Data-Driven Reinforcement Learning Control of Large CPS Networks using Multi-Stage Hierarchical Decompositions

Aranya Chakrabortty

Electrical & Computer Engineering, North Carolina State University

CPS-1931932

Project Goal

To design fast, resilient and cost-optimal controllers for extreme-dimensional CPS networks (such as electric power systems) in a model-free way using massive volumes of real-time streaming data.

Reinforcement Learning based Optimal Control

- Recent literature (Vrabie, Vamvoudakis & Lewis, Powell & Frazier, Jiang & Jiang, Zhang & Basar, Bhasin & Dixon, Fazel & Mesbahi, Mukherjee, Bai & Chakrabortty) have shown RL to be highly successful for model-free LQR optimal control

- System: $\dot{x} = Ax + Bu$
- Cost functional: $J = \int_0^\infty (x^T Q x + u^T R u)\, dt$
- control law that minimizes the value of the cost: $u = -Kx$
  - $K = R^{-1} B^T P$
  - $A^T P + PA - PBR^{-1}B^T P + Q = 0$

RL learns K by solving the Riccati equation using only $x(t)$ and $u(t)$, no model

- Can we learn real-time Reinforcement Learning based LQR controllers for large and complex CPS networks in a tractable way?

Problem Formulation

Plant

$u_0 \rightarrow$ $\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0$

A, B unknown

$u(t)$ $\qquad$ $x(t)$

$K^{k+1} = \xi\left(K^k, x(t_k), u(t_k)\right), \quad K^0 = K_{LS}$

$u(t_k) = K^k x(t_k)$

RL Controller

Central controller learns the full-dimensional optimal controller K

Fundamental Steps of RL

1. Apply exploration noise
   $u(t) = u_0(t) \text{ for } 0 \le t \le t_\ell$
2. NN training step: Wait and collect enough measurements of $x(t)$ and $u_0(t)$ for solving a least-squares estimate of the Riccati matrix P
   $$\Phi(x,\, u_0)p = V(x, u_0)$$
   Must have least $n(n+1)/2$ rows to be full column-rank — Vectorized form of P — Vector of measurements of "value function" J
3. Generate an initial estimate K^0 from P^0
4. For $t > t_\ell$ apply $u(t) = K^k x(t)$
5. Iterate:
   $$K^{k+1} = \xi(K^k,\, x(t_k),\, u(t_k))$$
   ξ : Nonlinear RL function

Main challenge is Curse of Dimensionality!

A simple 10-dimensional K can easily take 1 hour of learning time in a standard PC!

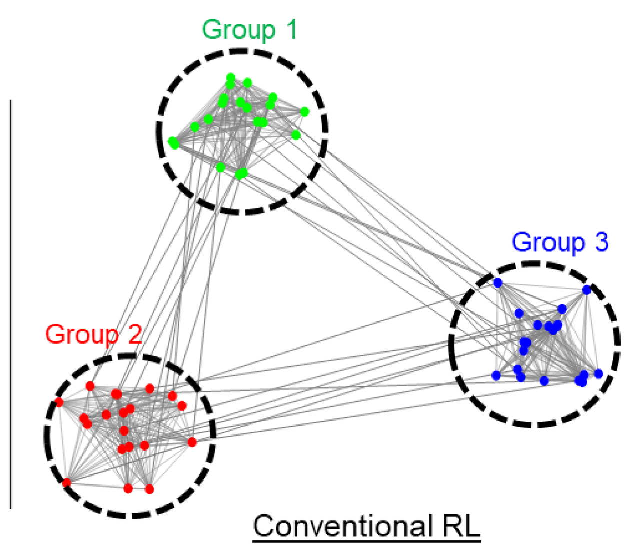
1. Define projection matrices P that bring out a possible decomposition of local and global control objectives through low-rank controllability subspaces.

2. Design artificial neural networks (NN) that can quickly learn and predict P from data.

3. Design online *hierarchical* Reinforcement Learning (RL)-based optimal controllers that independently learn the local and global controls.

This controller will be significantly more scalable and faster than centralized RL due to its hierarchical structure involving smaller-dimensional state feedback, and parallelization of the learning loops.

