Robust Energy Storage Operation via Generative Wasserstein DRO

Gen-WDRO Framework for Decision-Making Under Distribution Shift

Han Xu & Christopher Yeh

California Institute of Technology

NeurIPS 2025 Climate Change Workshop

Problem & Motivation

Climate Challenge

Energy storage is critical for renewable energy integration and emissions reduction

Battery Operation Problems

- Battery operators must decide charging schedules under **uncertain** electricity prices
- Distribution **shifts** from increasing renewables cause prediction models to fail

Need: Robust decision-making under uncertainty

Limitations of Existing Approaches

❌ Estimate-Then-Optimize (ETO)

Two-stage approach cannot directly improve downstream decision performance

❌ End-to-End Non-Robust

No robustness guarantees under distribution shift

⚠ End-to-End Robust Optimization

Worst-case optimization over entire uncertainty set is too conservative

⚠ Requires Convex Sets

Limited to convex uncertainty sets for tractability

✓ Gen-WDRO: End-to-End Distributionally Robust

Optimizes over a Wasserstein ball around learned distribution - less conservative because unlikely events receive less weight, while maintaining robustness guarantees

Problem Formulation

Contextual Information

Original Problem:

$$z^*(x) := \arg\min_z \mathrm{E}^{P}[\, f(x, y, z) \mid x \,] \text{ s.t. } g(x, z) \le \mathbf{0}$$

- Notice Conditional Distribution of $P(y|x)$ is **unknown**.
- But we can learn it which might be **inaccurate**.

Distributionally robust optimization (DRO) formulation:

$$z^*(x) := \arg\min_{z} \max_{Q \in \mathcal{B}_\rho\left(\widehat{\mathrm{P}_\theta}(\cdot|x)\right)} \mathrm{E}^{Q}[f(x, y, z)] \quad \text{s.t. } g(x, z) \le \mathbf{0}$$

- Even the learned distribution is inaccurate, the decision has robustness
- **Q1**: How to learn the distribution? **Q2**: How to solve DRO?

Problem Formulation

Future Outcome

Original Problem:

$$z^*(x) \coloneqq \arg\min_z \mathrm{E}^{P}[\,f(x, y, z) \mid x\,] \text{ s.t. } g(x, z) \le 0$$

- Notice Conditional Distribution of $P(y|x)$ is **unknown**.
- But we can learn it which might be **inaccurate**.

Distributionally robust optimization (DRO) formulation:

$$z^*(x) \coloneqq \arg\min_{z} \max_{Q \in \mathcal{B}_\rho\left(\widehat{\mathrm{P}_\theta}(\cdot|x)\right)} \mathrm{E}^{Q}[f(x, y, z)] \quad \text{s.t. } g(x, z) \le 0$$

- Even the learned distribution is inaccurate, the decision has robustness
- **Q1**: How to learn the distribution? **Q2**: How to solve DRO?

Problem Formulation

Decision variables

Original Problem:

$$\boldsymbol{z}^*(\boldsymbol{x}) \coloneqq \arg\min_{\mathrm{z}} \mathrm{E}^{\boldsymbol{P}}[\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \mid \boldsymbol{x}] \text{ s.t. } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{z}) \leq \boldsymbol{0}$$

- Notice Conditional Distribution of $P(\boldsymbol{y}|\boldsymbol{x})$ is **unknown**.
- But we can learn it which might be **inaccurate**.

Distributionally robust optimization (DRO) formulation:

$$\boldsymbol{z}^*(\boldsymbol{x}) \coloneqq \arg\min_{\boldsymbol{z}} \max_{\boldsymbol{Q} \in \mathcal{B}_\rho\left(\widehat{\mathrm{P}_{\boldsymbol{\theta}}}(\cdot|\boldsymbol{x})\right)} \mathrm{E}^{\boldsymbol{Q}}[\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})] \quad \text{s.t. } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{z}) \leq \boldsymbol{0}$$

- Even the learned distribution is inaccurate, the decision has robustness
- **Q1**: How to learn the distribution? **Q2**: How to solve DRO?

