Robust Decisions via Generative Wasserstein Distributionally Robust Optimization

Han Xu **Christopher Yeh**
Department of Computing and Mathematical Sciences
California Institute of Technology
{hxu2,cyeh}@caltech.edu

Abstract

Decision-making under uncertainty is fundamental to many real-world applications, from energy systems to finance, where agents must make optimal decisions without knowing future outcomes. This work proposes Gen-WDRO, a novel generative Wasserstein distributionally robust optimization framework that combines conditional normalizing flows with distributionally robust optimization for robust decision-making under distribution shift. Our approach learns conditional distributions via normalizing flows, constructs Wasserstein ambiguity sets around these learned distributions, and employs neural networks to adaptively determine robustness radii. We prove that under linear cost structures, the resulting distributionally robust problem can be reformulated as a tractable convex optimization problem, enabling efficient end-to-end training that simultaneously improves performance and enhances robustness against distribution shift. Experiments on battery storage management under distribution shift demonstrate that Gen-WDRO achieves superior robustness with the best CVaR performance, validating the effectiveness of adaptive uncertainty quantification for robust decision-making.

1 Introduction

Machine learning techniques are widely used for predicting uncertain quantities for decision-making under uncertainty in many real-world problems including energy systems [4, 8], finance [3], supply chain management [7], among many other domains. Traditional predict-then-optimize and estimate-then-optimize (ETO) approaches first train predictive models for prediction accuracy, then use predictions in downstream optimization [5]. However, this separation between training and optimization causes misalignment: prediction errors optimized for statistical metrics may severely degrade decision quality which is also known as the "optimizer's curse" [14]. End-to-end (E2E) approaches (also known as decision-focused learning [11]) address this by training models to directly minimize downstream decision costs using differentiable optimization [4].

Recent E2E approaches have explored incorporating risk aversion by predicting an uncertainty set instead of a point estimate for the uncertain parameter, and then solving a robust optimization (RO) problem [2, 17, 20] to minimize the worst-case cost over the learned uncertainty set. However, the E2E-RO approach has two limitations. First, to ensure tractability, E2E-RO requires the uncertainty set to be convex, which limits the expressivity of the uncertainty representation. Additionally, the worst-case optimization approach can lead to overly conservative solutions [15]. Due to space constraints, we include a longer discussion about related works in Appendix A.

Instead of learning an uncertainty set for the uncertain parameter, an alternative approach is to quantify uncertainty around a learned conditional distribution of the parameter. **In this work, we propose a novel generative Wasserstein distributionally robust optimization framework (Gen-WDRO)**

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Generative AI in Finance.

that combines generative modeling with distributionally robust optimization for end-to-end learning. Our approach provides a flexible uncertainty representation, and we demonstrate robustness against distribution shift in experiments on a battery storage participation in day-ahead electricity markets.

2 Problem Formulation and Methodology

2.1 Problem Statement

Consider an agent that observes an input $x \in \mathbb{R}^m$ and must make a decision $z \in \mathbb{R}^p$ without knowing the future outcome $y \in \mathbb{R}^n$. Once z is chosen and y is revealed, the agent incurs a task loss $f(x, y, z)$ defined by a loss function $f : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$. We assume that the decision z must adhere to hard constraints $g(x, z) \leq 0$ which do not depend on the future outcome y. The future outcome y is uncertain and only influences the cost function $f(x, y, z)$. Let $\mathbb{P}$ denote the unknown joint distribution of (x, y). The agent's goal is to minimize the expected cost while satisfying the constraints:

$$z^*(x) := \arg\min_z \mathbb{E}^{\mathbb{P}}[f(x, y, z) \mid x] \quad \text{s.t. } g(x, z) \leq 0 \tag{1}$$

where $\mathbb{E}^{\mathbb{P}}[\cdot \mid x]$ denotes conditional expectation given x.

