Enforcing Policy Feasibility Constraints through Differentiable Projection for Energy Optimization

Bingqing Chen, Priya L. Donti, Kyri Baker, J. Zico Kolter, and Mario Bergés

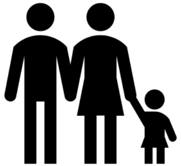
Best Paper Runner-up at ACM e-energy'21

Carnegie Mellon University

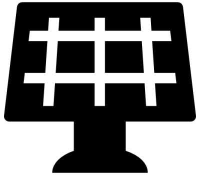
Autonomous energy systems should satisfy constraints

Application 1: HVAC control in buildings to minimize energy consumption, without violating occupants' thermal comfort constraints.

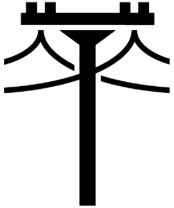
subject to



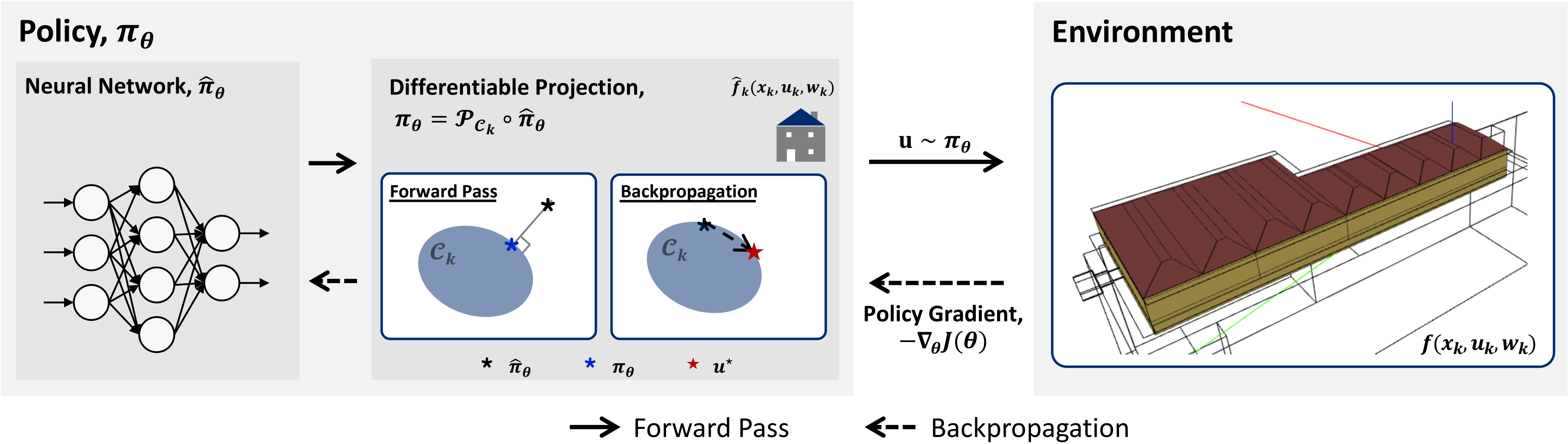
Application 2: Control of inverters to minimize renewable energy curtailment, without violating constraints of the distribution network.



subject to



PROF: Projected Feasibility

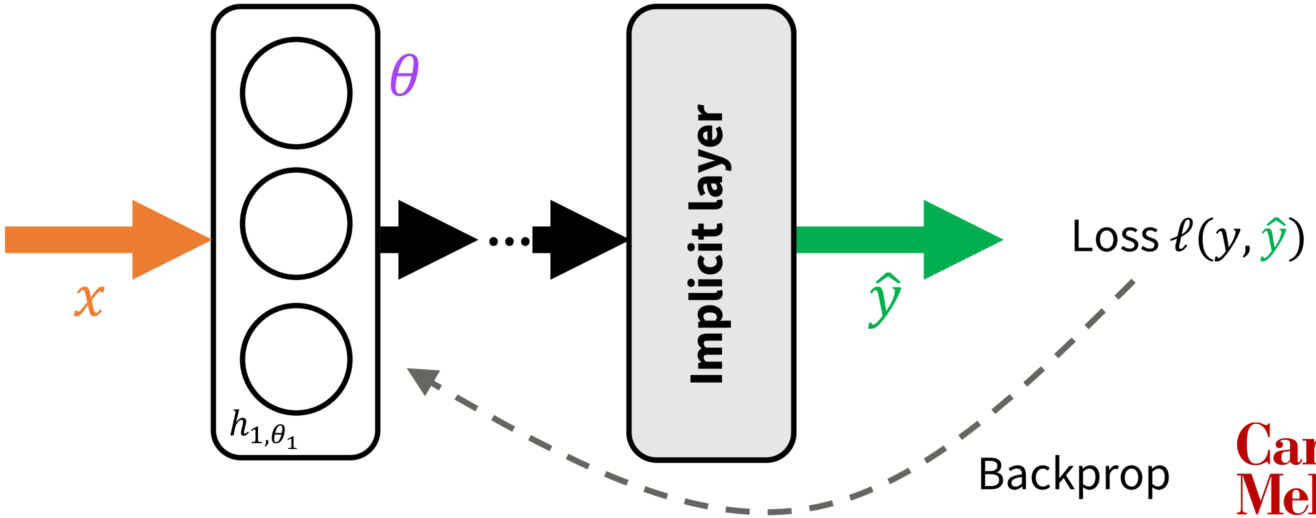


Background: Neural Networks and Differentiable Learning

Neural network = composition of non-linear, parameterized, differentiable functions

$$h_\theta = h_{1,\theta_1} \circ \cdots \circ h_{L,\theta_L}$$

Recent interest in enriching the set of functions that can be accommodated ("implicit layers"), such as **optimization problems** and **physical equations**.