Proposed Approach



Three Primary Questions

- Q1: How to transition from model-based to model-free (data-driven) control?
- Q2: How to take decisions in real-time based on data only?
- Q3: Make control highly scalable with network size

Conventional RL

Con: Learns $u(t) = Kx(t)$ for the whole network at one shot

- Large learning time
- Dense K requires ${}^nC_2$ comm. links

Pro: $u(t)$ is optimal

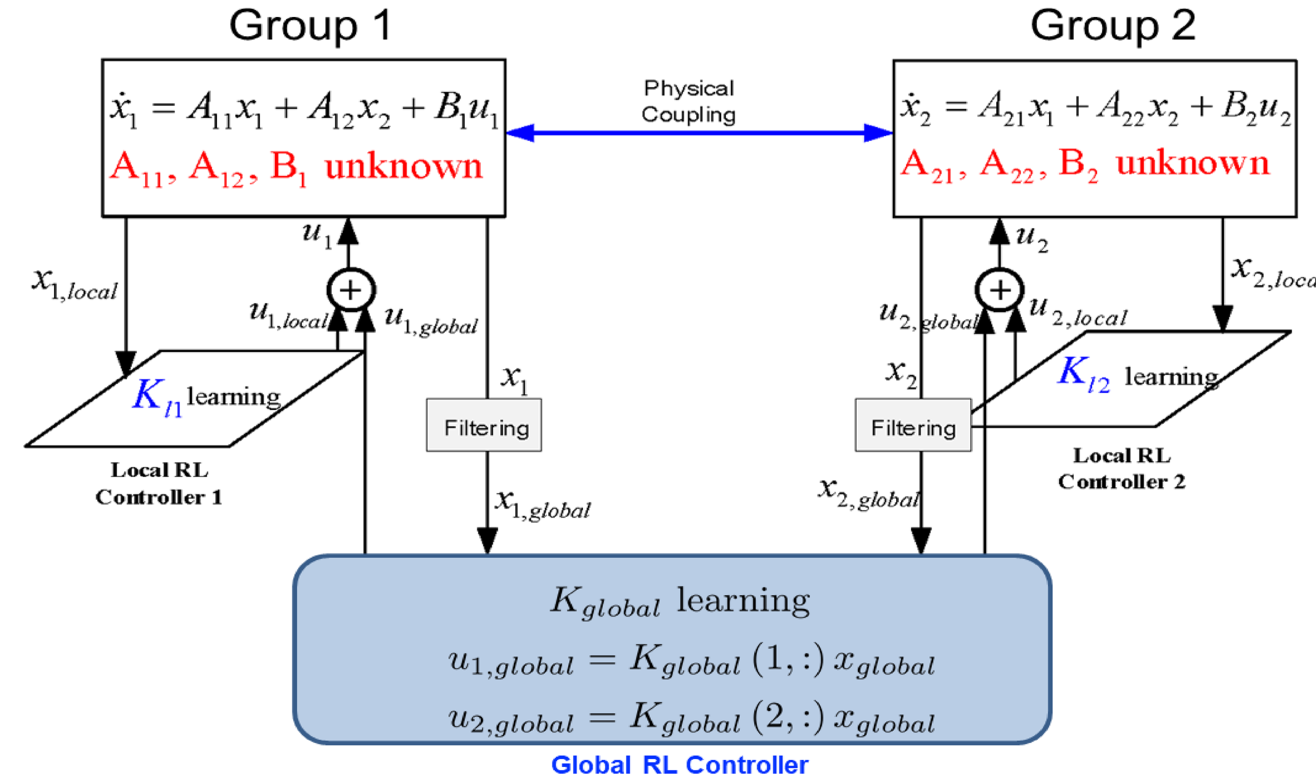
Hierarchical RL

Pro: Learns $u(t)$ in multiple stages (in parallel):

