GENERATIVE DISTRIBUTIONALLY ROBUST OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Generative models increasingly supply scenarios for downstream optimization, but the resulting decisions are vulnerable to model misspecification. Standard transport-based distributionally robust optimization (DRO) accepts arbitrary nominal samplers without requiring worst-case laws to respect generator structure, while existing generative ambiguity-set methods enforce such structure but typically rely on model-specific access or training-data certificates. We introduce Generative Distributionally Robust Optimization (GDRO), a general framework that accepts any pretrained conditional generative model as its nominal model and optimizes against adversarial laws induced by a chosen conditional generator family. The key is to pair sampler representations of conditional laws with a Sinkhorn certificate that compares nominal and adversarial output laws from samples alone. This enables generator-constrained robust optimization across explicit and implicit nominal models, controlling distributional proximity at the decision context without requiring nominal likelihoods or gradients. We develop a stochastic primal–dual algorithm for an expected mini-batch formulation and establish bounds linking Sinkhorn proximity to changes in expected downstream loss and quantifying the Lagrangian relaxation gap. Experiments span convex contextual newsvendor problems with explicit generators on synthetic and M5 retail demand data, and nonconvex robot navigation with the implicit SocialGAN generator. Relative to nominal decisions, GDRO reduces rare-context inventory regret by 63% on synthetic data and navigation collisions by 66%, while maintaining a 99.7% arrival rate.

1 INTRODUCTION

Conditional generative models increasingly provide the predictive distributions used in downstream optimization. Given a current context x, a pretrained model generates demand scenarios, trajectories, or other uncertain outcomes, and an optimizer chooses a decision from those samples. This interface is attractive because it accommodates rich, high-dimensional uncertainty, but it also passes any misspecification of the generator directly into the decision. The practical question is therefore not only how to generate realistic conditional outcomes, but how to robustify a downstream decision when the available generative model may be wrong.

Distributionally robust optimization (DRO) addresses model misspecification by optimizing decisions against a set of plausible distributions (Mohajerin Esfahani & Kuhn, 2018; Blanchet & Murthy, 2019; Gao & Kleywegt, 2023; Gao et al., 2024; Rahimian & Mehrotra, 2022; Wang et al., 2026; Yang et al., 2025). Standard Wasserstein and Sinkhorn formulations can use any conditional generator as a nominal sampler at the decision context x, requiring only sample access to the model (Mohajerin Esfahani & Kuhn, 2018; Wang et al., 2026). However, proximity to the nominal law alone does not require worst-case distributions to respect the dependence patterns, temporal dynamics, or manifold structure encoded by a chosen generator family. This distinction matters when high-dimensional data concentrate near low-dimensional manifolds, as posited by the manifold hypothesis (Song & Ermon, 2019).

Parametric DRO methods (Michel et al., 2021; 2022) and generative ambiguity-set methods, including GAS-DRO (Wen & Yang, 2026) and diffusion ambiguity-set DRO (Wen & Yang, 2025), incorporate structure by parameterizing adversarial distributions or likelihood ratios. Their stated

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

formulations, however, control deviations from the nominal law through density evaluations, learned likelihood ratios, or model-specific reconstruction and score-matching losses. They therefore do not directly provide a common robustification layer for arbitrary implicit, proprietary, or otherwise frozen conditional generators available only through samples, with proximity certified from nominal and adversarial output samples alone. Moreover, if a conditional extension of GAS-DRO bounds reconstruction loss only on average over training contexts, it need not ensure proximity between the adversarial and nominal laws at the decision context x.

To address these limitations, we develop a DRO framework that accepts any sampleable pretrained conditional generator as the nominal model, restricts adversarial laws to a chosen conditional generator family, and controls their proximity to the nominal law at the decision context x using output samples alone. We call the resulting ambiguity set *generator-faithful*: every admissible worst-case law is induced by the selected adversarial generator family, rather than being an arbitrary distribution in the ambient outcome space. When the nominal architecture is available, a natural choice is to use the same generator family for both nominal and adversarial models, thereby preserving its inductive structure. More generally, the framework accommodates any black-box nominal sampler and permits the adversarial generator family to be chosen independently.

Our key observation is that the right interface is the *sampler–Sinkhorn pairing*. Every sampleable conditional generative model exposes its law operationally through a sampler, regardless of whether it has a tractable likelihood, score, or density. Sinkhorn divergence is naturally compatible with this interface: it is a transport discrepancy between probability laws, can be estimated from samples alone, and is differentiable with respect to generated outputs (Cuturi, 2013; Genevay et al., 2018; Feydy et al., 2019). It also carries decision-level meaning: under bounded output support, Sinkhorn proximity controls Wasserstein proximity, and Kantorovich–Rubinstein duality then bounds the change in expected loss uniformly over Lipschitz objectives (Theorem 3.1). Sinkhorn divergences have previously been used to fit generative models to data; here we use them to constrain a conditional adversary to remain close to the nominal law while maximizing expected downstream loss.

Generative distributionally robust optimization (GDRO). Let $P_{\hat{\phi}}(\cdot \mid x)$ denote the conditional law supplied by a pretrained nominal generator and $\{Q_\psi(\cdot \mid x) : \psi \in \Psi\}$ the conditional laws induced by a chosen adversarial generator family. Given a downstream loss $f : \mathcal{W} \times \mathcal{Y} \to \mathbb{R}$, let S_ε denote the debiased Sinkhorn divergence with regularization ε, and let ρ be the ambiguity radius. Our central population problem is

$$w^\star_{\rho,\varepsilon}(x) \in \arg\min_{w \in \mathcal{W}} \sup_{\psi \in \Psi:\, S_\varepsilon(Q_\psi(\cdot|x), P_{\hat{\phi}}(\cdot|x)) \le \rho} \mathbb{E}_{Y \sim Q_\psi(\cdot|x)}[f(w, Y)]\,. \tag{1}$$

The generator family determines which perturbations are structurally admissible, while the Sinkhorn constraint controls their deviation from the nominal conditional law. Both the worst-case loss and the proximity constraint are evaluated at the decision context x. Section 2 presents the exact sampler reformulation, a likelihood-free expected mini-batch formulation, and the resulting stochastic primal–dual algorithm.

Relation to prior work. FlowDRO (Xu et al., 2024; Kobyzev et al., 2021) uses neural flow maps to approximate a Wasserstein adversary, but its underlying ambiguity set is a Wasserstein ball without an explicit generator-family constraint. DRAGEN (Ren & Majumdar, 2022) generates adversarial environments under a Wasserstein constraint in a learned latent space, whereas GDRO constrains conditional output laws at the decision context. Decision-focused learning (Donti et al., 2017; Elmachtoub & Grigas, 2022; Costa & Iyengar, 2023; Wang et al., 2025; Ma et al., 2024) learns predictive models or uncertainty sets for downstream performance, while GDRO constructs a post-training ambiguity set around a frozen conditional sampler. Table 1 highlights the combination of capabilities that distinguishes GDRO.

Contributions. We introduce GDRO, to our knowledge the first context-local DRO framework that constrains a generator-family adversary by sample-based Sinkhorn proximity to an arbitrary frozen conditional sampler. We derive exact sampler and expected mini-batch formulations, develop a stochastic primal–dual algorithm, and establish downstream risk bounds, relaxation-gap, and finite-time optimization guarantees. Across conditional VAE, CVAE–LSTM, and implicit SocialGAN generators, GDRO reduces rare-context inventory regret by 63% and navigation collisions by 66% relative to nominal decisions.

Table 1: Structural capabilities of representative DRO formulations. ✓ denotes native support, △ variant-dependent or model-specific support, and × no support.

Method	Arbitrary frozen nominal sampler	Generator-family adversary	Output-law transport	Active-context certificate
Conditional Wasserstein / Sinkhorn DRO	✓	×	✓	✓
Parametric adversary DRO	×	△	×	△
Diffusion ambiguity-set DRO / GAS-DRO	×	✓	×	△
GDRO (ours)	✓	✓	✓	✓

2 Generative Distributionally Robust Optimization (GDRO)

2.1 Conditional Population Formulation

Let $\mathcal{X}$ be the context space, $\mathcal{Y}$ the outcome space, $\mathcal{W} \subseteq \mathbb{R}^d$ the decision set, and $f : \mathcal{W} \times \mathcal{Y} \to \mathbb{R}$ the downstream loss. We fix the active decision context $x \in \mathcal{X}$ throughout. Write

$$P_0^x := P_{\hat{\phi}}(\cdot \mid x), \qquad Q_\psi^x := Q_\psi(\cdot \mid x), \quad \psi \in \Psi,$$

for the nominal conditional law and a conditional law in the chosen adversarial generator family.

For $\varepsilon > 0$, probability measures P, Q on $\mathcal{Y}$, and a ground cost $c : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_+$, define entropically regularized optimal transport by

$$\mathrm{OT}_\varepsilon(P, Q) := \inf_{\pi \in \Pi(P,Q)} \left\{ \int c(y, y') \, \mathrm{d}\pi(y, y') + \varepsilon \, \mathrm{KL}(\pi \,\|\, P \otimes Q) \right\}, \tag{2}$$

where $\Pi(P, Q)$ denotes the couplings of P and Q. The debiased Sinkhorn divergence (Feydy et al., 2019) is

$$S_\varepsilon(P, Q) := \mathrm{OT}_\varepsilon(P, Q) - \tfrac{1}{2} \mathrm{OT}_\varepsilon(P, P) - \tfrac{1}{2} \mathrm{OT}_\varepsilon(Q, Q). \tag{3}$$

For boundedly supported laws on a Euclidean outcome space and the squared-Euclidean ground cost $c(y, y') = \|y - y'\|_2^2$, S_ε is nonnegative, positive definite, and continuous for weak convergence (Feydy et al., 2019; Peyré & Cuturi, 2019).

For an ambiguity radius $\rho \geq 0$, the conditional population GDRO problem is

$$V_{\rho,\varepsilon}(x) := \inf_{w \in \mathcal{W}} \sup_{\psi \in \Psi:\ S_\varepsilon(Q_\psi^x, P_0^x) \leq \rho} \mathbb{E}_{Y \sim Q_\psi^x}[f(w, Y)]. \tag{4}$$

The feasible laws are generator-faithful: every adversarial law is induced by the selected generator family. When the nominal architecture is available, we include its parameter $\hat{\phi} \in \Psi$, so that $Q_{\hat{\phi}}^x = P_0^x$.