Example: Projection Layer

Consider the projection operation

$$\mathcal{P}_{\mathcal{C}}(\hat{u}) = \underset{u \in \mathcal{C}}{\operatorname{argmin}} \frac{1}{2} ||u - \hat{u}||_2^2$$

For linear constraints $\mathcal{C} = \{u : Au = b, Gu \leq h\}$, can write and differentiate through KKT conditions

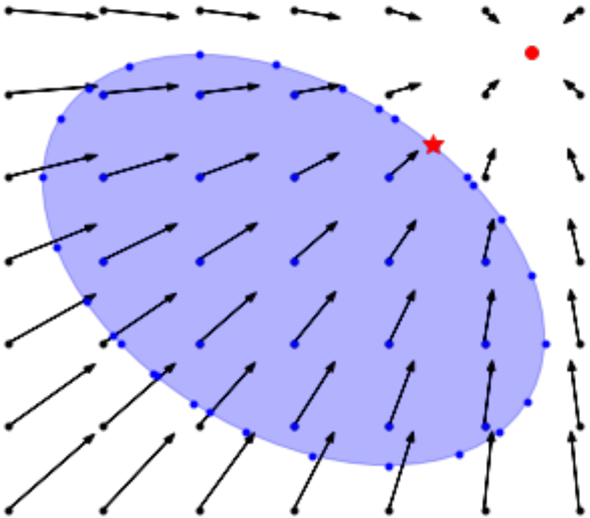
$$\begin{aligned} \operatorname{diag}(\lambda^\star)(Gu^\star - h) &= 0 \\ Au^\star - b &= 0 \\ u^\star - \hat{u} + A^T\nu^\star + G^T\lambda^\star &= 0 \end{aligned} \quad \longrightarrow \quad \begin{aligned} \mathrm{d}u^\star - \mathrm{d}\hat{u} + \mathrm{d}A^T\nu^\star + A^T\mathrm{d}\nu^\star + \mathrm{d}G^T\lambda^\star + G^T\mathrm{d}\lambda^\star &= 0 \\ \mathrm{d}Au^\star + A\mathrm{d}u^\star - \mathrm{d}b &= 0 \\ \operatorname{diag}(Gu^\star - h)\mathrm{d}\lambda + \operatorname{diag}(\lambda^\star)(\mathrm{d}Gu^\star + G\mathrm{d}u^\star - \mathrm{d}h) &= 0 \end{aligned}$$

Note: Can also differentiate through general convex projections

See also:

- Amos, B., Kolter, J.Z. (2017). OptNet: Differentiable Optimization as a Layer in Neural Networks. *ICML*.
- Agrawal, A., Amos, B., Barratt, S., Boyd, S., Diamond, S., & Kolter, J.Z. (2019). Differentiable convex optimization layers. *NeurIPS*.

"Post-hoc" projections enforce constraints, but in a way that is hidden from the neural network during learning.



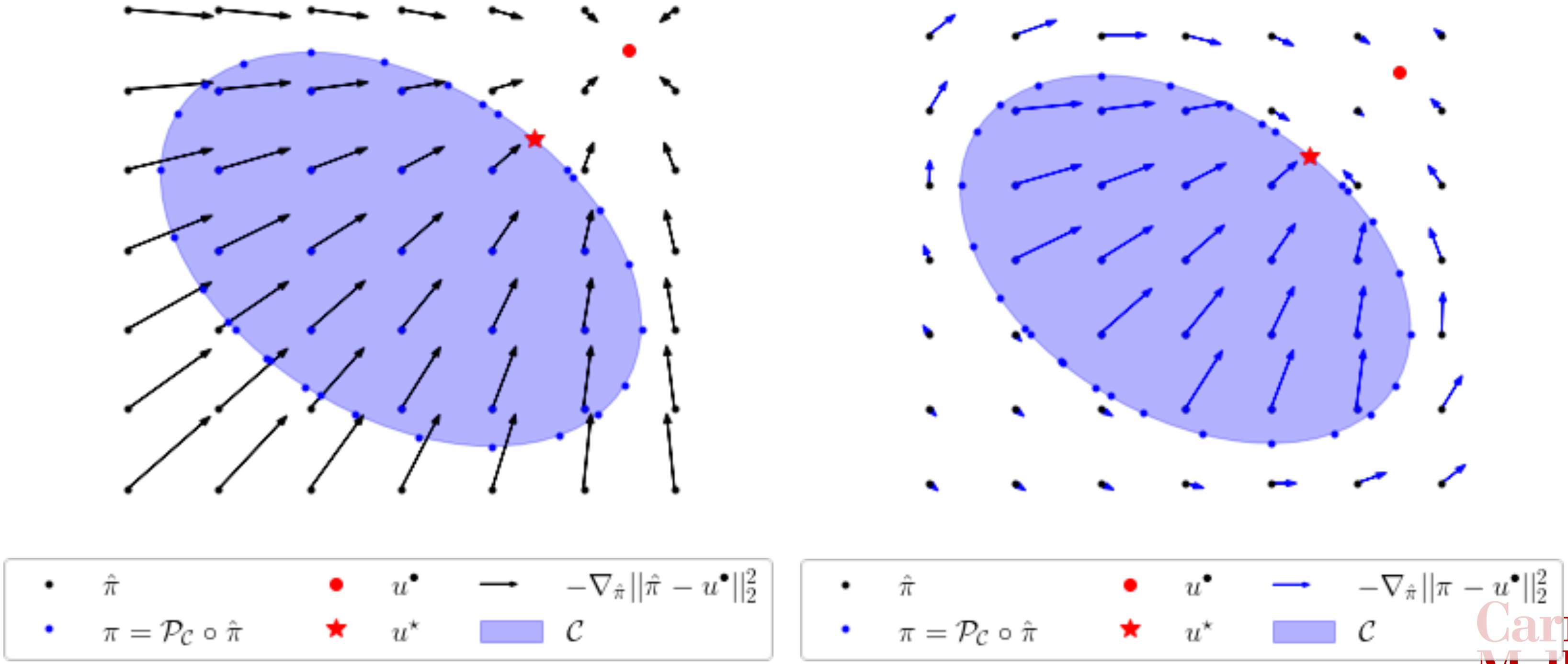
$\hat{\pi}$ · $u^{\bullet}$ · $-\nabla_{\hat{\pi}} \|\hat{\pi} - u^{\bullet}\|_2^2$

