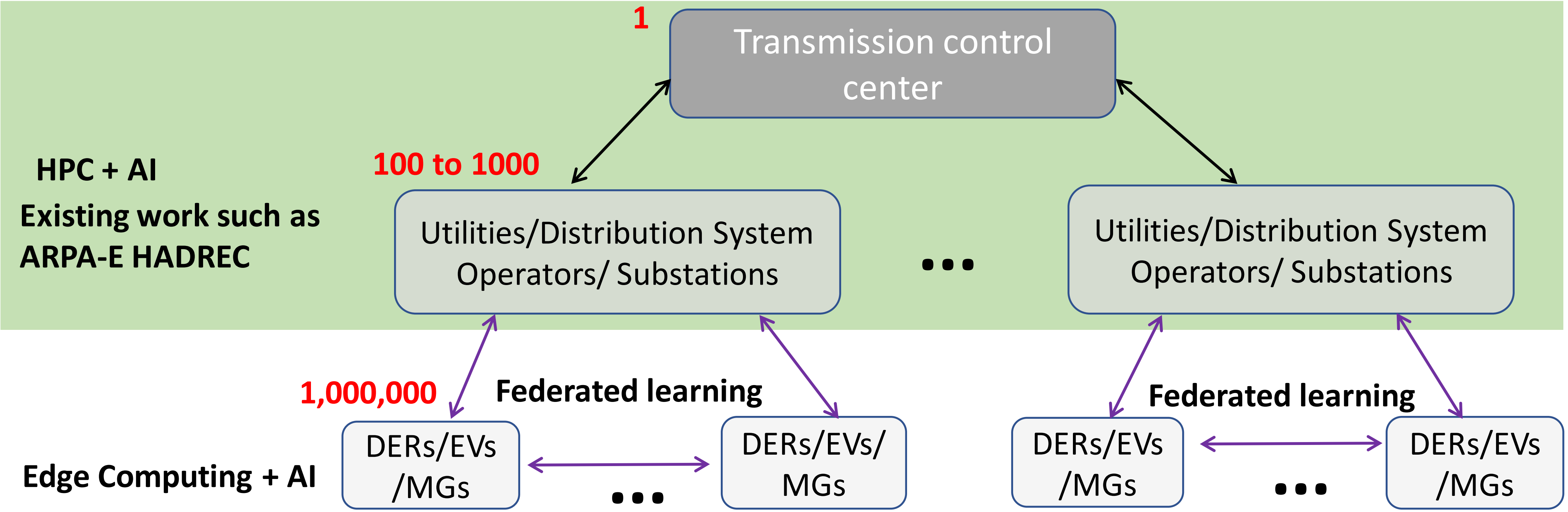
- Human-AI collaboration for intelligent grid control
- How to make DRL agent more trustworthy?
  - Good performance
  - Know its know limitations and alert human operators in advance
- "Learning to run power network with trust" formulation and competition [1]
- Research directions
  - Reinforcement learning from human feedback (RLHF)
  - Off-line reinforcement learning from historical human control actions
  - Explainable reinforcement learning (XRL)



[1] Marot, Antoine, Benjamin Donnot, Karim Chaouache, Adrian Kelly, Qiuhua Huang, Ramij-Raja Hossain, and Jochen L. Cremer. "Learning to run a power network with trust." *Electric Power Systems Research* 212 (2022): 108487.

Distributed Control with Edge AI and Coordination with Centralized Control

- Distributed control with edge computing and AI helps manage millions of DERs.
- Coordinating centralized and distributed control as well as computing is critical for large-scale clean energy integration and FERC 2222 compliance.



Other Potential Applications

- AI-Copilot for system operators
- Next-generation decision support tools for preventive control and network optimization
  - Optimal transmission switching [1]
- Mitigation plan for N-1-1 and N-2 contingencies during planning
- AI-bot for dispatcher(Operator) training system (DTS/OTS)
  - Generate training scenarios and optimal control actions

[1] Yoon, Deunsol, et al. "Winning the l2rpn challenge: Power grid management via semi-markov afterstate actor-critic." *International Conference on Learning Representations*. 2021.