2.2 Sampler Representation

Let the nominal and adversarial conditional samplers be

$$G_{\hat{\phi}} : \mathcal{Z}_0 \times \mathcal{X} \to \mathcal{Y}, \quad Z_0 \sim \zeta_0, \qquad G_\psi : \mathcal{Z}_\mathrm{A} \times \mathcal{X} \to \mathcal{Y}, \quad Z_\mathrm{A} \sim \zeta_\mathrm{A}.$$

Their conditional output laws are the pushforwards

$$P_0^x = \big(G_{\hat{\phi}}(\cdot, x)\big)_\# \zeta_0, \qquad Q_\psi^x = \big(G_\psi(\cdot, x)\big)_\# \zeta_\mathrm{A}. \tag{5}$$

Consequently, equation 4 has the exact sampler form

$$V_{\rho,\varepsilon}(x) = \inf_{w \in \mathcal{W}} \sup_{\psi \in \Psi:\ S_\varepsilon((G_\psi(\cdot,x))_\# \zeta_\mathrm{A}, (G_{\hat{\phi}}(\cdot,x))_\# \zeta_0) \leq \rho} \mathbb{E}_{Z_\mathrm{A} \sim \zeta_\mathrm{A}}[f(w, G_\psi(Z_\mathrm{A}, x))]. \tag{6}$$

This representation requires only nominal output samples; it does not require likelihoods, scores, or an inverse map.

2.3 Expected Mini-batch Formulation

Computing the population Sinkhorn divergence in equation 6 exactly is generally infeasible. We therefore state explicitly the expected mini-batch problem targeted by the re-sampled implementation.

Fix $\kappa \in \Pi(\zeta_0, \zeta_{\mathrm{A}})$ and draw

$$\mathcal{B} = \{(Z_{0,i}, Z_{\mathrm{A},i})\}_{i=1}^{M} \overset{\text{iid}}{\sim} \kappa^{\otimes M}.$$

The product coupling, which draws the two clouds independently, is always available. When the samplers share a latent interface, one may instead use common random numbers by setting $Z_{0,i} = Z_{\mathrm{A},i} = Z_i$. This couples only the Monte Carlo draws: Sinkhorn still optimizes over all couplings of the two empirical output measures. All results below allow any fixed κ.

For a realized batch $\mathcal{B}$, define

$$\widehat{P}^{\mathcal{B}}_{G_{\hat{\phi}},x} := \frac{1}{M}\sum_{i=1}^{M} \delta_{G_{\hat{\phi}}(Z_{0,i},x)}, \qquad \widehat{P}^{\mathcal{B}}_{G_{\psi},x} := \frac{1}{M}\sum_{i=1}^{M} \delta_{G_{\psi}(Z_{\mathrm{A},i},x)}. \tag{7}$$

The realized batch objective and certificate are

$$\widehat{F}^{\mathcal{B}}_{M}(w,\psi) := \frac{1}{M}\sum_{i=1}^{M} f(w, G_{\psi}(Z_{\mathrm{A},i}, x)), \tag{8}$$

$$\widehat{s}^{\mathcal{B}}_{M}(\psi) := S_{\varepsilon}\Big(\widehat{P}^{\mathcal{B}}_{G_{\psi},x}, \widehat{P}^{\mathcal{B}}_{G_{\hat{\phi}},x}\Big). \tag{9}$$

The re-sampled algorithm targets their expectations

$$F(w,\psi) := \mathbb{E}_{\mathcal{B}\sim\kappa^{\otimes M}}[\widehat{F}^{\mathcal{B}}_{M}(w,\psi)] = \mathbb{E}_{Y\sim Q^{x}_{\psi}}[f(w,Y)], \tag{10}$$

$$s_{M,\kappa}(\psi) := \mathbb{E}_{\mathcal{B}\sim\kappa^{\otimes M}}[\widehat{s}^{\mathcal{B}}_{M}(\psi)]. \tag{11}$$

At finite M, $s_{M,\kappa}$ is an expected mini-batch certificate, not generally the population Sinkhorn divergence (Fatras et al., 2020). To simplify notation, we henceforth fix (M,κ) and write $s(\psi) := s_{M,\kappa}(\psi)$. The finite-$M$ re-sampled GDRO problem is

$$V_{M,\kappa}(x) := \inf_{w\in\mathcal{W}} \sup_{\psi\in\Psi:\ s(\psi)\le\rho} F(w,\psi). \tag{12}$$

The stochastic algorithm re-samples $\mathcal{B}$ across updates, as shown in Algorithm 1.

2.4 Stochastic Primal–Dual Optimization

For fixed w, the expected mini-batch inner problem is

$$\sup_{\psi\in\Psi} F(w,\psi) \quad \text{subject to} \quad s(\psi) \le \rho. \tag{13}$$

Its expected Lagrangian and a batch estimator are

$$\mathcal{L}(w,\psi,\mu) := F(w,\psi) + \mu(\rho - s(\psi)), \qquad \mu \ge 0, \tag{14}$$

$$\widehat{\mathcal{L}}^{\mathcal{B}}(w,\psi,\mu) := \widehat{F}^{\mathcal{B}}_{M}(w,\psi) + \mu(\rho - \widehat{s}^{\mathcal{B}}_{M}(\psi)). \tag{15}$$

For a new batch drawn after (w,ψ,μ) is fixed,

$$\mathbb{E}\Big[\widehat{\mathcal{L}}^{\mathcal{B}}(w,\psi,\mu)\Big] = \mathcal{L}(w,\psi,\mu), \tag{16}$$

$$\mathbb{E}\Big[\nabla_{\psi}\widehat{\mathcal{L}}^{\mathcal{B}}(w,\psi,\mu)\Big] = \nabla_{\psi}\mathcal{L}(w,\psi,\mu), \tag{17}$$

where the second identity requires the standard dominated-gradient condition (Genevay et al., 2018; Fatras et al., 2020). The first is used in the dual analysis; the second justifies stochastic gradients.

Model flexibility. As discussed above, GDRO requires only conditional samples from the nominal model, without likelihoods, scores, or differentiability. This separates nominal sampling from adversarial optimization: users can retain their existing nominal model and choose a separate, trainable adversarial architecture suited to their application.

Algorithm 1 Stochastic Mini-batch Primal–Dual GDRO

Require: Context x; nominal sampler $G_{\hat{\phi}}$; adversarial sampler G_{ψ}; coupling κ; radius ρ; batch size M; dual bound $\mu_{\max}$; step sizes $\eta_\psi, \eta_\mu, \eta_w$; iteration counts J, K, T.

Initialize $w \in \mathcal{W}$, $\psi \in \Psi$, and $\mu_{1,1} \in [0, \mu_{\max}]$.
for $t = 1, \dots, T$ **do**
  $w^{(t)} \leftarrow w$.
  for $k = 1, \dots, K$ **do**
    for $j = 1, \dots, J$ **do**
      Draw $\mathcal{B}^{\psi}_{t,k,j} \sim \kappa^{\otimes M}$.
      $\psi \leftarrow \psi + \eta_\psi \nabla_\psi \widehat{\mathcal{L}}^{\mathcal{B}^{\psi}_{t,k,j}}(w^{(t)}, \psi, \mu_{t,k})$.
    end for
    Set $\psi_{t,k} \leftarrow \psi$.
    Draw a fresh certificate batch $\mathcal{B}^{s}_{t,k} \sim \kappa^{\otimes M}$ after $\psi_{t,k}$ is fixed.
    $\widehat{s}_{t,k} \leftarrow \widehat{s}_M^{\mathcal{B}^{s}_{t,k}}(\psi_{t,k})$.
    $\mu_{t,k+1} \leftarrow \Pi_{[0,\mu_{\max}]}(\mu_{t,k} + \eta_\mu(\widehat{s}_{t,k} - \rho))$.
    Draw a fresh outer batch $\mathcal{B}^{w}_{t,k} \sim \kappa^{\otimes M}$ and set $\widehat{g}_{t,k} \leftarrow \nabla_w \widehat{F}_M^{\mathcal{B}^{w}_{t,k}}(w^{(t)}, \psi_{t,k})$.
  end for
  $\mu_{t+1,1} \leftarrow \mu_{t,K+1}$.
  $w \leftarrow \Pi_{\mathcal{W}}\Big(w^{(t)} - \eta_w K^{-1} \sum_{k=1}^{K} \widehat{g}_{t,k}\Big)$.
end for
return $\overline{w} = T^{-1} \sum_{t=1}^{T} w^{(t)}$.

3 Theoretical Analysis

We first show how the population and expected mini-batch Sinkhorn constraints control downstream risk. We then bound the Lagrangian relaxation gap and establish inner and joint optimization guarantees for the expected mini-batch problem. All proofs are in Appendix A.

3.1 Setup

Fix the context x and mini-batch coupling κ. We use the following assumptions:

- (A1) $f(w, y)$ is L_f-Lipschitz in y, uniformly in w;
- (A2) $\mathcal{W} \subset \mathbb{R}^d$ is convex and compact, with diameter $D_{\mathcal{W}}$;
- (A3) the generator outputs are uniformly bounded: $\|G_{\hat{\phi}}(z_0, x)\|_2 \le R$ for ζ_0-almost every z_0, and $\|G_\psi(z_{\mathrm{A}}, x)\|_2 \le R$ for every $\psi \in \Psi$ and ζ_{A}-almost every z_{A}.

Note that Assumption (A2) is used only for the joint outer guarantee, which additionally requires convexity of $f(\cdot, y)$. Fix $\varepsilon, \rho > 0$ and squared-Euclidean transport cost. Write $F_0(w) := \mathbb{E}_{P_0^x}[f(w, Y)]$, $\mathcal{A} := \{\psi \in \Psi : s(\psi) \le \rho\}$, $\phi(w) := \max_{\psi \in \mathcal{A}} F(w, \psi)$, and $\phi^* := \inf_{w \in \mathcal{W}} \phi(w)$. Assume $\mathcal{A} \neq \emptyset$ and that the inner maximum is attained. Define

$$\gamma_M := \varepsilon \log M, \qquad B := \max\{\rho, 4R^2\}. \tag{18}$$

3.2 Downstream Risk Control

Let $\mathcal{P}_R$ denote the probability laws supported on the closed Euclidean ball of radius R, and define

$$\omega_{\varepsilon,R}(r) := \sup\{W_1(P, Q) : P, Q \in \mathcal{P}_R,\ S_\varepsilon(P, Q) \le r\}. \tag{19}$$

Theorem 3.1 (Downstream Risk Control)**.** *Under* (A1) *and* (A3)*, the following statements hold.*

1. *The function $\omega_{\varepsilon,R}$ is finite and nondecreasing, and $\omega_{\varepsilon,R}(r) \downarrow 0$ as $r \downarrow 0$. Hence, if $S_\varepsilon(Q^x_\psi, P^x_0) \le \rho$, then, for every $w \in \mathcal{W}$,*

$$\Big|\mathbb{E}_{Q^x_\psi}[f(w, Y)] - \mathbb{E}_{P^x_0}[f(w, Y)]\Big| \le L_f \omega_{\varepsilon,R}(\rho). \tag{20}$$

2. *For the expected mini-batch certificate associated with any fixed $\kappa \in \Pi(\zeta_0, \zeta_A)$,*

$$|F(w,\psi) - F_0(w)| \le L_f \sqrt{s(\psi) + \gamma_M} \qquad \text{for every } w \in \mathcal{W}. \tag{21}$$

In particular, every $\psi \in \mathcal{A}$ satisfies

$$|F(w,\psi) - F_0(w)| \le L_f \sqrt{\rho + \gamma_M}. \tag{22}$$

These bounds show that the Sinkhorn certificate limits changes in downstream expected loss relative to the nominal model. The proof is in Appendix A.2. Pointwise consistency of $s_{M,\kappa}$ as $M \to \infty$ is established in Lemma A.5.