The above problem cannot be solved directly since the conditional distribution $\mathbb{P}(y \mid x)$ is not known. Instead, we assume the agent has access to a set of i.i.d. samples $\{(x_i, y_i)\}_{i=1}^N$ from the joint distribution $\mathbb{P}$. Given the samples, we then learn a conditional distribution $\hat{\mathbb{P}}_\theta(\cdot \mid x)$ with learnable parameters θ to estimate the conditional expectation $\mathbb{E}^{\mathbb{P}}[f(x, y, z) \mid x]$. However, the learned conditional distribution may not be accurate, or distribution shifts may occur, leading to poor decision-making performance under the learned distribution. To address this issue, we propose using distributionally robust optimization (DRO) to ensure that the decision is robust against an ambiguity set of distributions $\mathcal{B}_\rho(\hat{\mathbb{P}}_\theta(\cdot \mid x))$ that is centered at the learned conditional distribution $\hat{\mathbb{P}}_\theta(\cdot \mid x)$ with some distribution metric $d(\cdot, \cdot)$:

$$\mathcal{B}_\rho(\hat{\mathbb{P}}_\theta(\cdot \mid x)) := \left\{ \mathbb{Q} \in \mathcal{P}(\mathbb{R}^n) : \ d(\mathbb{Q}, \hat{\mathbb{P}}_\theta(\cdot \mid x))) \leq \rho \right\}. \tag{2}$$

In this case, the agent's robust decision policy can be formulated as the following DRO problem:

$$z^*(x) := \arg\min_z \max_{\mathbb{Q} \in \mathcal{B}_\rho(\hat{\mathbb{P}}_\theta(\cdot \mid x))} \mathbb{E}^{\mathbb{Q}}[f(x, y, z)] \quad \text{s.t. } g(x, z) \leq 0 \tag{3}$$

2.2 Ambiguity Set Construction and Problem Tractability

Our challenge is to ensure tractability and differentiability of the robust decision policy (3) so that we can learn a conditional distribution $\hat{\mathbb{P}}_\theta(\cdot \mid x)$ from the data samples in an end-to-end way. We model the conditional distribution $\hat{\mathbb{P}}_\theta(\cdot \mid x)$ with a conditional normalizing flow (CNF) [19]. Because computing an ambiguity set directly around a CNF model is intractable, we instead approximate the learned CNF with an empirical distribution. For each input x, we sample M predictions $\{\hat{y}_i\}_{i=1}^M$ from the learned CNF to construct a discrete empirical distribution $\hat{\mathbb{Q}}_M(y) = \frac{1}{M} \sum_{i=1}^M \delta_{\hat{y}_i}$. Then, we construct the ambiguity set (2) using the Wasserstein metric

$$d_{\mathrm{W}}(\mathbb{Q}, \hat{\mathbb{Q}}_M) = \inf_{\gamma \in \Gamma(\mathbb{Q}, \hat{\mathbb{Q}}_M)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \|y - y'\| \, d\gamma(y, y') \tag{4}$$

where $\Gamma(\mathbb{Q}, \hat{\mathbb{Q}}_M)$ is the set of all joint distributions γ on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals $\mathbb{Q}$ and $\hat{\mathbb{Q}}_M$, and $\| \cdot \|$ is any norm on $\mathbb{R}^n$. In our experiments, we use the 2-norm $\| \cdot \|_2$.

The following theorem shows that the DRO problem (3) based on this ambiguity set can be reformulated as a tractable convex optimization problem under certain conditions.

Theorem 1. *Suppose the task loss function has the form* $f(x, y, z) = \tilde{f}(x, z) + \langle y, Fz \rangle$ *where* $F \in \mathbb{R}^{n \times p}$*, and both* $\tilde{f}(x, z)$ *and the constraint function* $g(x, z)$ *are convex in* z*. Then, problem* (3) *with ambiguity set* $\mathcal{B}_\rho(\hat{\mathbb{Q}}_M)$ *based on the Wasserstein metric can be reformulated as the convex*

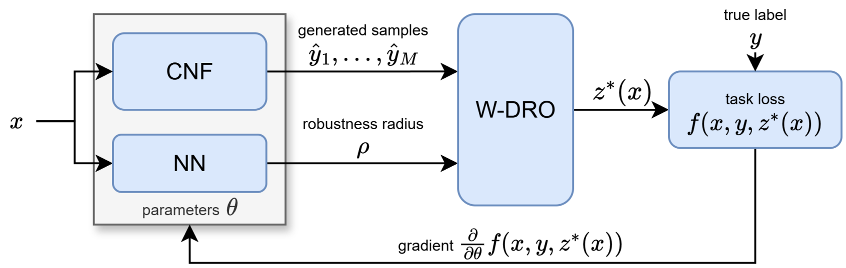

Figure 1: Our proposed Gen-WDRO method utilizes a CNF to generate a discrete distribution and a DRO layer to improve robustness. The parameters are updated using gradients from the task loss.