A Summary

- Intelligent power system control is key to safeguard the grid while enabling grid transformation and modernization.
- Solving such a grand challenge requires the convergence of Physics, AI, Control and Computing.
- We developed an integrated framework, open-sourced environments and advanced DRL algorithms.
- We showed promising results on large systems.
- Trustworthy AI, human-AI collaboration and AI-enhanced T&D coordination are important future research directions.

Publications

1. A. Marot, B. Donnot, K. Chaouache, A. Kelly, **Q. Huang**, R. R. Hossain, & J. L. Cremer. “Learning to run a power network with trust.” Electric Power Systems Research, 212, 108487, 2022
2. Gao, W., Fan, R., Huang, R., Huang, Q., Du, Y., Qiao, W., Wang, S., Gao, D.W.: Improving DFIG performance under fault scenarios through evolutionary reinforcement learning based control. *IET Gener. Transm. Distrib.* 16, 3825– 3836 (2022).
3. R. Huang, Y. Chen, T. Yin, Q. Huang, J. Tan, W. Yu, X. Li, A. Li, Y. Du, "Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning," in IEEE Transactions on Power Systems, vol. 37, no. 6, pp. 4168-4178, Nov. 2022
4. D. Yan, Q. Huang, R. Huang, T. Yin, J. Tan, W. Yu, X. Li. “Physics-informed Evolutionary Strategy based Control for Mitigating Delayed Voltage Recovery”, IEEE Trans. on Power Systems, vol. 37, no. 5, pp. 3516-3527, Sept. 2022
5. R. Huang, Y. Chen, T. Yin, X. Li, A. Li, J. Tan, W. Yu, Y. Liu, Q. Huang “Accelerated Derivative-free Deep Reinforcement Learning Based Load Shedding for Emergency Voltage Control,” in IEEE Transactions on Power Systems, vol. 37, no. 1, pp. 14-25, Jan. 2022
6. Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, Z. Huang. “Adaptive Power System Emergency Control Using Deep Reinforcement Learning,” *IEEE Transactions on Smart Grid,* vol. 11, no. 2, pp. 1171-1182, March 2020
7. X. Sun, X. Li, S. Datta, X. Ke, Q. Huang, R. Huang, Z. Jason Hou, “Smart Sampling for Reduced and Representative Power System Scenario Selection,” *IEEE Open Access Journal of Power and Energy*, vol. 8, pp. 293-302, 2021, doi: 10.1109/OAJPE.2021.3093278.
8. R. Hossain, Q. Huang, R. Huang, “Graph Convolutional Network-Based Topology Embedded Deep Reinforcement Learning for Voltage Stability Control”, *in IEEE Transactions on Power Systems,*

Questions?

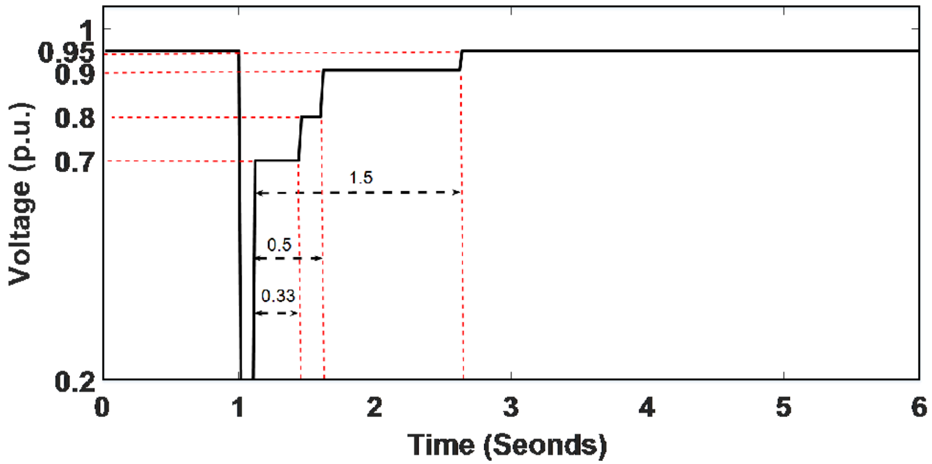
Qiuhua Huang

qiuhuahuang@mines.edu

BACKUP SLIDES

Reward Function Design: An Example

- RL agents try to maximize the accumulative rewards
  - Opposite to the objective function of optimal control
  - In general, no convexity requirement → more flexibility in design
- The reward function is usually application-specific



