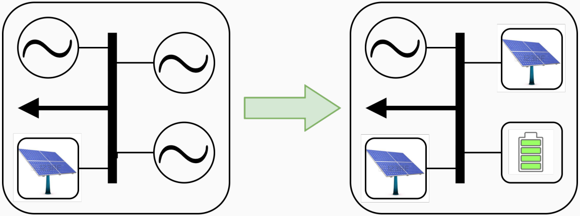
# A Set-Theoretic Approach to Safe Reinforcement Learning in Power Systems

*Daniel Tabas and Baosen Zhang*
*Department of Electrical and Computer Engineering*
*University of Washington, Seattle*

ICML 2021 Workshop: Tackling Climate Change with Machine Learning
July 23, 2021

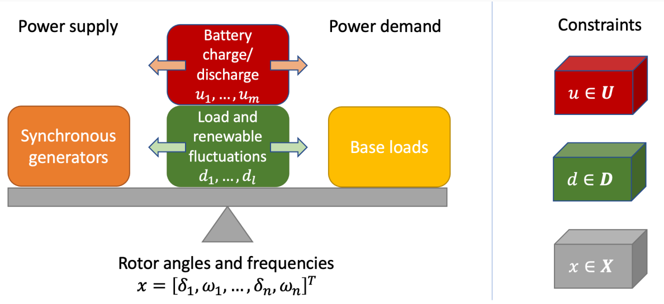
# Introduction



Power systems are transitioning from synchronous generator-based to inverter-based.

- More flexibility, less inherent stability
- Enables and necessitates new control techniques
- Increasing complexity → difficult to find good policy
- New control policies must be **safe**

# Model



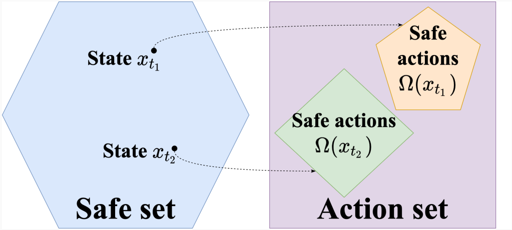

Power system stability requires balancing power supply with power demand.

- Loads and renewables fluctuate, but energy storage can be used to balance the fluctuations
- System operates under constraints: capacity limits $(U, D)$ and safety constraints $(X)$

# Previous work

- Lyapunov stability [1, 3, 7] and robust control guarantees [4, 5]
  - Guarantee stability but not hard constraint satisfaction
- Optimization-based safety filters [2, 8]
  - Calling an optimization solver in real time may not be practical
- Geometric approaches [9]
  - Can lead the system into states that are safe but have no safe action

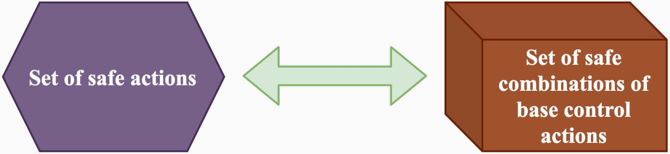
# Policy network architecture




The current set of safe actions is a moving target for a policy network.

- Changing geometry
- Difficult to parameterize
- Difficult to enforce as action constraints without solving projection (LP) at best

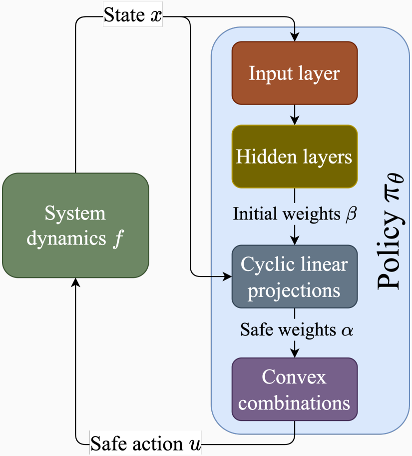
# Output feature map




The set of **safe combinations** of base control actions has simpler geometry than the set of safe actions. We can reach the set of safe combos through a small number of iterative linear projections.

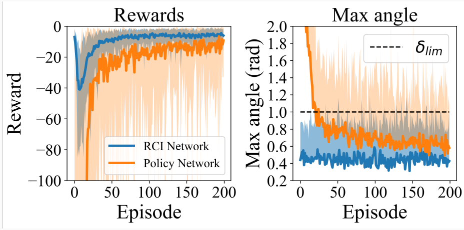
- **Base control actions:** safe actions associated with each vertex of the safe set
- **Safe combos:** any convex weights that generate the current state when applied to the vertices of the safe set

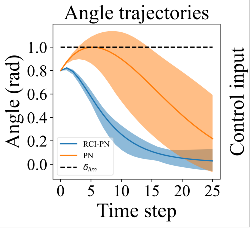
# Policy network architecture




- Policy parameterized by a neural network and trained using RL (DDPG algorithm [6])

# Simulations




Compared performance of our method (blue) to a policy network trained with a soft penalty on constraint violations (orange). Our proposed method demonstrated:

- Better rewards
- Far fewer constraint violations throughout training
- Better constraint satisfaction during testing

# Conclusions and future work

Conclusions:

- Proposed computationally efficient safe RL paradigm for power systems

Future work:

- Devise an output layer that has a closed-form solution
- Investigate robustness of learned policies to topology changes

# References


[1] Felix Berkenkamp, Matteo Turchetta, Angela P Schoellig, and Andreas Krause. Safe model-based reinforcement learning with stability guarantees. In *Advances in Neural Information Processing Systems*, volume 2017-Decem, pages 909–919, 2017.

[2] Richard Cheng, Gábor Orosz, Richard M. Murray, and Joel W. Burdick. End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks. In *33rd AAAI Conference on Artificial Intelligence*, pages 3387–3395, 2019.

[3] Wenqi Cui and Baosen Zhang. Reinforcement Learning for Optimal Frequency Control: A Lyapunov Approach. *arXiv preprint: 2009.05654v3*, 2021.

[4] Priya L. Donti, Melrose Roderick, Mahyar Fazlyab, and J. Zico Kolter. Enforcing robust control guarantees within neural network policies. In *International Conference on Learning Representations*, pages 1–26, 2021.

[5] R. Matthew Kretchmar, Peter M. Young, Charles W. Anderson, Douglas C. Hittle, Michael L. Anderson, and Christopher C. Delnero. Robust reinforcement learning control with static and dynamic stability. *International Journal of Robust and Nonlinear Control*, 11(15):1469–1500, 2001.

[6] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 2016.

[7] Theodore J Perkins and Andrew G Barto. Lyapunov design for safe reinforcement learning control. *Journal of Machine Learning Research*, 3(4-5):803–832, 2002.

[8] Kim Peter Wabersich and Melanie N. Zeilinger. A predictive safety filter for learning-based control of constrained nonlinear dynamical systems. *Automatica*, 129:109597, 2021.

[9] Liyuan Zheng, Yuanyuan Shi, Lillian J. Ratliff, and Baosen Zhang. Safe reinforcement learning of control-affine systems with vertex networks. *arXiv preprint: arXiv 2003.09488v1*, 2020.