3.3 Lagrangian Relaxation of the Expected Mini-Batch Problem

Define

$$Q_w(\mu) := \sup_{\psi \in \Psi} \{F(w,\psi) + \mu(\rho - s(\psi))\}, \qquad \phi_D(w) := \inf_{\mu \ge 0} Q_w(\mu), \qquad \phi_D^* := \inf_{w \in \mathcal{W}} \phi_D(w). \tag{23}$$

Weak duality, $\phi(w) \le \phi_D(w)$, always holds, whereas strong duality is not guaranteed because the inner problem is generally nonconcave. Yet the Sinkhorn constraint allows us to bound the relaxation gap explicitly.

Theorem 3.2 (Control of the Lagrangian relaxation gap)**.** *Assume* (A1) *and* (A3)*, and suppose the adversarial family contains a nominal parameter $\hat{\phi}$ satisfying $Q^x_{\hat{\phi}} = P^x_0$ and $s(\hat{\phi}) \le \rho$. Then, for every $w \in \mathcal{W}$,*

$$0 \le \phi_D(w) - \phi(w) \le C_{\rho,M}, \qquad C_{\rho,M} := L_f \sqrt{\rho + \gamma_M}. \tag{24}$$

If $\phi^* > -\infty$, then

$$0 \le \phi_D^* - \phi^* \le C_{\rho,M}. \tag{25}$$

The proof is in Appendix A.3. $C_{\rho,M}$ bounds the relaxation gap, separately from optimization error.

3.4 Optimization Guarantees for the Expected Mini-Batch Problem

We now analyze the inner and outer updates, assuming that the adversarial parameter maximizes the expected mini-batch Lagrangian at each step, i.e.,

$$\psi_k \in \arg\max_{\psi \in \Psi} \mathcal{L}(w,\psi,\mu_k), \qquad Q_w(\mu_k) = \mathcal{L}(w,\psi_k,\mu_k). \tag{26}$$

Assumption 3.1 (Dual regularity)**.** For every $w \in \mathcal{W}$, the minimization defining $\phi_D(w)$ in equation 23 has a selected minimizer $\mu^*(w) \ge 0$ satisfying $\overline{\mu}^* := \sup_{w \in \mathcal{W}} \mu^*(w) < \infty$.

Theorem 3.3 (Inner and joint optimization guarantees)**.** *Suppose the condition equation 26 holds, and let $\mu_{1,1} \in [0, \mu_{\max}]$.*

(i) Inner guarantee. *Under* (A1) *and* (A3)*, for every fixed $w \in \mathcal{W}$ and every block initialization $\mu_1 \in [0, \mu_{\max}]$, the dual step size $\eta_\mu = \mu_{\max}/(B\sqrt{K})$ gives*

$$\phi_D(w) - \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}[F(w,\psi_k)] \le \frac{B\mu_{\max}}{\sqrt{K}}. \tag{27}$$

If Assumption 3.1 also holds and $\mu_{\max} > \overline{\mu}^*$, then the constraint $s(\psi) \le \rho$ is satisfied on average in expectation up to an $O(K^{-1/2})$ error:

$$\frac{1}{K}\sum_{k=1}^{K} \mathbb{E}[s(\psi_k)] \le \rho + \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}. \tag{28}$$

(ii) Joint outer guarantee. *Under* (A1)–(A3) *and Assumption 3.1, suppose additionally that $\mu_{\max} > \overline{\mu}^*$ and $f(\cdot, y)$ is convex for every y. Write $\psi_k^{(t)}$ for the ψ_k in equation 26 with $w = w^{(t)}$ and $\mu_k = \mu_{t,k}$. With $(w^{(t)}, \psi_k^{(t)})$ fixed, draw a new batch $\mathcal{B}^w_{t,k}$ independently of the preceding history*

$\mathcal{G}_{t,k}$. *Define* $\widehat{g}_{t,k} := \nabla_w \widehat{F}_M^{\mathcal{B}^w_{t,k}}(w^{(t)}, \psi_k^{(t)})$ *and assume* $\|\widehat{g}_{t,k}\|_2 \le G_w$ *almost surely. Convexity and independent sampling then give*

$$\mathbb{E}[\widehat{g}_{t,k} \mid \mathcal{G}_{t,k}] \in \partial_w F(w^{(t)}, \psi_k^{(t)}). \tag{29}$$

Use $\psi_k^{(t)}$ *in the dual and outer updates with*

$$\eta_\mu = \frac{\mu_{\max}}{B\sqrt{K}}, \qquad \eta_w = \frac{D_{\mathcal{W}}}{G_w\sqrt{T}}. \tag{30}$$

Then $\overline{w} = T^{-1}\sum_{t=1}^{T} w^{(t)}$ *satisfies*

$$\mathbb{E}[\phi_D(\overline{w})] - \phi_D^* \le \frac{D_{\mathcal{W}} G_w}{\sqrt{T}} + \frac{B\mu_{\max}}{\sqrt{K}} + \overline{\mu}^* \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}. \tag{31}$$

If, in addition, the adversarial family contains a nominal parameter $\hat{\phi}$ *with* $Q^x_{\hat{\phi}} = P_0^x$ *and* $s(\hat{\phi}) \le \rho$, *then*

$$\mathbb{E}[\phi(\overline{w})] - \phi^* \le \underbrace{\frac{D_{\mathcal{W}} G_w}{\sqrt{T}}}_{\text{outer optimization}} + \underbrace{\frac{B\mu_{\max}}{\sqrt{K}}}_{\text{inner optimization}} + \underbrace{\overline{\mu}^* \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}}_{\text{constraint slack in equation 28}} + \underbrace{L_f\sqrt{\rho + \gamma_M}}_{\text{relaxation gap}}. \tag{32}$$

The above guarantees convergence in expected objective value for the relaxed problem as $K, T \to \infty$. For the constrained mini-batch problem, the optimization and constraint slack terms in the bound vanish, leaving only the explicit bound on the possible relaxation gap provided by GDRO's Sinkhorn constraint (Theorem 3.2). Proofs are given in Appendices A.4 and A.5 for Part (i), and in Appendix A.6 for Part (ii).

4 Experiments

Setup and baselines. We evaluate convex contextual newsvendor decisions with explicit conditional generators on synthetic demand and M5 retail data (Makridakis et al., 2022), and nonconvex robot navigation with the pretrained implicit SocialGAN generator (Gupta et al., 2018). All methods share the fitted nominal generator. NOMINAL optimizes over its samples; KL reweights that support; W2 perturbs outcomes directly; and GDRO optimizes a generator-family adversary under a context-local Sinkhorn certificate. We also include GAS-DRO for newsvendor and a true-distribution ORACLE for synthetic demand. Implementation details are in Appendix B.1.

Evaluation. Tables report means $\pm$ sample standard deviations over five optimization sampling seeds, with the nominal generator fixed. Newsvendor metrics are average cost and oracle regret; Q95, CVaR_{95}, stockout, and average order characterize the risk–conservatism trade-off. Navigation metrics cover safety, arrival, minimum distance, and terminal gap. See Appendix B for full details.

4.1 Contextual Newsvendor

Synthetic rare-tail benchmark. Rare contexts may activate a shifted demand component that the fitted conditional VAE may underrepresent. Decisions minimize asymmetric multi-product newsvendor cost; the known data-generating process permits oracle-regret evaluation (Appendix B.2). Table 2 shows that GDRO achieves the lowest non-oracle average cost and regret in both context groups, reducing rare-context regret by 63% relative to Nominal. KL improves some tail metrics but places larger orders and incurs higher average cost.

M5 retail demand. We select 20 products with the fewest zero-demand observations and use a CVAE–LSTM conditioned on the previous 10 days to generate next-day demand. Decisions are evaluated against realized demand on held-out rolling test windows. Risk scores adapt radii and inner budgets (Appendix B.3). Table 3 shows that GDRO has the lowest average cost, with improved Q_{95} relative to Nominal and orders close to W2. GAS-DRO achieves the lowest stockout rate through substantially larger orders, incurring the highest average cost.

Table 2: Synthetic contextual newsvendor results on frequent and rare contexts.

Slice	Method	Avg. cost	Regret	CVaR_{95}	Avg. order
Frequent	Oracle	422.91 ± 0.17	0.00 ± 0.00	1260.18 ± 1.33	25.09 ± 0.00
	Nominal	436.52 ± 0.50	13.62 ± 0.40	1436.14 ± 10.49	24.31 ± 0.04
	KL	449.94 ± 1.11	27.04 ± 1.13	1159.63 ± 16.00	26.20 ± 0.09
	W2	436.55 ± 0.23	13.65 ± 0.17	1312.20 ± 13.02	24.98 ± 0.07
	GAS-DRO	431.24 ± 0.35	8.34 ± 0.23	1356.94 ± 5.61	24.69 ± 0.02
	GDRO	$\mathbf{429.47 \pm 0.18}$	$\mathbf{6.57 \pm 0.04}$	1257.18 ± 1.52	25.16 ± 0.00
Rare	Oracle	515.83 ± 1.15	0.00 ± 0.00	1800.17 ± 2.66	28.39 ± 0.03
	Nominal	540.92 ± 3.64	25.09 ± 3.30	2290.53 ± 49.58	26.42 ± 0.20
	KL	538.31 ± 2.74	22.47 ± 2.77	1911.41 ± 47.26	28.21 ± 0.26
	W2	535.87 ± 3.75	20.04 ± 3.00	2156.93 ± 37.33	26.97 ± 0.11
	GAS-DRO	532.62 ± 2.61	16.79 ± 1.71	2211.65 ± 28.24	26.71 ± 0.08
	GDRO	$\mathbf{525.10 \pm 1.38}$	$\mathbf{9.26 \pm 0.24}$	2075.83 ± 9.09	27.26 ± 0.01

Table 3: M5 contextual newsvendor results on held-out test windows.