Bus voltage performance requirement

Voltage criteria

Load shedding

Invalid actions

$$Reward = \begin{cases} c_1 \sum_i \Delta V_i - c_2 \sum_j \Delta P_j(p.u.) - c_3 u_{invalid} \\ -10000, \text{ if } V_i(t) < 0.95, \quad t > T_{pf} + 4 \end{cases}$$

Large-penalty for non-acceptable performance

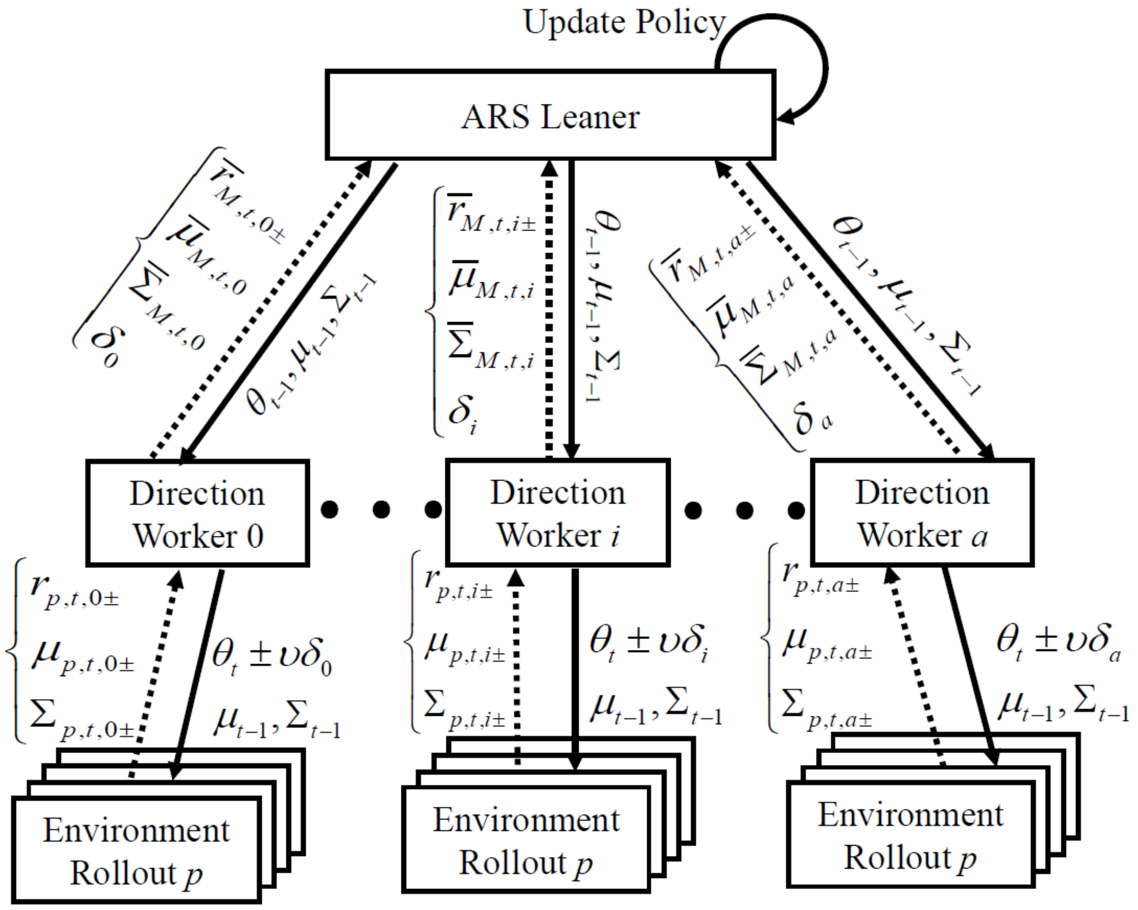
$$\Delta V_i(t) = \begin{cases} \min\{V_i(t) - 0.7, 0\}, \text{ if } T_{post\_fault} < t < T_{post\_fault} + 0.33 \\ \min\{V_i(t) - 0.8, 0\}, \text{ if } T_{post\_fault} + 0.33 < t < T_{post\_fault} + 0.5 \\ \min\{V_i(t) - 0.9, 0\}, \text{ if } T_{post\_fault} + 0.5 < t < T_{post\_fault} + 1.5 \\ \min\{V_i(t) - 0.95, 0\}, \text{ if } T_{post\_fault} + 1.5 < t \end{cases}$$