Problem Formulation

Unknown conditional distribution

Original Problem:

$$\boldsymbol{z}^*(\boldsymbol{x}) := \arg\min_z \mathrm{E}^{\boldsymbol{P}}[\,\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \mid \boldsymbol{x}\,] \text{ s.t. } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{z}) \leq \boldsymbol{0}$$

- Notice Conditional Distribution of $P(\boldsymbol{y}|\boldsymbol{x})$ is **unknown**.
- But we can learn it which might be **inaccurate**.

Ambiguity Set of Probability

Distributionally robust optimization (DRO) formulation:

$$\boldsymbol{z}^*(\boldsymbol{x}) := \arg\min_{\boldsymbol{z}} \max_{\boldsymbol{Q}\in\mathcal{B}_\rho\left(\widehat{\mathrm{P}_{\boldsymbol{\theta}}}(\cdot|\boldsymbol{x})\right)} \mathrm{E}^{\boldsymbol{Q}}[\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})] \quad \text{s.t. } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{z}) \leq \boldsymbol{0}$$

- Even the learned distribution is inaccurate, the decision has robustness
- **Q1**: How to learn the distribution? **Q2**: How to solve DRO?

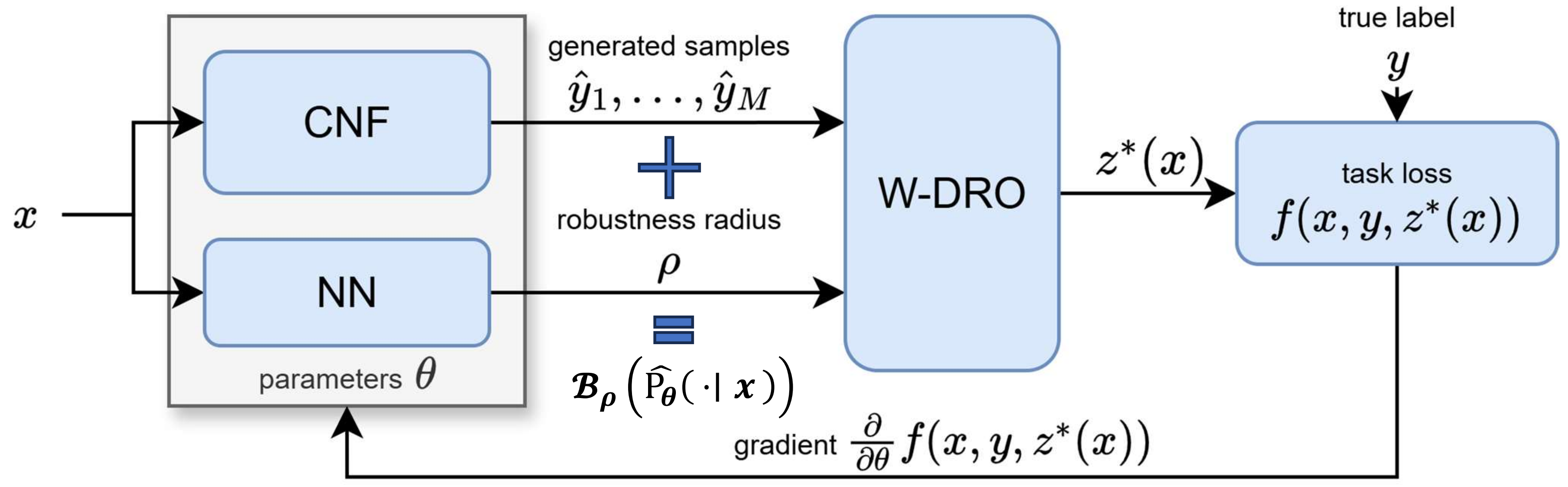
Gen-WDRO Framework

1. Conditional Normalizing Flow

2. Wasserstein Ambiguity Set

3. Adaptive Radius

4. End-to-End Training



Experimental Setup

Application: Grid-Scale Battery Storage

Day-ahead electricity market participation with optimal charging/discharging schedule

Inputs

- Past electricity prices
- Temperature data
- Energy load forecast
- Calendar features