optimization problem

$$z^*(x) := \arg\min_z \ \rho \|Fz\|_\star + \tilde{f}(x, z) + \frac{1}{N} \sum_{i=1}^{M} \langle \hat{y}_i, Fz \rangle \quad \text{s.t. } g(x, z) \leq 0 \tag{5}$$

where $\|\cdot\|_\star$ *is the dual norm of the norm* $\|\cdot\|$ *used in the definition of the Wasserstein distance* (4)*.*

The proof of this theorem is a simple application of [12, Corollary 5.1] by reformulating the inner maximization problem in (3).

Notice that the choice of the radius ρ is flexible. In our approach, we employ a neural network to learn the radius ρ as a function of the input x. In practice, the uncertainty of the learned distribution can vary depending on the specific input x, so it is important for the radius ρ to adapt accordingly.

2.3 End-to-end Training

Algorithm 1 End-to-end training of the Gen-WDRO method.

function TRAIN(training data $D = \{(x_i, y_i)\}_{i=1}^N$, initial model parameters θ)
  for mini-batch $B \subset \{1, \ldots, N\}$ **do**
    for $i \in B$ **do**
      Compute the gradient of the likelihood loss g_i^{lik}.
      Sample M independent samples $\{\hat{y}_i\}$ from the CNF neural network.
      Solve the convex optimization problem (5).
      Compute the gradient of the task loss: $g_i^{\text{task}} = \partial f(x_i, y_i, z^*(x_i))/\partial\theta$.
    Update parameters θ using the gradients $\sum_{i \in B} \alpha_{\text{task}} g_i^{\text{task}} + \alpha_{\text{lik}} g_i^{\text{lik}}$

After reformulating the DRO problem as a convex optimization problem, we propose a methodology in Algorithm 1 to enable end-to-end training of the decision-making policy. For simplicity, the parameters of both the CNF model and the neural network for ρ are collectively denoted as θ. Our end-to-end training approach uses mini-batch gradient descent to minimize a weighted loss function

$$\ell_i(\theta) = \alpha_{\text{lik}} \log \hat{\mathbb{P}}_\theta(y_i \mid x_i) + \alpha_{\text{task}} f(x_i, y_i, z^*(x_i))$$

which accounts for both the likelihood loss of the CNF neural network and the task loss of the decision-making policy. The gradient of the likelihood loss $g_i^{\text{lik}} := \frac{\partial}{\partial \theta} \log \hat{\mathbb{P}}_\theta(y_i \mid x_i)$ is computed as in [19]. The gradient of the task loss on a single instance is $g_i^{\text{task}} := \frac{\partial f}{\partial z}|_{(x_i, y_i, z^*)} \frac{\partial z^*}{\partial \theta}|_{x_i}$, where $\frac{\partial z^*}{\partial \theta}|_{x_i}$ can be computed by differentiating through the Karush-Kuhn-Tucker (KKT) conditions of the convex reformulated DRO problem (5) [1]. Specifically, $\frac{\partial z^*}{\partial \theta}$ can be expressed as $\frac{\partial z^*}{\partial \theta} = \sum_{i=1}^M \frac{\partial z^*}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial \theta} + \frac{\partial z^*}{\partial \rho} \frac{\partial \rho}{\partial \theta}$, and the gradients $\frac{\partial z^*}{\partial \hat{y}_i}$ and $\frac{\partial z^*}{\partial \rho}$ can be computed by differentiating through the convex optimization problem.