Meet the minimum performance requirement

Reward Function

R. Huang, Y. Chen, T. Yin, Q. Huang, J. Tan, W. Yu, X. Li, A. Li, Y. Du, “Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning,” IEEE Transactions on Power Systems, 2022

Two-level Parallelism of PARS

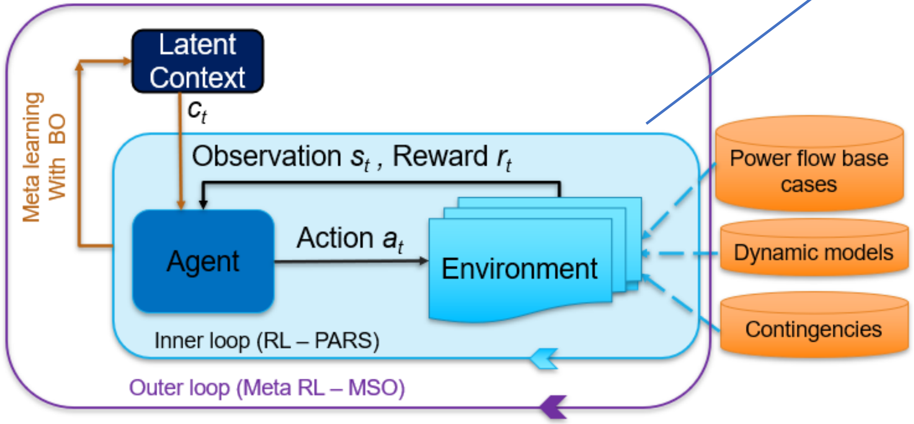


RLGC/HADREC RLGC/HADREC RLGC/HADREC

Parallel Augmented Random Search (ARS) Algorithm

- Highly scalable when combined with high performance computing and grid simulator [1] (see results in the next slide)
- Guided search by exploiting estimated gradients [2]

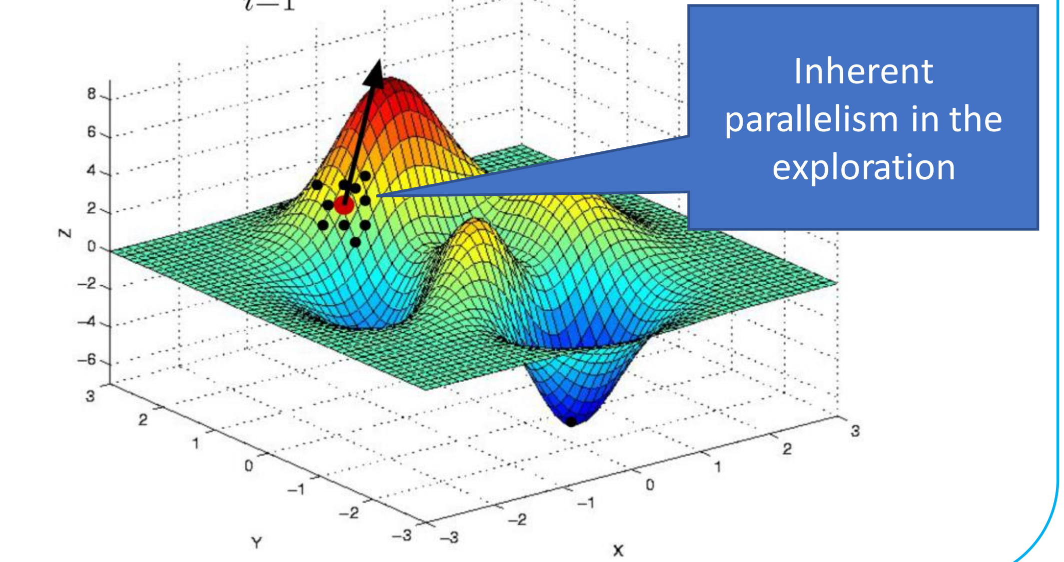
$$\varepsilon_i = \alpha\varepsilon + \boxed{(1-\alpha)U\varepsilon'},\ \ \varepsilon \sim N(0,\sigma I_n)\ and\ \boxed{\varepsilon' \sim N(0,\sigma I_k)}$$