Method	Avg. cost	Q_{95}	CVaR_{95}	Stockout	Avg. order
Nominal	1228.98 ± 7.08	2677.72 ± 88.16	3441.57 ± 25.58	0.1101 ± 0.0009	64.60 ± 0.04
KL	1408.33 ± 14.78	2474.82 ± 93.16	3192.02 ± 137.92	0.0878 ± 0.0014	71.63 ± 0.24
W2	1227.78 ± 6.59	2582.77 ± 76.88	3340.52 ± 74.23	0.0999 ± 0.0014	65.65 ± 0.04
GAS-DRO	1730.62 ± 20.17	2587.14 ± 30.17	2945.99 ± 63.97	0.0257 ± 0.0042	86.20 ± 0.63
GDRO	$\mathbf{1211.95 \pm 1.34}$	2433.86 ± 13.23	3174.41 ± 5.07	0.0890 ± 0.0007	66.54 ± 0.02

4.2 Pedestrian-Conditioned Robot Navigation with SocialGAN

Setup. The official pretrained SocialGAN model generates 12 future pedestrian frames from 8 observed frames. The robot chooses a 12-step velocity plan to reach a goal while avoiding pedestrians, using only observed history and generated futures; realized futures are reserved for evaluation. We use 200 fixed tasks from 50 test contexts, with four start–goal queries per context, screened using nominal samples for nontrivial straight-line collision risk and initial clearance. A collision, near miss, and arrival use thresholds of 0.5m, 0.8m, and 0.25m, respectively. Task construction, losses, adaptive radii, and budgets are detailed in Appendix B.4.

Results. Across five seeds, GDRO reduces collisions from 21.2 to 7.2 and near misses from 106.2 to 42.8 relative to nominal decisions, with $199.4/200$ arrivals (Table 4). W2 has a larger mean minimum distance but more collisions, fewer arrivals, and a larger terminal gap. Figure 1 illustrates GDRO's collision avoidance. W2 produces a distorted route after perturbing trajectory support directly. The main comparison uses $M = 8$; Appendix B.4, Table 7 reports sensitivity to $M \in \{8, 16, 32, 64\}$.

Worst-case trajectory structure. On an additional fixed task (Appendix B.4 Figure 3), GDRO retains smooth adversarial pedestrian motion, while W_2 introduces abrupt turns: the 95th-percentile discrete acceleration is 0.053 versus 1.152 (0.037 nominal). Even at the smallest tested radius, maximum W_2 acceleration remains 19.7 times the nominal maximum. The GDRO plan reaches the goal safely, whereas W_2 collides and fails to arrive, illustrating the value of generator structure beyond transport proximity.

Together, these results show that GDRO reduces newsvendor costs and improves navigation safety relative to nominal decisions.

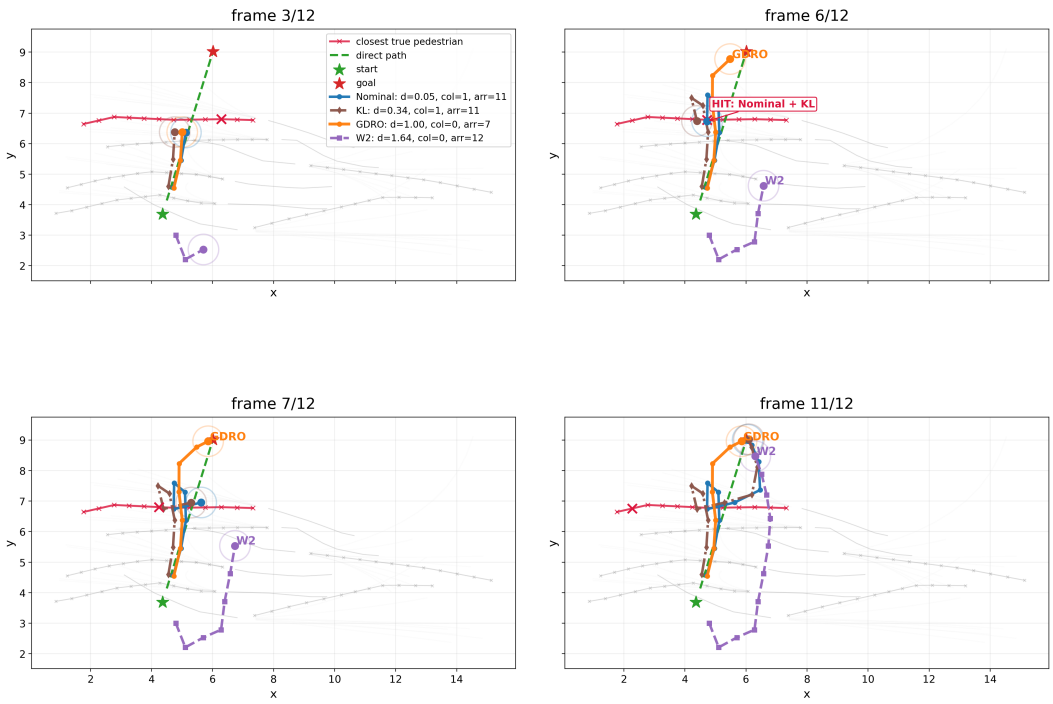

Figure 1: SocialGAN navigation example. Nominal and KL collide with the realized pedestrian future; GDRO avoids collision and reaches the goal, while W2 produces a distorted route.

Table 4: SocialGAN navigation on 200 fixed test tasks. GDRO uses $M = 8$ per Sinkhorn evaluation.

Method	Collisions	Near misses	Arrivals	Min dist.	Terminal gap
Nominal	21.2 ± 1.30	106.2 ± 2.49	200.0 ± 0.00	0.792 ± 0.012	0.011 ± 0.001
KL	16.0 ± 3.54	92.8 ± 3.56	199.8 ± 0.45	0.833 ± 0.018	0.019 ± 0.003
GDRO	$\mathbf{7.2 \pm 1.64}$	$\mathbf{42.8 \pm 3.83}$	199.4 ± 0.89	0.950 ± 0.010	0.042 ± 0.003
W2	27.2 ± 3.27	56.8 ± 6.14	136.8 ± 4.97	1.093 ± 0.023	0.286 ± 0.025

5 Conclusion

We proposed GDRO, a context-local distributionally robust optimization framework for decision-making with learned conditional generators. GDRO restricts adversarial laws to a chosen generator family and controls their proximity to a frozen nominal sampler through a sample-based Sinkhorn certificate. We developed an expected mini-batch formulation and established downstream risk bounds, relaxation-gap, and finite-time optimization guarantees.

Experiments with explicit conditional generators and an implicit SocialGAN generator show improved inventory decisions and safer robot navigation. Inspection of adversarial pedestrian trajectories illustrates the benefit of generator structure through smoother motion and a safer robot decision. The generator-family restriction provides useful structural inductive bias, although it does not by itself guarantee preservation of the underlying data manifold. Establishing formal conditions for such preservation is an interesting direction for future work.

AI use statement

During the preparation of this submission, the authors used a generative AI tool to assist with literature search and citation checking, LaTeX and language editing, verification of mathematical derivations and proofs developed by the authors, and code-level optimization suggestions. The core re-

search idea, GDRO formulation, and principal implementation were developed by the authors. All AI-assisted suggestions were independently reviewed, checked, and verified by the authors. The authors take full responsibility for the final content of this work, including its claims, proofs, experiments, and code.

Reproducibility statement

The population and empirical formulations, optimization procedure, and assumptions are stated in the main text. Complete proofs are provided in Appendix A. Experimental details, including baseline implementations, data construction and preprocessing, model architectures, hyperparameters, ambiguity radii, optimization budgets, fixed task sets, and evaluation metrics, are provided in Appendix B. The complete anonymized code, configurations, fixed task sets, and released evaluation artifacts are available at `https://github.com/reproducibility-research/gdro-artifacts`.

References

Jose Blanchet and Karthyek Murthy. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research*, 44(2):565–600, 2019. doi: 10.1287/moor.2018.0936.

Giorgio Costa and Garud N. Iyengar. Distributionally robust end-to-end portfolio construction. *Quantitative Finance*, 23(10):1465–1482, 2023. doi: 10.1080/14697688.2023.2236148.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, volume 26, 2013. URL `https://papers.nips.cc/paper_files/paper/2013/hash/af21d0c97db2e27e13572cbf59eb343d-Abstract.html`.

Priya L. Donti, Brandon Amos, and J. Zico Kolter. Task-based end-to-end model learning in stochastic optimization. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL `https://papers.nips.cc/paper/2017/hash/3fc2c60b5782f641f76bcefc39fb2392-Abstract.html`.

Adam N. Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68 (1):9–26, 2022. doi: 10.1287/mnsc.2020.3922.

Kilian Fatras, Younes Zine, Rémi Flamary, Rémi Gribonval, and Nicolas Courty. Learning with minibatch Wasserstein: Asymptotic and gradient properties. In *Proceedings of the Twenty-Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 2131–2141. PMLR, 2020. URL `https://proceedings.mlr.press/v108/fatras20a.html`.

Jean Feydy, Thibault Séjourné, François-Xavier Vialard, Shun-ichi Amari, Alain Trouvé, and Gabriel Peyré. Interpolating between optimal transport and MMD using sinkhorn divergences. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 2681–2690. PMLR, 2019. URL `https://proceedings.mlr.press/v89/feydy19a.html`.

Rui Gao and Anton J. Kleywegt. Distributionally robust stochastic optimization with wasserstein distance. *Mathematics of Operations Research*, 48(2):603–655, 2023. doi: 10.1287/moor.2022.1275.

Rui Gao, Xi Chen, and Anton J. Kleywegt. Wasserstein distributionally robust optimization and variation regularization. *Operations Research*, 72(3):1177–1191, 2024. doi: 10.1287/opre.2022.2383.

Aude Genevay, Gabriel Peyré, and Marco Cuturi. Learning generative models with Sinkhorn divergences. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pp. 1608–1617. PMLR, 2018. URL `https://proceedings.mlr.press/v84/genevay18a.html`.

Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social GAN: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2255–2264, 2018. doi: 10.1109/CVPR.2018.00240.

Zhaolin Hu and L. Jeff Hong. Kullback–Leibler divergence constrained distributionally robust optimization. Optimization Online, 2012. URL https://optimization-online.org/2012/11/3677/. Preprint.

Ivan Kobyzev, Simon J. D. Prince, and Marcus A. Brubaker. Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11):3964–3979, 2021. doi: 10.1109/TPAMI.2020.2992934.