3 Experimental Results

We consider the problem of grid-scale battery storage participating in day-ahead electricity markets, where the operator seeks to optimally schedule charging and discharging operations for the next day to

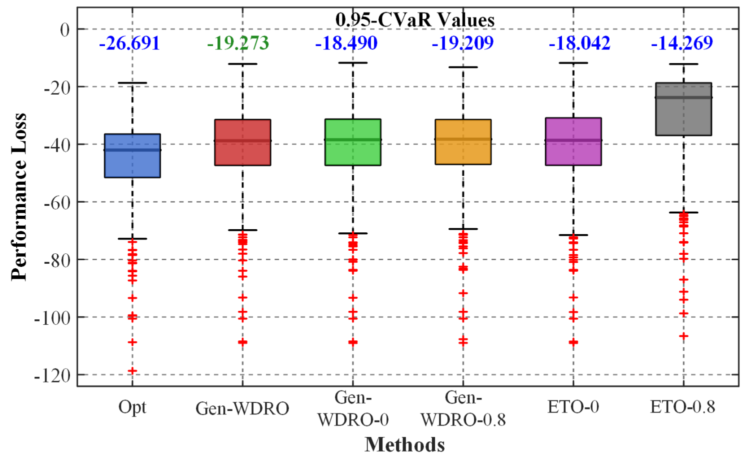

Figure 2: Performance comparison on the battery management problem. Each method is tested on 438 test cases, and 0.95-CVaR is adopted for robustness comparison. Lower values are better.

maximize profit through price arbitrage while satisfying operational constraints. We follow the same setup as in Donti et al. [4]. The input features x include the past day's prices and temperature, the next day's energy load forecast and temperature forecast, binary indicators of weekends or holidays, and yearly sinusoidal features. The operator utilizes these features to predict future energy prices $y \in \mathbb{R}^T$ over a T-step horizon, and decides how much and when to charge ($z^{\text{in}} \in \mathbb{R}^T$) or discharge ($z^{\text{out}} \in \mathbb{R}^T$) the battery, together represented by the decision variable z. Due to space constraints, we specify the specific form of task loss f and constraint g in Appendix B. Code will be released on GitHub upon publication.

We evaluate our approach on this problem under distribution shift, where the test set is corrupted with Gaussian $\mathcal{N}(0, 0.3)$ noise added to standardized future electricity prices. This simulates the increased volatility expected from growing renewable energy penetration.

Figure 2 compares our Gen-WDRO approach (with trainable radius) against other baselines: ETO-0 and ETO-0.8, representing two-stage estimate-then-optimize approaches using fixed-radius DRO layers with $\rho = 0$ and $\rho = 0.8$ ($\rho = 0.8$ is chosen as the mean radius from our Gen-WDRO model), and fixed-radius variants Gen-WDRO-0 and Gen-WDRO-0.8. The "Opt" benchmark represents the optimum obtained using ground-truth prices for optimization.

Our Gen-WDRO method demonstrates superior robustness with the best 0.95-CVaR performance,[1] and outperforms baselines. Additionally, Gen-WDRO achieves more stable performance compared to fixed-radius approaches and substantially outperforms ETO methods. This shows that learning the uncertainty radius end-to-end, rather than fixing it, leads to better performances.

4 Conclusion

We have presented a generative Wasserstein distributionally robust optimization (Gen-WDRO) method for end-to-end decision-making under uncertainty. By integrating generative modeling with DRO, our approach enables flexible conditional distribution learning and enhances robustness against distribution shifts. Experimental results demonstrate the method's effectiveness for robust decision-making in electricity markets despite distributional changes, indicating potential for broader financial applications such as portfolio construction and risk management.

[1]The α-conditional value-at-risk (CVaR) of a random variable X is defined as $\mathbb{E}[X \mid X \geq q_\alpha(X)]$, where $q_\alpha(X)$ is the α-quantile of X [16]. CVaR is a common metric for evaluating robustness [17, 18].

Acknowledgments and Disclosure of Funding

The authors acknowledge support from a Quad Fellowship, NSF Grants (CNS-2146814, CPS-2136197, CNS-2106403, and NGSDI-2105648), S2I Initiative of Caltech, and the Caltech Resnick Sustainability Institute.

References

[1] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and J. Z. Kolter. Differentiable Convex Optimization Layers. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://papers.nips.cc/paper_files/paper/2019/hash/9ce3c52fc54362e22053399d3181c638-Abstract.html.

[2] A. Chenreddy and E. Delage. End-to-end Conditional Robust Optimization, Mar. 2024. URL http://arxiv.org/abs/2403.04670. arXiv:2403.04670 [cs].