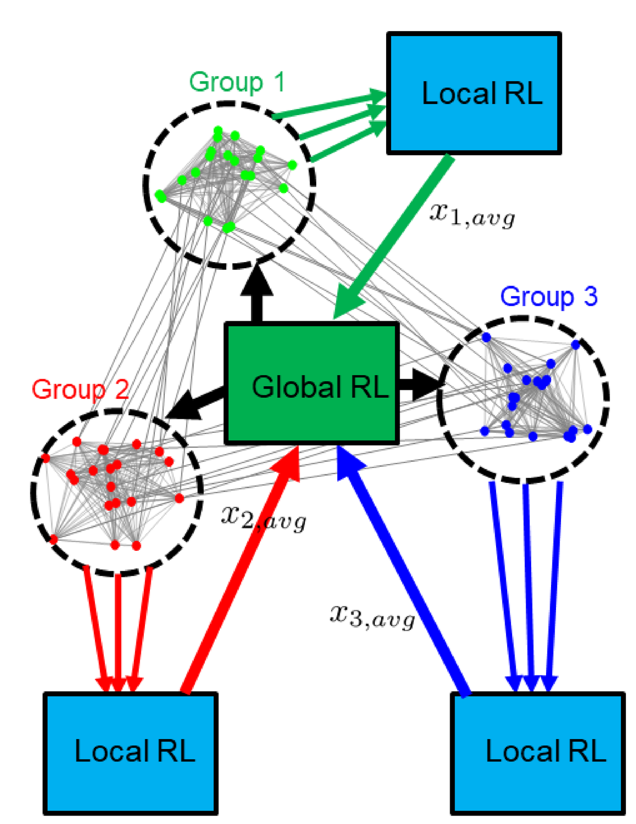
$$u(t) = u_{local}(t) + u_{global}(t)$$

u_{local} is group-wise decentralized, and u_{global} is an aggregate input in reduced dimension

Con: $u(t)$ is sub-optimal



Hierarchical Reinforcement Learning



Proposed Hierarchical RL

Benefits:

1. Local RL & Global RL can go on in parallel – saves learning time
2. Data privacy is preserved between groups
3. Reduces inter-group communication

Drawbacks:

1. Conventional RL produces optimal controller:
$$K^* = K_{optimal}$$

The proposed hierarchical RL will be super-fast but will produce only a sub- optimal controller:

$$\begin{aligned} u'(t) &= u_{local}(t) + u_{global}(t) \\ &= K_{local}x_l + K_{global}x_g \end{aligned}$$

Smaller r will improve global learning time, but degrade closed-loop performance

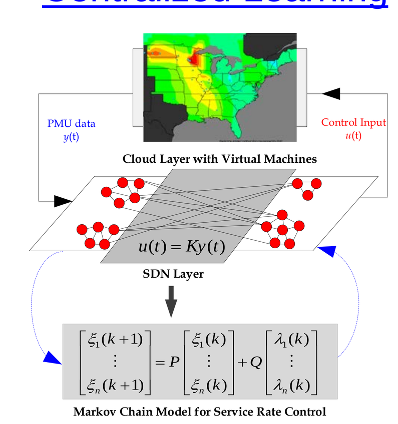
How to group agents:

1. Time-scale separation
   - P is averaging - several papers on identifying P (coherent groups) from PMU data, redundancy in observability

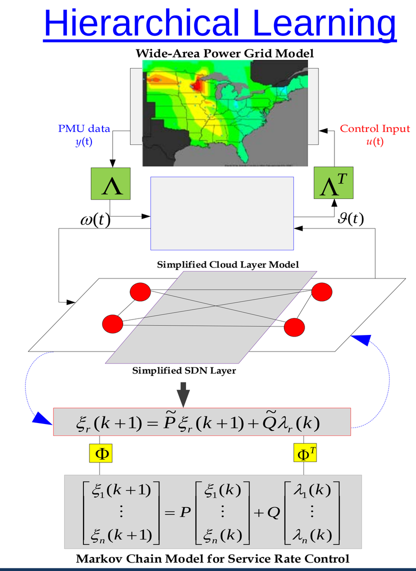
2. Low-rank controllability
   - P follows from SVD of empirical controllability gramian
   - Only gens with most influence on inter-area mode are identified