Xutao Ma, Chao Ning, and Wenli Du. Differentiable distributionally robust optimization layers. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 33880–33901. PMLR, 2024. URL https://proceedings.mlr.press/v235/ma24j.html.

Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: Online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(81):2503–2528, 2012. URL http://jmlr.org/papers/v13/mahdavi12a.html.

Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4):1346–1364, 2022. doi: 10.1016/j.ijforecast.2021.11.013. URL https://doi.org/10.1016/j.ijforecast.2021.11.013.

Paul Michel, Tatsunori Hashimoto, and Graham Neubig. Modeling the second player in distributionally robust optimization. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=ZDnzZrTqU9N.

Paul Michel, Tatsunori Hashimoto, and Graham Neubig. Distributionally robust models with parametric likelihood ratios. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=a34GrNaYEcS.

Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1–2):115–166, 2018. doi: 10.1007/s10107-017-1172-1.

Hongseok Namkoong and John C. Duchi. Stochastic gradient methods for distributionally robust optimization with f-divergences. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Angelia Nedić and Asuman Ozdaglar. Subgradient methods for saddle-point problems. *Journal of Optimization Theory and Applications*, 142(1):205–228, 2009.

Gabriel Peyré and Marco Cuturi. Computational optimal transport with applications to data sciences. *Foundations and Trends in Machine Learning*, 11(5–6):355–607, 2019. doi: 10.1561/2200000073. URL https://doi.org/10.1561/2200000073.

Hamed Rahimian and Sanjay Mehrotra. Frameworks and results in distributionally robust optimization. *Open Journal of Mathematical Optimization*, 3:4, 2022. doi: 10.5802/ojmo.15. URL https://ojmo.centre-mersenne.org/articles/10.5802/ojmo.15/.

Allen Z. Ren and Anirudha Majumdar. Distributionally robust policy learning via adversarial environment generation. *IEEE Robotics and Automation Letters*, 7(2):1379–1386, 2022. doi: 10.1109/LRA.2021.3139949.

Aman Sinha, Hongseok Namkoong, and John Duchi. Certifying some distributional robustness with principled adversarial training. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=Hk6kPgZA-.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*, volume 32, pp. 11895–11907, 2019.

Irina Wang, Bart Van Parys, and Bartolomeo Stellato. Learning decision-focused uncertainty sets in robust optimization. arXiv preprint arXiv:2305.19225v5, 2025. URL https://arxiv.org/abs/2305.19225v5.

Jie Wang, Rui Gao, and Yao Xie. Sinkhorn distributionally robust optimization. *Operations Research*, 74(3):1581–1603, 2026. doi: 10.1287/opre.2023.0294. URL https://doi.org/10.1287/opre.2023.0294.

Jiaqi Wen and Jianyi Yang. Distributionally robust optimization via diffusion ambiguity modeling, 2025. URL https://arxiv.org/abs/2510.22757.

Jiaqi Wen and Jianyi Yang. Distributionally robust optimization via generative ambiguity modeling. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=q67t0gFrdY.

Chen Xu, Jonghyeok Lee, Xiuyuan Cheng, and Yao Xie. Flow-based distributionally robust optimization. *IEEE Journal on Selected Areas in Information Theory*, 5:62–77, 2024. doi: 10.1109/JSAIT.2024.3370699. URL https://doi.org/10.1109/JSAIT.2024.3370699.

Yufeng Yang, Yi Zhou, and Zhaosong Lu. Nested stochastic algorithm for generalized Sinkhorn distance-regularized distributionally robust optimization. arXiv preprint arXiv:2503.22923v2, 2025. URL https://arxiv.org/abs/2503.22923v2.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML)*, 2003.

A PROOFS

This appendix contains full proofs for the results stated in Section 3. The population proof below concerns P_0^x and Q_ψ^x at the fixed context x. All optimization proofs use the expected mini-batch quantities introduced in Section 2.3, with the batch size M and coupling κ fixed:

$$F(w,\psi) := \mathbb{E}_{\mathcal{B}}[\widehat{F}_M^{\mathcal{B}}(w,\psi)], \qquad s(\psi) := \mathbb{E}_{\mathcal{B}}[\widehat{s}_M^{\mathcal{B}}(\psi)].$$

We write $\mathcal{A} := \{\psi : s(\psi) \le \rho\}$, $\phi(w) := \max_{\psi\in\mathcal{A}} F(w,\psi)$, $\phi_D(w) := \inf_{\mu\ge 0} Q_w(\mu)$, and retain $\gamma_M := \varepsilon \log M$ and $B := \max\{\rho, 4R^2\}$.

A.1 MINI-BATCH SINKHORN PROPERTIES

Lemma A.1 (Uniform Sinkhorn upper bound)**.** *Under* (A3)*, any two nominal or adversarial output laws P, Q satisfy*

$$0 \le S_\varepsilon(P,Q) \le 4R^2. \tag{33}$$

For each fixed $\psi \in \Psi$, the same bound holds almost surely for its empirical cloud. Consequently,

$$0 \le \widehat{s}_M^{\mathcal{B}}(\psi) \le 4R^2 \quad \textit{almost surely}, \qquad 0 \le s(\psi) \le 4R^2.$$

Proof. The independent coupling $P \otimes Q$ is feasible in equation 2 and has zero KL penalty. Since both laws are supported in the radius-R ball,

$$\|y - y'\|_2^2 \le (\|y\|_2 + \|y'\|_2)^2 \le 4R^2.$$

Consequently, $\mathrm{OT}_\varepsilon(P,Q) \le 4R^2$. The two self-transport terms are nonnegative, so $S_\varepsilon(P,Q) \le \mathrm{OT}_\varepsilon(P,Q) \le 4R^2$. Nonnegativity follows from the positivity of the debiased Sinkhorn divergence for squared-Euclidean cost (Feydy et al., 2019). For each fixed ψ, the empirical measures have the same support bound almost surely. Taking expectations gives $0 \le s(\psi) \le 4R^2$. For an adaptively selected ψ_k, the same conclusion holds conditionally when its certificate batch is drawn independently after ψ_k is fixed. □

Discrete convention. For equally weighted empirical measures $\alpha = M^{-1}\sum_i \delta_{a_i}$ and $\beta = M^{-1}\sum_j \delta_{b_j}$, define

$$\Pi_M := \left\{\pi \in \mathbb{R}_+^{M\times M} : \pi \mathbf{1}_M = M^{-1}\mathbf{1}_M,\ \pi^\top \mathbf{1}_M = M^{-1}\mathbf{1}_M\right\}.$$

Then

$$\mathrm{OT}_\varepsilon(\alpha,\beta) = \min_{\pi\in\Pi_M}\left\{\sum_{i,j=1}^M \pi_{ij}c(a_i,b_j) + \varepsilon\sum_{i,j=1}^M \pi_{ij}\log(M^2\pi_{ij})\right\}, \tag{34}$$

with $0\log 0 = 0$. The matrices are indexed by sample draws, including when sample locations coincide. Aggregating coincident locations cannot increase relative entropy; conversely, a coupling on distinct locations can be lifted by distributing its mass uniformly over the corresponding index pairs, preserving cost and relative entropy. Thus the indexed formulation has the same optimal value. Replacing the relative entropy by the corresponding negative-entropy convention changes every transport term by the same additive constant, which cancels in the debiased Sinkhorn divergence.

Lemma A.2 (Empirical Sinkhorn–Wasserstein comparison)**.** *For any two equally weighted M-point empirical laws α, β,*

$$\left|S_\varepsilon(\alpha,\beta) - W_2^2(\alpha,\beta)\right| \le \gamma_M := \varepsilon\log M. \tag{35}$$

Proof. Let $\alpha = M^{-1}\sum_i \delta_{a_i}$ and $\beta = M^{-1}\sum_j \delta_{b_j}$. The KL term in equation 34 is nonnegative, hence $\mathrm{OT}_\varepsilon(\alpha,\beta) \ge W_2^2(\alpha,\beta)$. The unregularized problem has an optimal permutation coupling. Its M nonzero entries equal $1/M$, so its KL divergence from the product weights $1/M^2$ is $\log M$. Therefore

$$0 \le \mathrm{OT}_\varepsilon(\alpha,\beta) - W_2^2(\alpha,\beta) \le \varepsilon\log M.$$

Applying the same argument to the two self-transport terms shows that $S_\varepsilon(\alpha,\beta) - W_2^2(\alpha,\beta)$ is one number in $[0,\gamma_M]$ minus half of each of two numbers in $[0,\gamma_M]$. It therefore lies in $[-\gamma_M,\gamma_M]$, proving equation 35. □

Lemma A.3 (Coupled mini-batch oracle)**.** *Let $\mathcal{F}$ be a sigma-field, let ψ be $\mathcal{F}$-measurable, and draw $\mathcal{B}\sim\kappa^{\otimes M}$ independently of $\mathcal{F}$. Then*

$$\mathbb{E}[\widehat{s}_M^{\mathcal{B}}(\psi)\mid\mathcal{F}] = s(\psi). \tag{36}$$

Likewise,

$$\mathbb{E}[\widehat{F}_M^{\mathcal{B}}(w,\psi)\mid\mathcal{F}] = F(w,\psi) \tag{37}$$

whenever (w,ψ) is $\mathcal{F}$-measurable and $\widehat{F}_M^{\mathcal{B}}(w,\psi)$ is integrable. If, in addition, $\widehat{s}_M^{\mathcal{B}}$ is almost surely differentiable at ψ and its gradient admits an integrable local envelope permitting differentiation under the expectation, then

$$\mathbb{E}[\nabla_\psi \widehat{s}_M^{\mathcal{B}}(\psi)\mid\mathcal{F}] = \nabla_\psi s(\psi). \tag{38}$$

Proof. Conditional on $\mathcal{F}$, the value of ψ is fixed and the law of the fresh batch remains $\kappa^{\otimes M}$. Therefore

$$\mathbb{E}[\widehat{s}_M^{\mathcal{B}}(\psi)\mid\mathcal{F}] = \mathbb{E}_{\mathcal{B}\sim\kappa^{\otimes M}}[\widehat{s}_M^{\mathcal{B}}(\psi)] = s(\psi),$$

which is equation 36. The proof of equation 37 is identical and also uses the fact that the adversarial coordinate of κ has marginal ζ_{A}. Finally, equation 38 follows by differentiation under the integral. For the product coupling, this is the expected mini-batch differentiation setting analyzed by Fatras et al. (2020). For a general fixed coupling, the same argument applies to the paired random element $(Z_0, Z_{\mathrm{A}}) \sim \kappa$ under the stated envelope condition. □

Corollary A.4 (Exact centering at the nominal generator)**.** *Suppose $\zeta_0 = \zeta_{\mathrm{A}}$ and the adversarial family contains the nominal sampler $G_{\hat\phi}$. In each new batch, use the same latent samples for the nominal and adversarial generators. Then*

$$\widehat{s}_M^{\mathcal{B}}(\hat\phi) = 0 \quad \textit{almost surely}, \qquad s(\hat\phi) = 0.$$

Thus $\hat\phi$ is strictly feasible for every $\rho > 0$.