[3] G. Costa and G. N. Iyengar. Distributionally robust end-to-end portfolio construction. *Quantitative Finance*, 23(10):1465–1482, Oct. 2023. ISSN 1469-7688. doi: 10.1080/14697688.2023.2236148. URL https://doi.org/10.1080/14697688.2023.2236148. Publisher: Routledge _eprint: https://doi.org/10.1080/14697688.2023.2236148.

[4] P. Donti, B. Amos, and J. Z. Kolter. Task-based End-to-end Model Learning in Stochastic Optimization. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://papers.nips.cc/paper_files/paper/2017/hash/3fc2c60b5782f641f76bcefc39fb2392-Abstract.html.

[5] A. N. Elmachtoub and P. Grigas. Smart "Predict, then Optimize". *Management Science*, 68(1):9–26, Jan. 2022. ISSN 0025-1909. doi: 10.1287/mnsc.2020.3922. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2020.3922. Publisher: INFORMS.

[6] J. Ho, A. Jain, and P. Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

[7] J. Jiang, X. Li, and J. Zhang. Online Stochastic Optimization with Wasserstein-Based Nonstationarity. *Management Science*, Mar. 2025. ISSN 0025-1909. doi: 10.1287/mnsc.2020.03850. URL https://pubsonline.informs.org/doi/full/10.1287/mnsc.2020.03850. Publisher: INFORMS.

[8] Z. Liang, Q. Li, A. Liu, and Y. Dvorkin. Prescribing Decision Conservativeness in Two-Stage Power Markets: A Distributionally Robust End-to-End Approach. In *2025 59th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6, Mar. 2025. doi: 10.1109/CISS64860.2025.10944764. URL https://ieeexplore.ieee.org/abstract/document/10944764. ISSN: 2837-178X.

[9] S. Luo, J. Guan, J. Ma, and J. Peng. A 3D Generative Model for Structure-Based Drug Design. *Advances in Neural Information Processing Systems*, 34:6229–6239, Dec. 2021. URL https://proceedings.neurips.cc/paper/2021/hash/314450613369e0ee72d0da7f6fee773c-Abstract.html?ref=https://githubhelp.com.

[10] X. Ma, C. Ning, and W. Du. Differentiable Distributionally Robust Optimization Layers. In *Proceedings of the 41st International Conference on Machine Learning*, pages 33880–33901. PMLR, July 2024. URL https://proceedings.mlr.press/v235/ma24j.html. ISSN: 2640-3498.

[11] J. Mandi, J. Kotary, S. Berden, M. Mulamba, V. Bucarey, T. Guns, and F. Fioretto. Decision-Focused Learning: Foundations, State of the Art, Benchmark and Future Opportunities. *Journal of Artificial Intelligence Research*, 80:1623–1701, Aug. 2024. ISSN 1076-9757. doi: 10.1613/jair.1.15320. URL https://www.jair.org/index.php/jair/article/view/15320.

[12] P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):115–166, Sept. 2018. ISSN 0025-5610, 1436-4646. doi: 10.1007/s10107-017-1172-1. URL http://link.springer.com/10.1007/s10107-017-1172-1.

[13] V. A. Nguyen, F. Zhang, S. Wang, J. Blanchet, E. Delage, and Y. Ye. Robustifying Conditional Portfolio Decisions via Optimal Transport, Apr. 2024. URL http://arxiv.org/abs/2103.16451. arXiv:2103.16451 [q-fin].

[14] J. E. Smith and R. L. Winkler. The Optimizer's Curse: Skepticism and Postdecision Surprise in Decision Analysis. *Management Science*, 52(3):311–322, Mar. 2006. ISSN 0025-1909. doi: 10.1287/mnsc.1050.0451. URL https://pubsonline.informs.org/doi/10.1287/mnsc.1050.0451. Publisher: INFORMS.

[15] A. Thiele. A note on issues of over-conservatism in robust optimization with cost uncertainty. *Optimization*, 59(7):1033–1040, Oct. 2010. ISSN 0233-1934. doi: 10.1080/02331930903395592. URL https://doi.org/10.1080/02331930903395592. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02331930903395592.