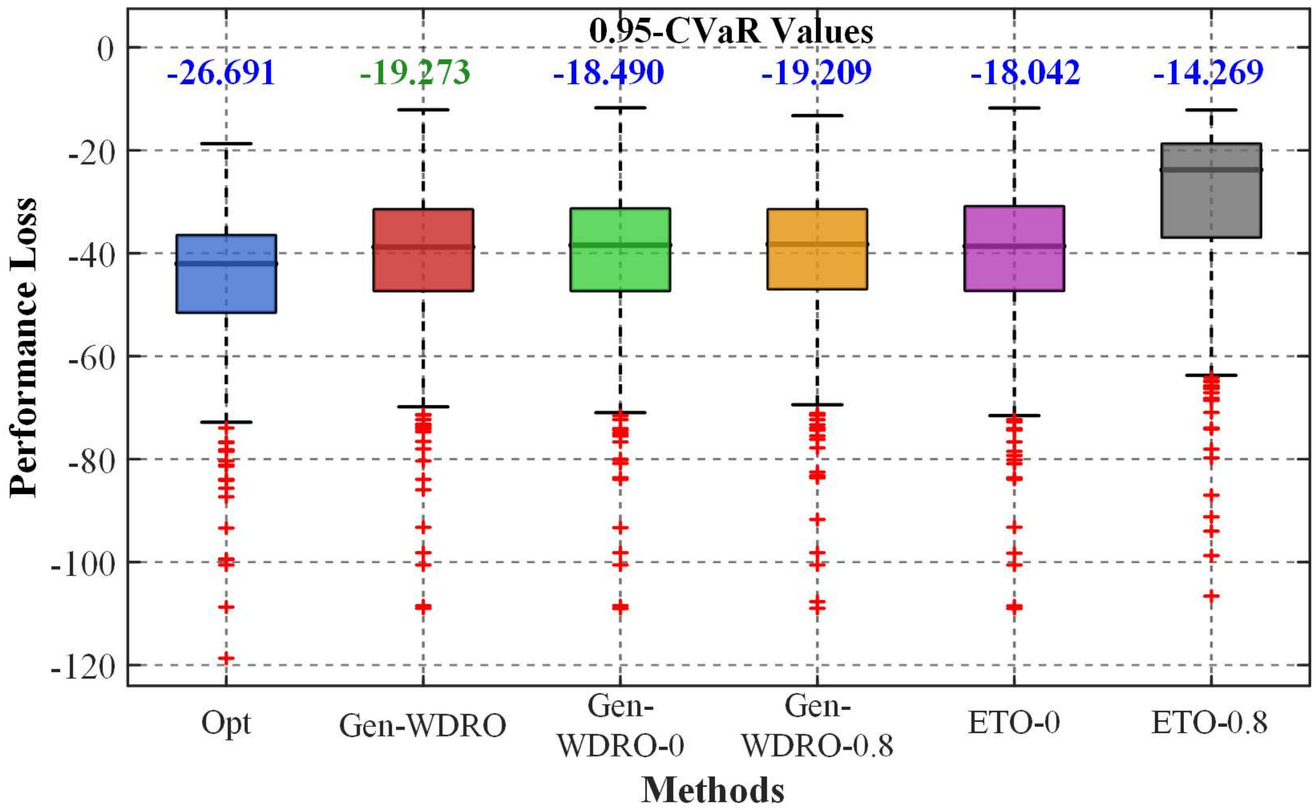
Decisions

- Charging schedule (z)
- Discharging schedule (z)
- 24-hour horizon

Distribution Shift Test

Gaussian noise N(0, 0.3) added to prices, simulating increased volatility from renewables

Results: Superior Robustness



Takeaways:

- **Adaptive Radius is necessary:** Our method with a learnable radius outperforms fixed-radius versions (Gen-WDRO-0 and Gen-WDRO-0.8), showing that adapting the uncertainty quantification is critical.
- **End-to-end training is necessary**: directly improving the decision performance via doing back propagation on generative model can indeed improve the performance.

Conclusion & Climate Impact

Contributions

- Novel integration of generative modeling with DRO
- Tractable convex reformulation enabling efficient end-to-end learning
- Adaptive uncertainty quantification via learned robustness radius

Impact on Climate Goals

By improving battery storage profitability and robustness, Gen-WDRO supports deployment of energy storage systems that enable greater renewable energy integration.