Proof. When $\psi = \hat{\phi}$, both generators use the same latent samples and therefore produce identical output samples in each batch. The debiased Sinkhorn divergence of a measure from itself is zero. Taking expectations gives $s(\hat{\phi}) = 0$. □

Lemma A.5 (Pointwise consistency of the certificate)**.** *Under* (A3)*, for every fixed* $\psi \in \Psi$ *and every fixed* $\kappa \in \Pi(\zeta_0, \zeta_{\mathrm{A}})$*,*

$$s_{M,\kappa}(\psi) \longrightarrow S_\varepsilon(Q_\psi^x, P_0^x) \qquad \text{as } M \to \infty. \tag{39}$$

Proof. Fix ψ and κ. Although the two coordinates of a pair drawn from κ may be dependent, the sequence of nominal coordinates is i.i.d. with law ζ_0, and the sequence of adversarial coordinates is i.i.d. with law ζ_{A}. Hence the two empirical output measures in equation 7 converge almost surely, weakly, to P_0^x and Q_ψ^x, respectively. Continuity of S_ε for weak convergence on a common compact support (Feydy et al., 2019) gives

$$\widehat{s}_M^{\mathcal{B}}(\psi) \longrightarrow S_\varepsilon(Q_\psi^x, P_0^x) \quad \text{almost surely.}$$

Lemma A.1 supplies the uniform bound $0 \le \widehat{s}_M^{\mathcal{B}}(\psi) \le 4R^2$, so dominated convergence proves equation 39. The result is pointwise in ψ; no uniform convergence over the generator family is claimed. □

Lemma A.6 (Concentration over fresh batches)**.** *Fix* $\psi \in \Psi$ *and draw independent batches* $\mathcal{B}_1, \dots, \mathcal{B}_L \sim \kappa^{\otimes M}$*. Let*

$$\overline{s}_L(\psi) := \frac{1}{L} \sum_{\ell=1}^{L} \widehat{s}_M^{\mathcal{B}_\ell}(\psi).$$

For $R > 0$ *and every* $u > 0$*,*

$$\mathbb{P}(|\overline{s}_L(\psi) - s(\psi)| \ge u) \le 2 \exp\left(-\frac{Lu^2}{8R^4}\right). \tag{40}$$

If $R = 0$*, both certificates vanish, and the probability is zero.*

Proof. The variables $\widehat{s}_M^{\mathcal{B}_\ell}(\psi)$ are independent, have common expectation $s(\psi)$, and lie in $[0, 4R^2]$ by Lemma A.1. For $R > 0$, Hoeffding's inequality gives equation 40; for $R = 0$, the variables are identically zero. □

A.2 Proof of Theorem 3.1

Proof. The radius-R ball is compact, so $\mathcal{P}_R$ is compact under weak convergence. On this space, S_ε is continuous, positive definite, and metrizes weak convergence (Feydy et al., 2019, Theorem 1); W_1 is also continuous. Since $W_1(P, Q) \le 2R$ on $\mathcal{P}_R$, the modulus $\omega_{\varepsilon,R}$ is finite, and it is nondecreasing by its definition.

If $\omega_{\varepsilon,R}(r)$ did not tend to zero, there would be $r_n \downarrow 0$ and $P_n, Q_n \in \mathcal{P}_R$ with $S_\varepsilon(P_n, Q_n) \le r_n$ but $W_1(P_n, Q_n)$ bounded away from zero. Compactness gives a subsequence converging to some P, Q. Continuity gives $S_\varepsilon(P, Q) = 0$, hence $P = Q$, while continuity of W_1 gives $W_1(P_n, Q_n) \to 0$, a contradiction. Thus $\omega_{\varepsilon,R}(r) \downarrow 0$.

For a population-feasible ψ, definition equation 19 gives $W_1(Q_\psi^x, P_0^x) \le \omega_{\varepsilon,R}(\rho)$.

Assumption (A1) and Kantorovich–Rubinstein duality now yield

$$\left| \mathbb{E}_{Q_\psi^x}[f(w, Y)] - \mathbb{E}_{P_0^x}[f(w, Y)] \right| \le L_f W_1(Q_\psi^x, P_0^x) \le L_f \omega_{\varepsilon,R}(\rho).$$

For the mini-batch statement, let $\alpha = \widehat{P}^{\mathcal{B}}_{G_\psi, x}$ and $\beta = \widehat{P}^{\mathcal{B}}_{G_{\hat{\phi}}, x}$. Assumption (A1), Kantorovich–Rubinstein duality, $W_1 \le W_2$, and equation 35 give

$$\left| \int f(w, y)\, \mathrm{d}\alpha(y) - \int f(w, y)\, \mathrm{d}\beta(y) \right| \le L_f W_1(\alpha, \beta) \le L_f \sqrt{\widehat{s}_M^{\mathcal{B}}(\psi) + \gamma_M}.$$

The expectations of the two empirical integrals are $F(w,\psi)$ and $F_0(w)$, respectively, since κ has marginals ζ_0 and ζ_{A}. Taking expectations and using Jensen's inequality,

$$\begin{aligned}|F(w,\psi)-F_0(w)| &\le \mathbb{E}_{\mathcal{B}}\left[\left|\int f(w,y)\,\mathrm{d}\alpha(y)-\int f(w,y)\,\mathrm{d}\beta(y)\right|\right]\\ &\le L_f\mathbb{E}_{\mathcal{B}}[\sqrt{\widehat{s}_M^{\mathcal{B}}(\psi)+\gamma_M}]\\ &\le L_f\sqrt{s(\psi)+\gamma_M}.\end{aligned}$$

For $\psi\in\mathcal{A}$, substituting $s(\psi)\le\rho$ gives $|F(w,\psi)-F_0(w)|\le L_f\sqrt{\rho+\gamma_M}$. □

A.3 Lagrangian Properties and Relaxation Gap

Convexity and finiteness. Suppose (A1)–(A3) hold, $f(w,y)$ is convex in w for every y, and $\mathcal{A}\neq\emptyset$. For each fixed ψ, $F(w,\psi)+\mu(\rho-s(\psi))$ is jointly convex in (w,μ), since $F(\cdot,\psi)$ is convex and $\rho-s(\psi)$ is independent of w. Its supremum over ψ therefore makes Q jointly convex.

To establish finiteness, fix $\psi_f\in\mathcal{A}$. The Lipschitz assumption and the common support diameter $2R$ give $|F(w,\psi)-F(w,\psi_f)|\le 2L_fR$ for every ψ. Consequently,

$$Q_w(0)=\sup_{\psi}F(w,\psi)\le F(w,\psi_f)+2L_fR<\infty.$$

Moreover, feasibility gives $\rho-s(\psi_f)\ge 0$, so for every $\mu\ge 0$,

$$Q_w(\mu)\ge F(w,\psi_f)+\mu(\rho-s(\psi_f))\ge F(w,\psi_f).$$

Taking the infimum over $\mu\ge 0$ thus yields $F(w,\psi_f)\le\phi_D(w)\le Q_w(0)<\infty$. Finally, partial minimization of the jointly convex function Q over $\mu\ge 0$ preserves convexity, so ϕ_D is convex and real-valued.

Strict feasibility and dual regularity. Under (A1) and (A3), suppose there is a fixed $\psi_0\in\Psi$ with $d_0:=\rho-s(\psi_0)>0$. For every w and $\mu\ge 0$,

$$Q_w(\mu)\ge F(w,\psi_0)+\mu d_0.$$

Hence $Q_w(\mu)\to\infty$ as $\mu\to\infty$. Since $Q_w(0)<\infty$ and Q_w is lower semicontinuous as a supremum of continuous affine functions of μ, it attains its minimum on $[0,\infty)$.

Choose a minimizer $\mu^*(w)$. Comparing its value with the candidate $\mu=0$ and using the same Lipschitz bound gives

$$\begin{aligned}F(w,\psi_0)+\mu^*(w)d_0 &\le Q_w(\mu^*(w))\le Q_w(0)\\ &\le F(w,\psi_0)+2L_fR.\end{aligned}$$

Subtracting $F(w,\psi_0)$ and dividing by $d_0>0$ bounds $\mu^*(w)$ uniformly in w. Thus Assumption 3.1 holds with

$$\overline{\mu}^*\le\frac{2L_fR}{\rho-s(\psi_0)}. \tag{41}$$

Proof of Theorem 3.2. Fix $w\in\mathcal{W}$ and recall that $C_{\rho,M}:=L_f\sqrt{\rho+\gamma_M}$. By Theorem 3.1, every $\psi\in\Psi$ satisfies

$$F(w,\psi)\le F_0(w)+L_f\sqrt{s(\psi)+\gamma_M}. \tag{42}$$

The function $g(u):=\sqrt{u+\gamma_M}$ is concave on $[0,\infty)$ and differentiable at $\rho>0$, with $g'(\rho)=1/(2\sqrt{\rho+\gamma_M})$. Its tangent at ρ therefore provides the global upper bound

$$\sqrt{u+\gamma_M}\le\sqrt{\rho+\gamma_M}+\frac{u-\rho}{2\sqrt{\rho+\gamma_M}},\qquad u\ge 0. \tag{43}$$

Choose the nonnegative multiplier

$$\mu_\rho:=L_fg'(\rho)=\frac{L_f}{2\sqrt{\rho+\gamma_M}}.$$

Substituting $u = s(\psi)$ into equation 43 and combining with equation 42 gives

$$F(w, \psi) \le F_0(w) + C_{\rho,M} + \mu_\rho\big(s(\psi) - \rho\big).$$

Adding $\mu_\rho(\rho - s(\psi))$ to both sides cancels the last term, yielding

$$F(w, \psi) + \mu_\rho(\rho - s(\psi)) \le F_0(w) + C_{\rho,M}.$$

This inequality holds for every $\psi \in \Psi$, so taking the supremum over ψ bounds $Q_w(\mu_\rho)$. Moreover, $\mu_\rho \ge 0$ is an admissible candidate in the infimum defining $\phi_D(w)$. Hence

$$\phi_D(w) = \inf_{\mu \ge 0} Q_w(\mu) \le Q_w(\mu_\rho) \le F_0(w) + C_{\rho,M}. \tag{44}$$

By the feasible-nominal assumption, $\hat{\phi} \in \mathcal{A}$ and $Q^x_{\hat{\phi}} = P^x_0$. Thus

$$\phi(w) \ge F(w, \hat{\phi}) = F_0(w).$$

Combining this lower bound with weak duality $\phi(w) \le \phi_D(w)$ and equation 44 gives

$$0 \le \phi_D(w) - \phi(w) \le F_0(w) + C_{\rho,M} - \phi(w) \le C_{\rho,M},$$

which proves equation 24.