$\pi = \mathcal{P}_{\mathcal{C}} \circ \hat{\pi}$ · $u^{\star}$ · $\mathcal{C}$

$\hat{\pi}$ · $u^{\bullet}$ · $-\nabla_{\hat{\pi}} \|\pi - u^{\bullet}\|_2^2$

$\pi = \mathcal{P}_{\mathcal{C}} \circ \hat{\pi}$ · $u^{\star}$ · $\mathcal{C}$

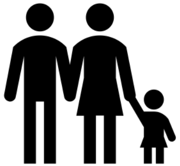
Backpropagating through the differentiable projection layer makes the neural network cognizant of the constraints in its learning.



Carnegie Mellon University

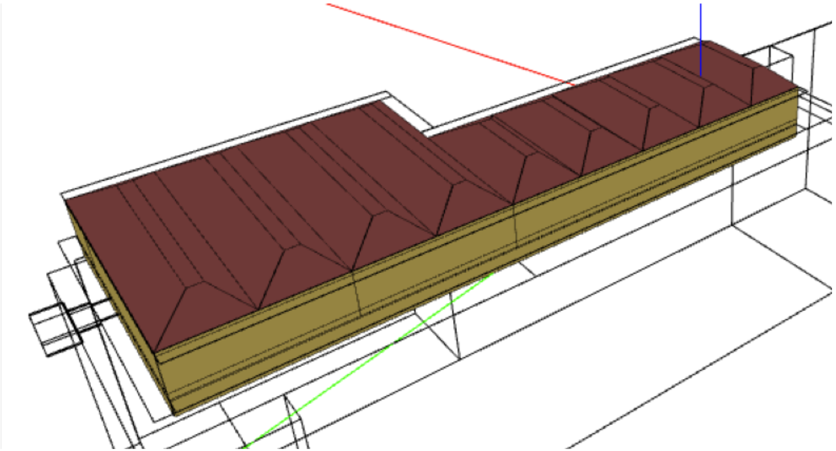
Experiment 1: Energy-efficient Building Operation

subject to

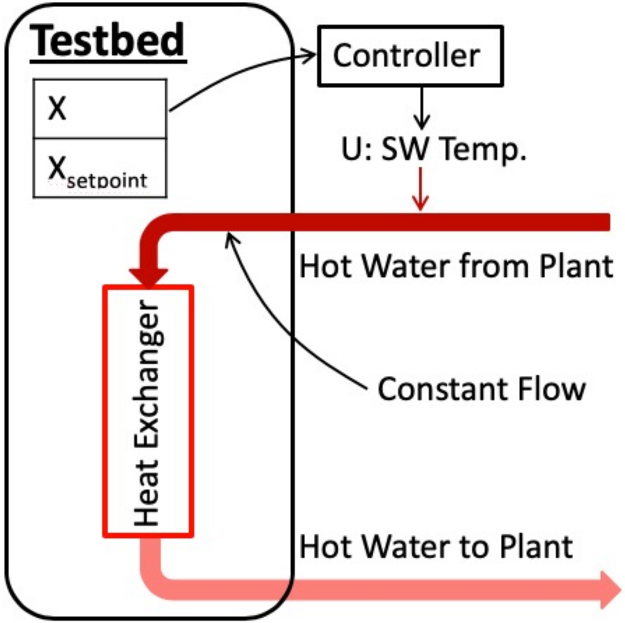


Carnegie Mellon University

Simulation Testbed



Intelligent Workplace
Margaret Morrison Hall, 4^{th} Floor
(✤ Zhang & Lam, 2018)