Application in Power System Controls

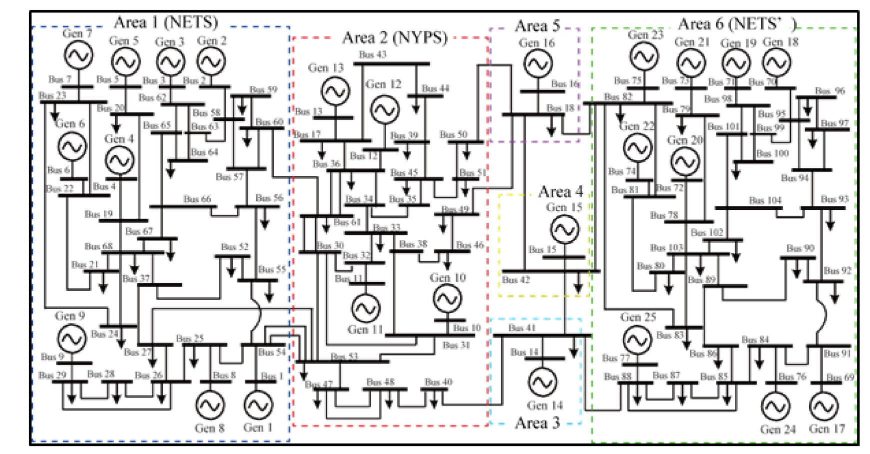
Centralized Learning



Hierarchical Learning



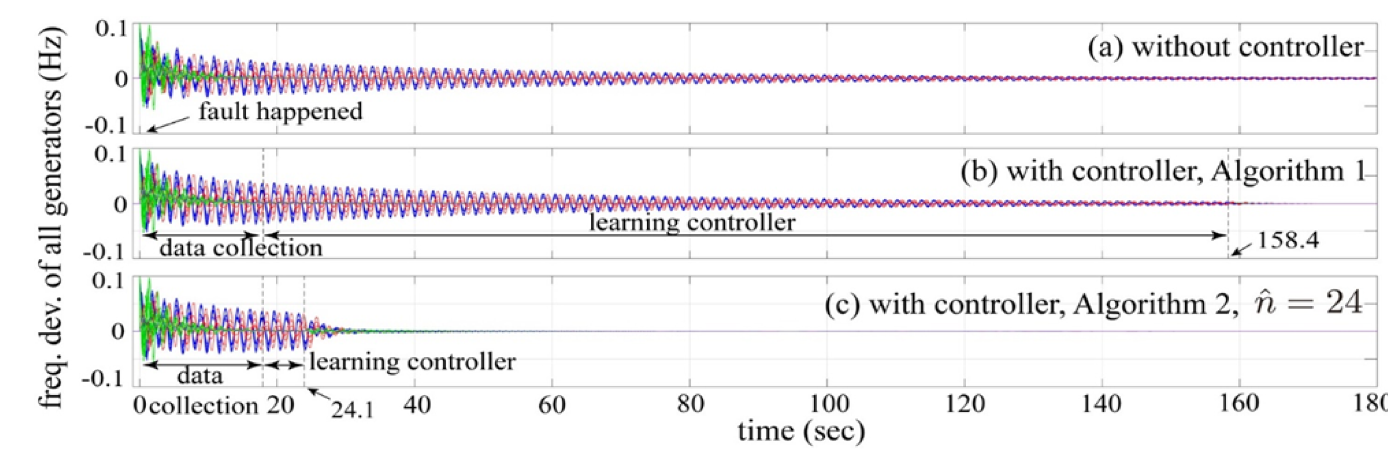
NETS+NYPA Power Grid Model



25 generators, 52 loads, 104 buses

6 groups (not coherent)