Since $C_{\rho,M}$ is independent of w, the pointwise inequalities $\phi(w) \le \phi_D(w) \le \phi(w) + C_{\rho,M}$ imply

$$\begin{aligned} \phi^* &\le \phi^*_D \\ &\le \inf_{w \in \mathcal{W}} \big\{\phi(w) + C_{\rho,M}\big\} = \phi^* + C_{\rho,M}. \end{aligned}$$

If $\phi^* > -\infty$, both optimal values are finite. Subtracting ϕ^* therefore yields $0 \le \phi^*_D - \phi^* \le C_{\rho,M}$, proving equation 25. □

A.4 Inner Convergence

For a fixed outer decision w, let $\mathcal{F}_k$ contain the history used to construct (ψ_k, μ_k) before a fresh certificate batch is drawn. Set

$$s_k := s(\psi_k), \qquad \widehat{s}_k := \widehat{s}^{\mathcal{B}^s_k}_M(\psi_k), \qquad b_k := \rho - s_k, \qquad \widehat{b}_k := \rho - \widehat{s}_k.$$

The fresh-batch construction gives

$$\mathbb{E}[\widehat{b}_k \mid \mathcal{F}_k] = b_k. \tag{45}$$

For an outer block t, take $w = w^{(t)}$ and write $\mu_k := \mu_{t,k}$, $\psi_k := \psi^{(t)}_k$, and $\mu_1 := \mu_{t,1} \in [0, \mu_{\max}]$. The bounds below hold for every initialization in this interval. They can therefore be applied conditionally on the history $\mathcal{H}_t$ at the start of a block, when $w^{(t)}$ and $\mu_{t,1}$ are fixed, and then averaged over that history.

Lemma A.7 (Expected inner-relaxation bound)**.** *Under* (A1)*,* (A3)*, equation 45, and equation 26, run the inner loop for K dual steps with $\eta_\mu = \mu_{\max}/(B\sqrt{K})$. Then*

$$\phi_D(w) - \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}[F(w, \psi_k)] \le \frac{B\mu_{\max}}{\sqrt{K}}. \tag{46}$$

The same bound holds with $\phi_D(w)$ replaced by $\phi(w)$.

Proof. The inner loop is projected primal–dual ascent on the Lagrangian $\mathcal{L}(w, \psi, \mu) = F(w, \psi) - \mu(s(\psi) - \rho)$: at each step k, ψ_k maximizes $\mathcal{L}(w, \cdot, \mu_k)$ and the dual variable is updated by projected descent,

$$\mu_{k+1} = \Pi_{[0,\mu_{\max}]}(\mu_k - \eta_\mu \widehat{b}_k), \qquad b_k := \rho - s(\psi_k), \qquad \widehat{b}_k := \rho - \widehat{s}_k.$$

We use the standard dual-subgradient analysis for constrained saddle-point problems (Zinkevich, 2003; Nedić & Ozdaglar, 2009).

Dual trajectory. Nonexpansiveness of projection gives, for any reference $\mu \in [0, \mu_{\max}]$,

$$|\mu_{k+1} - \mu|^2 \le |\mu_k - \mu|^2 - 2\eta_\mu(\mu_k - \mu)\widehat{b}_k + \eta_\mu^2 |\widehat{b}_k|^2.$$

Rearranging and summing over $k = 1, \dots, K$ gives

$$\frac{1}{K}\sum_{k=1}^{K}(\mu_k - \mu)\widehat{b}_k \le \frac{|\mu_1 - \mu|^2}{2K\eta_\mu} + \frac{\eta_\mu}{2K}\sum_{k=1}^{K}|\widehat{b}_k|^2. \tag{47}$$

Because the certificate batch is drawn after (μ_k, ψ_k) is fixed, $\mathbb{E}[\widehat{b}_k \mid \mathcal{F}_k] = b_k$, and hence

$$\begin{aligned}\mathbb{E}[(\mu_k - \mu)\widehat{b}_k] &= \mathbb{E}[(\mu_k - \mu)\mathbb{E}[\widehat{b}_k \mid \mathcal{F}_k]] \\ &= \mathbb{E}[(\mu_k - \mu)b_k].\end{aligned} \tag{48}$$

By Lemma A.1, $0 \le \widehat{s}_k \le 4R^2$, so $|\widehat{b}_k| \le B$. Taking expectations in equation 47 and using equation 48 yields

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[(\mu_k - \mu)b_k] \le \frac{\mathbb{E}[|\mu_1 - \mu|^2]}{2K\eta_\mu} + \frac{\eta_\mu B^2}{2}. \tag{49}$$

Taking $\mu = 0$, using $\mu_1 \le \mu_{\max}$, and substituting $\eta_\mu = \mu_{\max}/(B\sqrt{K})$ gives

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[\mu_k b_k] \le \frac{B\mu_{\max}}{\sqrt{K}}. \tag{50}$$

Duality gap. For every $\mu_k \ge 0$, weak duality gives

$$Q_w(\mu_k) = \sup_{\psi}\{F(w,\psi) + \mu_k(\rho - s(\psi))\} \ge \max_{\psi \in \mathcal{A}} F(w,\psi) = \phi(w),$$

since any feasible $\psi \in \mathcal{A}$ satisfies $s(\psi) \le \rho$ and hence $\mu_k(\rho - s(\psi)) \ge 0$. Moreover, $Q_w(\mu_k) \ge \phi_D(w) = \inf_{\mu \ge 0} Q_w(\mu)$. As ψ_k maximizes the expected Lagrangian at μ_k,

$$F(w, \psi_k) + \mu_k b_k = Q_w(\mu_k) \ge \phi_D(w).$$

Averaging over $k = 1, \dots, K$, taking expectations, and rearranging,

$$\phi_D(w) - \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[F(w,\psi_k)] \le \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[\mu_k b_k] \le \frac{B\mu_{\max}}{\sqrt{K}},$$

where the last step is equation 50. Since $\phi(w) \le \phi_D(w)$, the constrained version follows as well. An approximate expected oracle contributes its average Lagrangian error additively to this inner bound. □

A.5 Proof of Lemma A.8

Lemma A.8 (Expected average constraint slack)**.** *Under the conditions of Lemma A.7 and Assumption 3.1, assume $\mu_{\max} > \overline{\mu}^*$. Then*

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[s(\psi_k)] \le \rho + \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}. \tag{51}$$

The argument below follows the standard long-term-constraint analysis of Mahdavi et al. (2012). The comparison-multiplier construction is likewise standard and is used, for example, in the GAS-DRO analysis of Wen & Yang (2026). We include the relevant details here to keep the Sinkhorn specialization self-contained.

Proof. Fix the outer decision w and omit it from the notation. Write $s_k := s(\psi_k)$ and define

$$b_k := \rho - s_k, \qquad v_k := s_k - \rho = -b_k.$$

By Lemma A.1, $0 \le \widehat{s}_k \le 4R^2$, and hence $|\widehat{b}_k| \le B := \max\{\rho, 4R^2\}$.

We first upper-bound the projected-dual regret term. Applying equation 49 with reference multiplier $\mu = \mu_{\max}$ gives

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[(\mu_k - \mu_{\max})b_k] \le \frac{\mathbb{E}[|\mu_{\max} - \mu_1|^2]}{2K\eta_\mu} + \frac{\eta_\mu B^2}{2}.$$

Using $\mu_1 \in [0, \mu_{\max}]$ and $\eta_\mu = \mu_{\max}/(B\sqrt{K})$ yields

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[(\mu_k - \mu_{\max})b_k] \le \frac{B\mu_{\max}}{\sqrt{K}}. \tag{52}$$

We now lower-bound the same term by the average constraint violation. Recall the dual function

$$Q_w(\mu) := \sup_{\psi\in\Psi}\{F(w,\psi) + \mu(\rho - s(\psi))\}.$$

By the exact inner oracle condition equation 26,

$$Q_w(\mu_k) = F(w,\psi_k) + \mu_k b_k. \tag{53}$$

Let $\mu^*(w)$ be the selected minimizer of Q_w from Assumption 3.1. Since $Q_w(\mu_k) \ge Q_w(\mu^*(w))$, equation 53 implies

$$\mu_k b_k = Q_w(\mu_k) - F(w,\psi_k) \ge Q_w(\mu^*(w)) - F(w,\psi_k). \tag{54}$$

Moreover, evaluating the supremum defining $Q_w(\mu^*(w))$ at the candidate ψ_k gives

$$F(w,\psi_k) + \mu^*(w)b_k \le Q_w(\mu^*(w)). \tag{55}$$

Since $b_k = -v_k$, equation 55 is equivalent to

$$F(w,\psi_k) \le Q_w(\mu^*(w)) + \mu^*(w)v_k. \tag{56}$$

Substituting equation 56 into equation 54 gives

$$\mu_k b_k \ge Q_w(\mu^*(w)) - \big[Q_w(\mu^*(w)) + \mu^*(w)v_k\big] = -\mu^*(w)v_k.$$

Therefore, using $b_k = -v_k$,

$$\begin{aligned}
\sum_{k=1}^{K}(\mu_k - \mu_{\max})b_k &= \sum_{k=1}^{K}\mu_k b_k - \mu_{\max}\sum_{k=1}^{K}b_k \\
&= \sum_{k=1}^{K}\mu_k b_k + \mu_{\max}\sum_{k=1}^{K}v_k \\
&\ge -\mu^*(w)\sum_{k=1}^{K}v_k + \mu_{\max}\sum_{k=1}^{K}v_k \\
&= \big(\mu_{\max} - \mu^*(w)\big)\sum_{k=1}^{K}v_k.
\end{aligned} \tag{57}$$

Taking expectations in equation 57, combining with equation 52, and dividing by $\mu_{\max} - \mu^*(w) > 0$, gives

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[v_k] \le \frac{B\mu_{\max}}{(\mu_{\max} - \mu^*(w))\sqrt{K}}.$$

By Assumption 3.1, $\mu^*(w) \le \overline{\mu}^*$, and $\mu_{\max} > \overline{\mu}^*$. Hence

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[v_k] \le \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}.$$

Finally, since $v_k = s_k - \rho$,

$$\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[s_k] = \rho + \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[v_k] \le \rho + \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}.$$

This proves the claim. □

A.6 Proof of Theorem 3.3

Proof. Part (i) follows from Lemmas A.7 and A.8. For Part (ii), fix a deterministic comparator $w \in \mathcal{W}$. Under the convexity condition in Part (ii), $F(\cdot, \psi)$ is convex on $\mathcal{W}$ for every ψ. Moreover, ϕ_D is convex by joint convexity of Q and partial minimization over $\mu \geq 0$, as established in Appendix A.3.