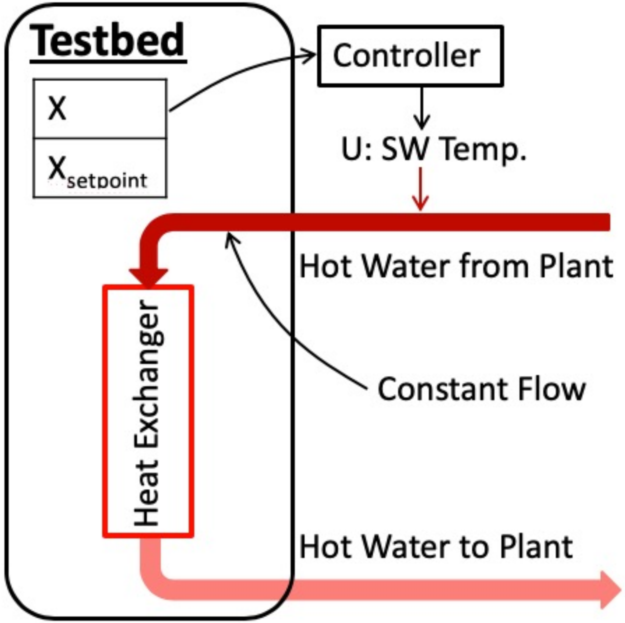


HVAC Schematic

Carnegie Mellon University

Simulation Testbed

Hot Water Pipes



HVAC Schematic

Carnegie Mellon University

Constructing the safe set with an approximate model

Safe Set

The set of all admissible actions that satisfy the end-use requirements and physical constrains.

$$\mathcal{C} = \left\{ [u_{t:t+T-1}] \;\middle|\; \begin{array}{l} x_{k+1} = \mathcal{T}(x_k, u_k); \\ \underline{u}_k \leq u_k \leq \overline{u}_k; \quad \forall k \in \{t, \ldots, t+T-1\} \\ \underline{x}_{k+1} \leq x_{k+1} \leq \overline{x}_{k+1}; \end{array} \right\}$$

Thermodynamics
Allowable Control
Thermal Comfort

Thermodynamics

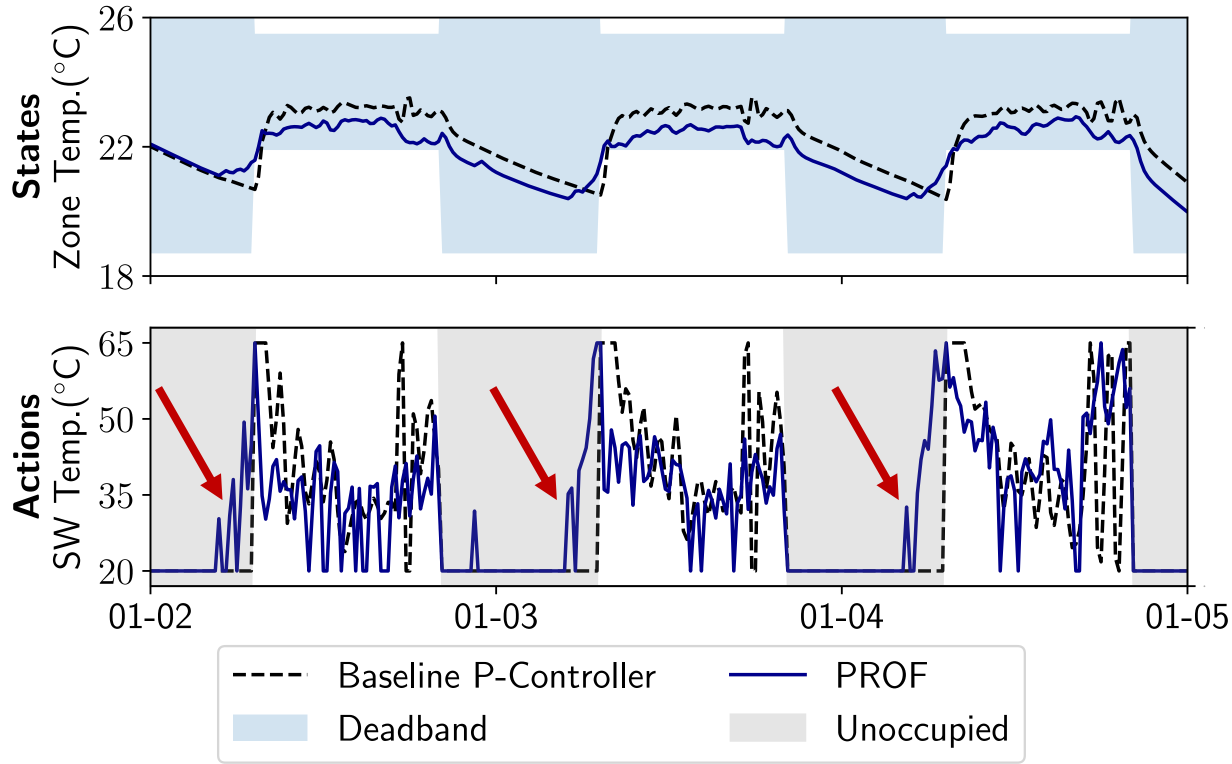
The building dynamics can be approximated as a linear system.

$$\underbrace{\begin{bmatrix} 1 & & & \\ -a & 1 & & \\ & \ddots & \ddots & \\ & & -a & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} T_{t+1} \\ T_{t+2} \\ \vdots \\ T_{t+T} \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} aT_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{x}_0} + \mathbf{B_u} \boxed{\underbrace{\begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+T-1} \end{bmatrix}}_{\mathbf{u}}} + \underbrace{\begin{bmatrix} T_{a,t} \\ T_{a,t+1} \\ \vdots \\ T_{a,t+T-1} \end{bmatrix}}_{\mathbf{D}} \underbrace{[1-a]}_{\mathbf{b_d}}$$

We consider the control actions, $\mathbf{u} \in \mathbb{R}^T$, over a 3-hour planning horizon.