[16] S. Uryasev and R. T. Rockafellar. Conditional Value-at-Risk: Optimization Approach. In S. Uryasev and P. M. Pardalos, editors, *Stochastic Optimization: Algorithms and Applications*, pages 411–435. Springer US, Boston, MA, 2001. ISBN 978-1-4757-6594-6. doi: 10.1007/978-1-4757-6594-6_17. URL https://doi.org/10.1007/978-1-4757-6594-6_17.

[17] I. Wang, C. Becker, B. Van Parys, and B. Stellato. Learning Decision-Focused Uncertainty Sets in Robust Optimization, July 2024. URL http://arxiv.org/abs/2305.19225. arXiv:2305.19225 [math].

[18] P. Z. Wang, J. Liang, S. Chen, F. Fioretto, and S. Zhu. Gen-DFL: Decision-Focused Generative Learning for Robust Decision Making, Feb. 2025. URL http://arxiv.org/abs/2502.05468. arXiv:2502.05468 [cs].

[19] C. Winkler, D. Worrall, E. Hoogeboom, and M. Welling. Learning Likelihoods with Conditional Normalizing Flows, Nov. 2023. URL http://arxiv.org/abs/1912.00042. arXiv:1912.00042 [cs].

[20] C. Yeh, N. Christianson, A. Wu, A. Wierman, and Y. Yue. End-to-End Conformal Calibration for Optimization Under Uncertainty, Sept. 2024. URL http://arxiv.org/abs/2409.20534. arXiv:2409.20534 [cs].

A Related Works

Robust End-to-End Learning. Donti et al. [4] pioneered task-based end-to-end learning, which trains models to minimize decision costs rather than prediction error. Differentiable optimization layers [1] enable tractable gradient computation through convex programs. Yeh et al. [20] utilize conformal prediction to calibrate uncertainty sets and employ input-convex neural networks to represent general convex uncertainty sets.

Distributionally Robust End-to-End Learning. Costa and Iyengar [3] developed KL-divergence-based distributionally robust portfolio optimization. Ma et al. [10] proposed generic differentiable DRO layers for mixed-integer problems using second-order cone ambiguity sets, mentioning the possibility of using Wasserstein distance to construct ambiguity sets but without experimental validation. Nguyen et al. [13] propose a conditional distributionally robust optimization strategy that solves conditional DRO problems without requiring explicit learning of the conditional distribution. Liang et al. [8] propose using DRO to solve two-stage decision-making problems in an end-to-end manner, though they do not learn conditional distributions.

Generative Models for Decision-Making. Our work is also related to generative modeling which has been applied in various fields such as image generation [6] and drug design [9]. In the context of decision-making, Wang et al. [18] employ generative models to learn conditional distributions for CVaR optimization. Rather than using robust optimization, they directly optimize the CVaR value as their objective function.

Unlike existing work, our approach integrates generative modeling with DRO for end-to-end learning under distribution shift, providing both flexible conditional distribution learning and enhanced robustness guarantees.

B Experimental Setting Details

The battery is characterized by its storage capacity B, charging efficiency γ, and maximum charging and discharging rates c_{in} and c_{out}, respectively. The objective function balances multiple goals: profit maximization, and maintaining operational flexibility by keeping the battery state of charge near 50% of capacity (weighted by parameter λ) to enable participation in multiple markets.

$$f(y, z) = \sum_{t=1}^{T} y_t(z_t^{\text{in}} - z_t^{\text{out}}) + \lambda \left\| z_t^{\text{state}} - \frac{B}{2} \right\|^2$$

The constraints are given by

$$\begin{gathered} z_0^{\text{state}} = B/2, \quad z_t^{\text{state}} = z_{t-1}^{\text{state}} - z_t^{\text{out}} + \gamma z_t^{\text{in}} \quad \forall t = 1, \dots, T \\ 0 \leq z_t^{\text{in}} \leq c^{\text{in}}, \quad 0 \leq z_t^{\text{out}} \leq c^{\text{out}}, \quad 0 \leq z_t^{\text{state}} \leq B. \end{gathered}$$

Following [4], we set $T = 24$, $B = 1$, $\gamma = 0.9$, $c^{\text{in}} = 0.5$, $c^{\text{out}} = 0.2$, and $\lambda = 0.1$.

We set the following parameters for our proposed Gen-WDRO model: $\alpha_{\text{task}} = 0.8$, $\alpha_{\text{lik}} = 0.2$, and $M = 10$.