Control via AVR/PSS

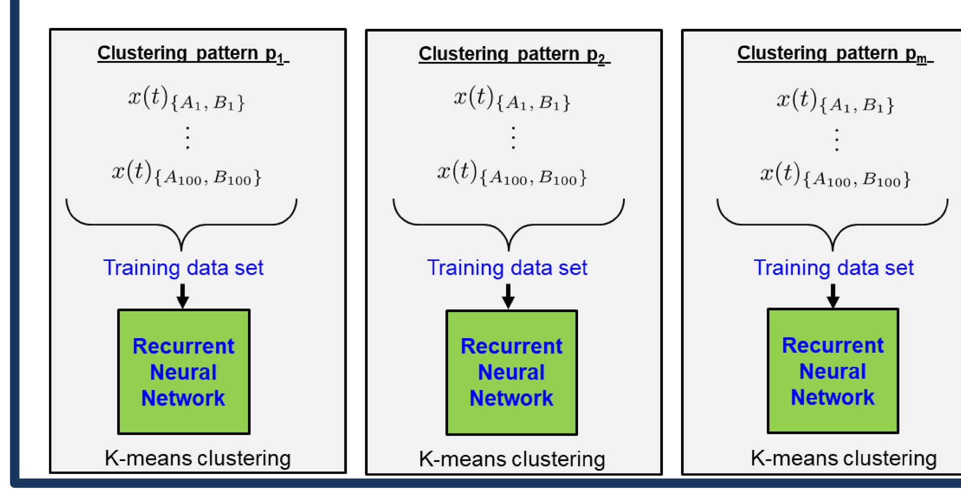


Sparsity-Constrained Learning

Neural networks will be trained with experimental data to identify clusters in real-time, as well as to identify desired sparse structures of local RL controllers

Train neural networks with prior data, predict groups from online data

Unsupervised methods - Kernel-based K-means, Graph product decompositions, Tensor decomposition theory (excellent survey paper by Sidiropoulos)



Theory: Impose sparse structure on local controllers to minimize communication:

- The local learning law
  $$K_{Li}^{k+1} = \xi_i(K_{Li}^k,\, x_i(t_k),\, u_i(t_k))$$

will, in general, produce a dense K_{Li} which means all-to-all communication

- To avoid this, we will augment the NN training step of RL with an additional sparsity constraint

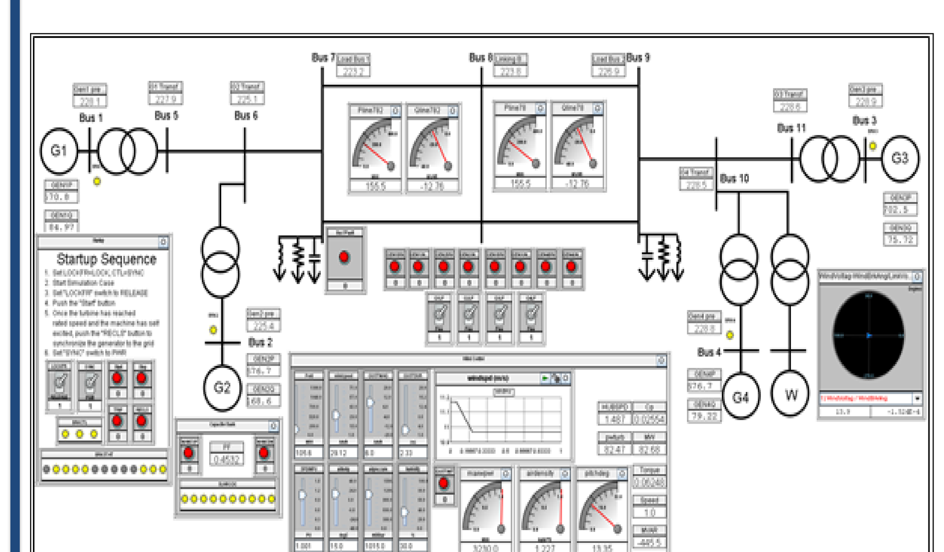
Conventional LS: $\Phi(x,\, u_0)p = V(x, u_0)$

Modified LS: $\min \|p\|_{\ell_1}$ s.t $\|\Phi(x,\, u_0)p - V(x, u_0)\|_2 \le \epsilon$

Imposes sparse structure on P and K

Must have least $n(n+1)/2$ rows to be full column-rank — Vectorized form of P — Vector of measurements of "value function" J

Experimental Testbed



1. Multi-vendor PMU-based hardware-in-loop simulation testbed at NCSU will be used for verification and validation
2. We will showcase resiliency of hierarchical RL based control for wide-area power oscillation damping
3. Sensitivity of RL control to CPS uncertainties such as delays, quantization, saturation and information loss will be tested

Broader Impacts

- Undergraduate, K-12 and minority education in power systems via Science House and FREEDM ERC programs at NC State
- New graduate curriculum on Machine Learning and Data Science
- International collaborations with Tokyo Institute of Technology in Japan
- Industry collaborations with power utilities (Duke Energy) and software vendors (SAS, ABB)