Carnegie Mellon University

At the beginning of the learning phase,

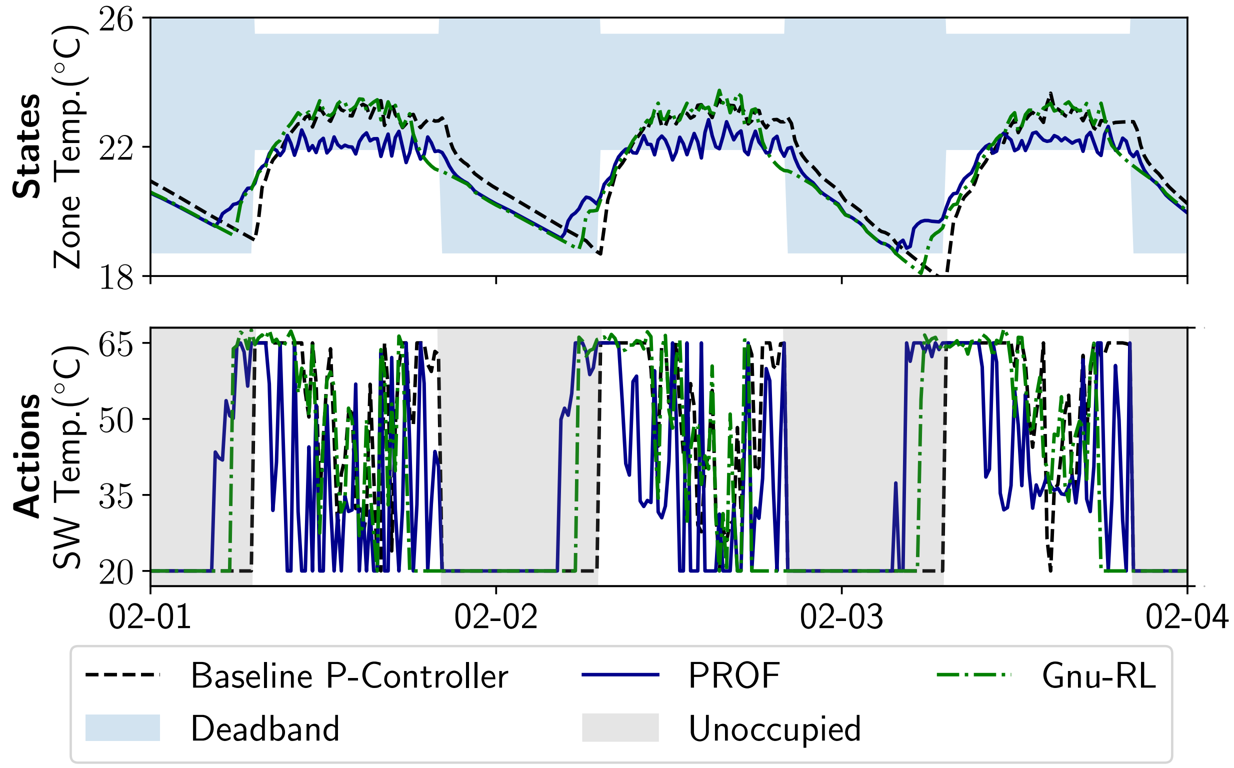


We pretrain PROF by imitating Baseline P-Controller

The differentiable projection layer enforces PROF to preheat the environment, even though this behaviour is not present in the expert demonstration.

Carnegie Mellon University

After learning for a month,



PROF learns an energy-saving behaviour by maintaining temperature just above the lower limit of allowable temperature.

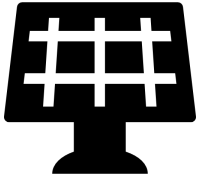
In comparison, Gnu-RL maintains temperature at the setpoint, which consumes more energy.

Carnegie Mellon University

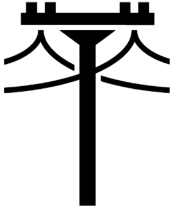
PROF improves energy-efficiency by 23.8% compared to the existing controller, while maintaining a comparable level of thermal comfort.

	Total Heating Demand (kWh)	**Predicted Percentage Dissatisfied**	
		Mean (%)	**STD (%)**
Existing Controller	43709	**9.45**	5.59
Agent #6 (Zhang & Lam, 2018)	37131	11.71	3.76
Gnu-RL (Chen et al., 2019)	34678	9.56	6.39
PROF	**33271**	9.68	**3.66**