Parallel ARS Algorithm

Basic idea: Estimate the gradient using random search

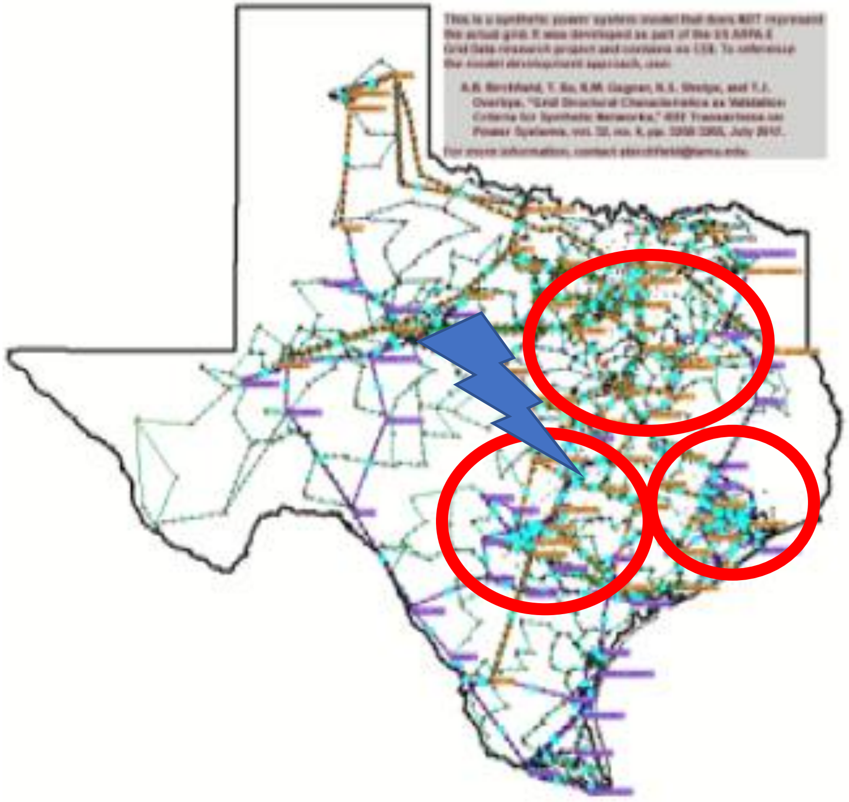
$$\nabla F(\theta) \approx \frac{1}{n\sigma}\sum_{i=1}^{n}\{F(\theta+\sigma\epsilon)\epsilon\} \quad \text{where} \quad \epsilon \sim N(0,I)$$



[1] R. Huang, Y. Chen, T. Yin, **Q. Huang**, J. Tan, W. Yu, X. Li, A. Li, Y. Du, "Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning," IEEE Transactions on Power Systems, accepted, 2022

[2] Y. Du, **Q. Huang**, R. Huang, T. Yin, J. Tan; W.Yu; X. Li, "Physics-informed Evolutionary Strategy based Control for Mitigating Delayed Voltage Recovery," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2021.3132328.

Physics-inspired training



Physics: the voltage stability problem in power systems are mostly local issues

1. Areas are loosely coupled for voltage problems

2. Yet, actions in two or three of the regions are required for faults near or at the boundary of the regions.

3. Solutions: divided training and then coordinative training