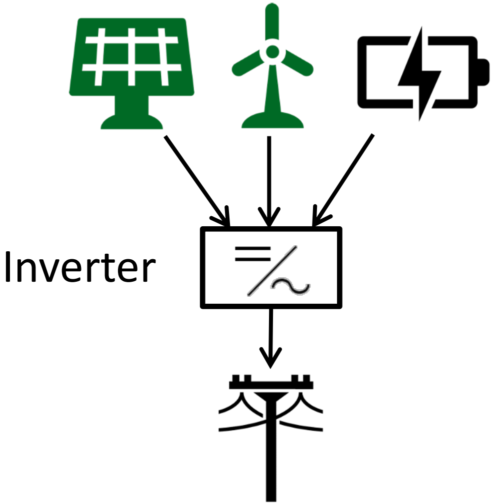
Experiment 2: Inverter Control

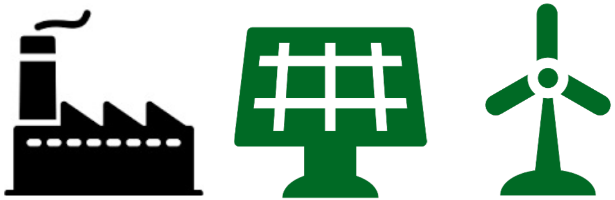


subject to



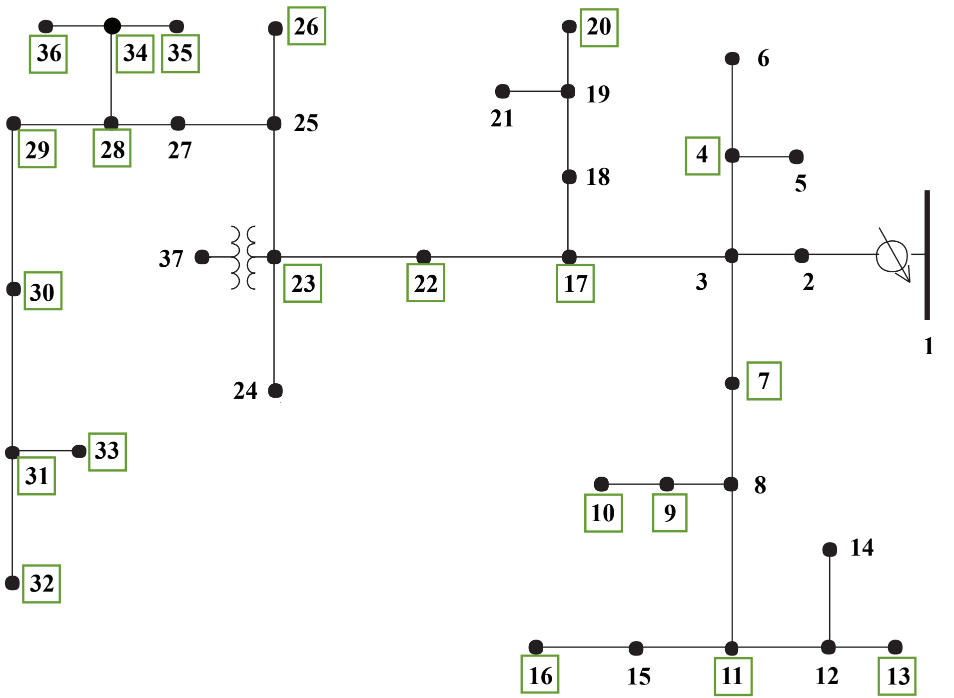
As the penetration of renewable energy resources continues to grow, the grid of the future will become inverter-dominated.





Kroposki, B., Johnson, B., Zhang, Y., Gevorgian, V., Denholm, P., Hodge, B. M., & Hannegan, B. (2017). Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy. *IEEE Power and Energy Magazine*, 15(2), 61-73.

Simulation Testbed: IEEE 37-bus Feeder System

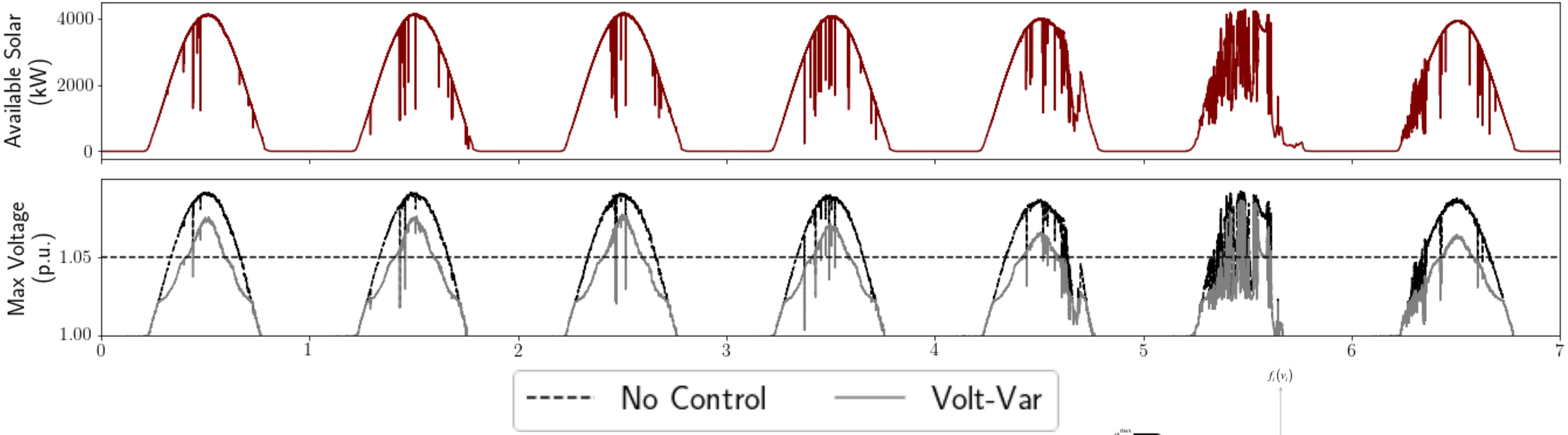


PV Panels

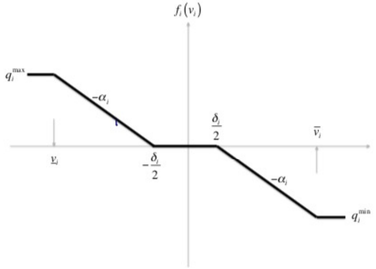
Baker, K., Bernstein, A., Dall'Anese, E., & Zhao, C. (2017). Network-cognizant voltage droop control for distribution grids. *IEEE Transactions on Power Systems*, *33*(2), 2098-2108.

Carnegie Mellon University

Over-voltage has become a common occurrence in areas with high renewables penetration. Volt-Var control still incurs violations 22% of the time.



Volt-Var control is recommended in IEEE 1547.8-2018.



Volt-Var Control Curve

Problem Formulation

Control the active and reactive power at each inverter p_i, q_i in order to minimize curtailment

$$C(\theta) = \min_{\mathbf{p},\mathbf{q}} \sum_{i=1}^{N} [p_i - p_{av,i}]_+, \quad \text{where } [\mathbf{p}, \mathbf{q}] = \pi_\theta$$

subject to system-level and device-level constraints

$$\mathcal{X} = \{v \mid 0.95 \leq v \approx Hu \leq 1.05\},$$

$$\mathcal{U}_i(k) = \{(p_i, q_i) : 0 \leq p_i \leq p_{av,i}(k), p_i^2 + q_i^2 \leq s_i^2\}$$

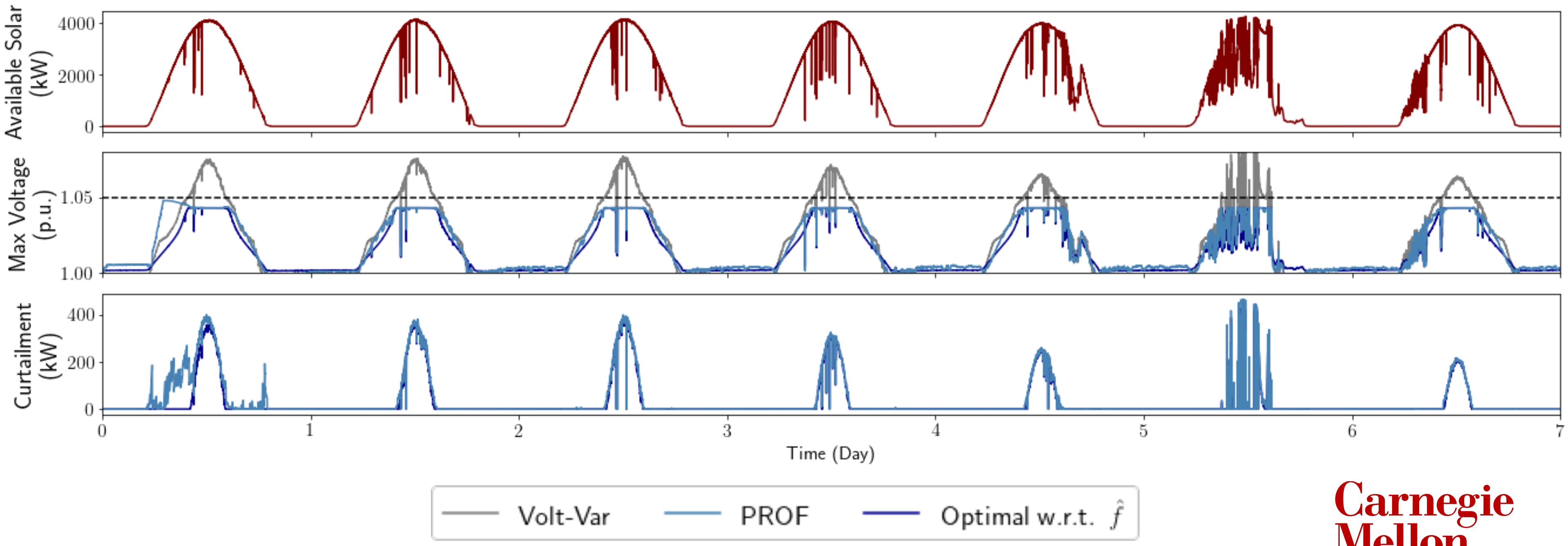
$$\mathcal{U}(k) := \mathcal{U}_1(k) \times \cdots \times \mathcal{U}_N(k).$$

given a linearized model of the distribution network.

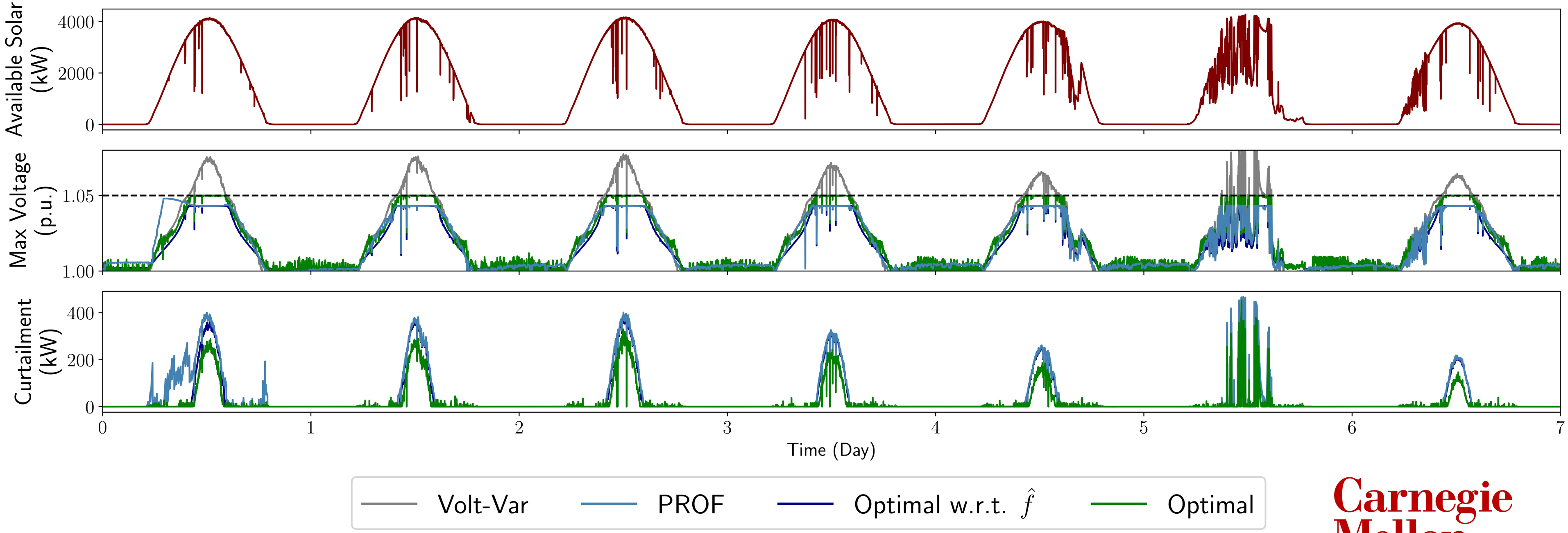
Bolognani, S., & Dörfler, F. (2015, September). Fast power system analysis via implicit linearization of the power flow manifold. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (pp. 402-409). IEEE.

Carnegie Mellon University

PROF satisfies voltage constraints throughout the experiment and learns to minimize curtailment as well as possible within its conservative safe set.



Due to the conservativeness of its safe set, PROF curtails more energy than the optimal solution, which is expensive to compute.



Carnegie Mellon University

Take-aways

PROF is a method that enforces convex operational constraints within neural policies.

The result is a powerful neural policy that can flexibly optimize performance on the true dynamics, while satisfying constraints.

In the building control case, PROF outperforms other RL agents, while mostly maintaining temperature within the deadband.

In the inverter control setting, PROF satisfies the constraints 100% of the time over more than half a million time steps.

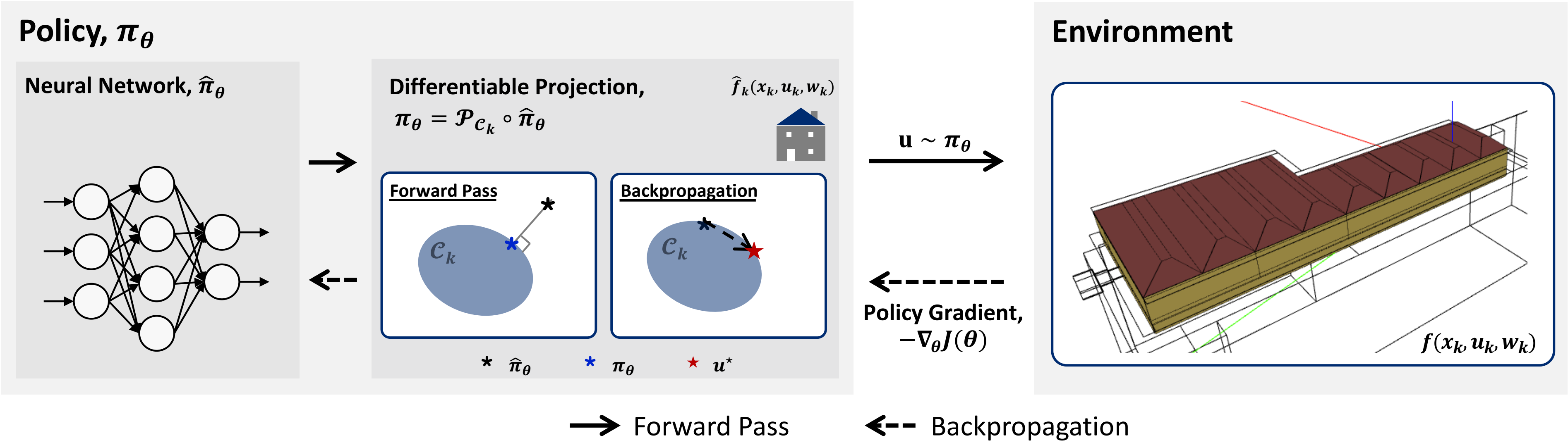
Limitations & Future Work

Assume an approximate model is available to construct the set of feasible actions

The set of feasible actions is static, i.e., not updated with new observations.

Vision: Integrate domain knowledge and physical constraints into learning-based methods.

PROF: Projected Feasibility



Thank you!

Carnegie Mellon University