We first show that the averaged inner gradient $\hat{g}^{(t)} = K^{-1} \sum_{k=1}^{K} \widehat{g}_{t,k}$ satisfies the expected approximate subgradient inequality

$$\begin{aligned} \mathbb{E}[\phi_D(w^{(t)})] - \phi_D(w) \leq \mathbb{E}[\langle \hat{g}^{(t)}, w^{(t)} - w \rangle] \\ + \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B \mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}. \end{aligned} \tag{58}$$

For each k, convexity of the sample objective gives

$$\widehat{F}_M^{\mathcal{B}_{t,k}^w}(w^{(t)}, \psi_k^{(t)}) - \widehat{F}_M^{\mathcal{B}_{t,k}^w}(w, \psi_k^{(t)}) \leq \langle \widehat{g}_{t,k}, w^{(t)} - w \rangle.$$

The batch is independent of $\mathcal{G}_{t,k}$, which fixes $(w^{(t)}, \psi_k^{(t)})$. Taking conditional expectations and using equation 37 yields

$$F(w^{(t)}, \psi_k^{(t)}) - F(w, \psi_k^{(t)}) \leq \left\langle \mathbb{E}[\widehat{g}_{t,k} \mid \mathcal{G}_{t,k}], w^{(t)} - w \right\rangle.$$

Since w is arbitrary, this is the subgradient relation equation 29. Averaging over k and taking expectations yields

$$\mathbb{E}\left[\frac{1}{K} \sum_{k=1}^{K} [F(w^{(t)}, \psi_k^{(t)}) - F(w, \psi_k^{(t)})] \right] \leq \mathbb{E}[\langle \hat{g}^{(t)}, w^{(t)} - w \rangle]. \tag{59}$$

Next, we compare $F(w, \psi_k^{(t)})$ to $\phi_D(w)$. Evaluating the supremum defining $Q_w(\mu^*(w))$ at $\psi_k^{(t)}$ gives

$$F(w, \psi_k^{(t)}) + \mu^*(w)(\rho - s(\psi_k^{(t)})) \leq Q_w(\mu^*(w)) = \phi_D(w).$$

Therefore,

$$F(w, \psi_k^{(t)}) - \phi_D(w) \leq \mu^*(w)(s(\psi_k^{(t)}) - \rho).$$

Averaging the preceding pointwise bound over $k = 1, \ldots, K$ and taking expectations gives

$$\mathbb{E}\left[\frac{1}{K} \sum_{k=1}^{K} F(w, \psi_k^{(t)}) \right] - \phi_D(w) \leq \mu^*(w) \left(\frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[s(\psi_k^{(t)})] - \rho \right). \tag{60}$$

Conditionally on the history $\mathcal{H}_t$ at the start of block t, the decision $w^{(t)}$ and inherited multiplier $\mu_{t,1}$ are fixed. Lemma A.8 applies to that block, uniformly over $\mu_{t,1} \in [0, \mu_{\max}]$. Taking the outer expectation gives

$$\frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[s(\psi_k^{(t)})] \leq \rho + \frac{B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}.$$

Substituting this bound into equation 60 and using the fixed nonnegative multiplier $\mu^*(w) \leq \overline{\mu}^*$ gives

$$\begin{aligned} \mathbb{E}\left[\frac{1}{K} \sum_{k=1}^{K} F(w, \psi_k^{(t)}) \right] - \phi_D(w) &\leq \frac{\mu^*(w) B \mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}} \\ &\leq \frac{\overline{\mu}^* B \mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}. \end{aligned} \tag{61}$$

By Lemma A.7, applied conditionally on $\mathcal{H}_t$ and then averaged,

$$\mathbb{E}[\phi_D(w^{(t)})] \leq \mathbb{E}\left[\frac{1}{K} \sum_{k=1}^{K} F(w^{(t)}, \psi_k^{(t)}) \right] + \frac{B\mu_{\max}}{\sqrt{K}}. \tag{62}$$

Combining equation 62, equation 61, and equation 59 gives

$$\begin{aligned}
\mathbb{E}[\phi_D(w^{(t)})] - \phi_D(w) &\le \mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K} F(w^{(t)}, \psi_k^{(t)})\right] - \phi_D(w) + \frac{B\mu_{\max}}{\sqrt{K}} \\
&= \mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K} [F(w^{(t)}, \psi_k^{(t)}) - F(w, \psi_k^{(t)})]\right] \\
&\quad + \left(\mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K} F(w, \psi_k^{(t)})\right] - \phi_D(w)\right) + \frac{B\mu_{\max}}{\sqrt{K}} \\
&\le \mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K} [F(w^{(t)}, \psi_k^{(t)}) - F(w, \psi_k^{(t)})]\right] \\
&\quad + \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}} \\
&\le \mathbb{E}[\langle \hat{g}^{(t)}, w^{(t)} - w\rangle] \\
&\quad + \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}},
\end{aligned}$$

which proves equation 58.

Now use the projected outer update

$$w^{(t+1)} = \Pi_{\mathcal{W}}\big(w^{(t)} - \eta_w \hat{g}^{(t)}\big).$$

By nonexpansiveness of projection and the bound $\|\hat{g}^{(t)}\| \le K^{-1}\sum_k \|\widehat{g}_{t,k}\| \le G_w$,

$$\begin{aligned}
\|w^{(t+1)} - w\|^2 &\le \|w^{(t)} - \eta_w \hat{g}^{(t)} - w\|^2 \\
&= \|w^{(t)} - w\|^2 - 2\eta_w \langle \hat{g}^{(t)}, w^{(t)} - w\rangle + \eta_w^2 \|\hat{g}^{(t)}\|^2 \\
&\le \|w^{(t)} - w\|^2 - 2\eta_w \langle \hat{g}^{(t)}, w^{(t)} - w\rangle + \eta_w^2 G_w^2.
\end{aligned}$$

Rearranging, taking expectations, and using equation 58,

$$\begin{aligned}
\mathbb{E}[\phi_D(w^{(t)})] - \phi_D(w) \le \frac{\mathbb{E}[\|w^{(t)} - w\|^2] - \mathbb{E}[\|w^{(t+1)} - w\|^2]}{2\eta_w} \\
+ \frac{\eta_w G_w^2}{2} + \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}.
\end{aligned} \tag{63}$$

Summing equation 63 over $t = 1, \dots, T$ and using $\|w^{(1)} - w\| \le D_{\mathcal{W}}$ gives

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\phi_D(w^{(t)})] - \phi_D(w) \le \frac{D_{\mathcal{W}}^2}{2\eta_w T} + \frac{\eta_w G_w^2}{2} \\
+ \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}.
\end{aligned}$$

By convexity of ϕ_D,

$$\phi_D(\overline{w}) \le \frac{1}{T}\sum_{t=1}^{T} \phi_D(w^{(t)}).$$

Therefore,

$$\begin{aligned}
\mathbb{E}[\phi_D(\overline{w})] - \phi_D(w) \le \frac{D_{\mathcal{W}}^2}{2\eta_w T} + \frac{\eta_w G_w^2}{2} \\
+ \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}}.
\end{aligned}$$

For arbitrary $a > 0$, choose $w = w_a$ satisfying $\phi_D(w_a) \le \phi_D^* + a$, and let $a \downarrow 0$. Finally, substituting

$$\eta_w = \frac{D_{\mathcal{W}}}{G_w\sqrt{T}}$$

gives

$$\frac{D_{\mathcal{W}}^2}{2\eta_w T} + \frac{\eta_w G_w^2}{2} = \frac{D_{\mathcal{W}} G_w}{\sqrt{T}}.$$

Thus

$$\mathbb{E}[\phi_D(\overline{w})] - \phi_D^* \le \frac{D_{\mathcal{W}} G_w}{\sqrt{T}} + \frac{B\mu_{\max}}{\sqrt{K}} + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}},$$

which proves equation 31.

Under the additional feasible-nominal assumption, Theorem 3.2 gives $0 \le \phi_D^* - \phi^* \le L_f\sqrt{\rho + \gamma_M}$. Together with weak duality $\phi(\overline{w}) \le \phi_D(\overline{w})$, this implies

$$\begin{aligned}\mathbb{E}[\phi(\overline{w})] - \phi^* &\le \mathbb{E}[\phi_D(\overline{w})] - \phi_D^* + (\phi_D^* - \phi^*) \\ &\le \frac{D_{\mathcal{W}} G_w}{\sqrt{T}} + \frac{B\mu_{\max}}{\sqrt{K}} \\ &\quad + \frac{\overline{\mu}^* B\mu_{\max}}{(\mu_{\max} - \overline{\mu}^*)\sqrt{K}} + L_f\sqrt{\rho + \gamma_M},\end{aligned}$$

which is equation 32. □

B Experimental Details

We evaluate GDRO on two downstream decision problems. The first is contextual newsvendor, where the decision is convex and the optimal order for a fixed scenario distribution has a weighted-quantile form. We study this task on both a synthetic benchmark, used to validate the rare-tail mechanism under a known data-generating process, and the real M5 retail demand dataset. The second downstream problem is robot navigation among pedestrians, where the decision is nonconvex and the nominal model is the official implicit SocialGAN trajectory generator. Across all experiments, the nominal generator defines P_0^x, the adversarial generator defines Q_ψ^x, and each method chooses a decision $w \in \mathcal{W}$ before the realized outcome Y is revealed.

B.1 Baselines and Solvers

All methods use the same fitted nominal generator within each benchmark. They differ only in how they form the scenario distribution used by the downstream decision solver.

Nominal. The nominal baseline optimizes the decision using Monte Carlo samples from P_0^x. This is the standard sample-average decision induced by the fitted generator.

KL. The KL baseline follows the empirical f-divergence DRO formulation of Namkoong & Duchi (2016). At each context, it keeps the nominal sample cloud fixed and adversarially reweights the samples within a KL ball. For a fixed decision, the KL-constrained worst-case weights have the exponential-tilt form (Hu & Hong, 2012); we choose the tilt temperature to match the prescribed radius. The downstream decision is then updated under the weighted samples: by weighted critical quantiles for newsvendor and by weighted first-order optimization for SocialGAN. Thus this baseline can emphasize high-loss nominal samples but cannot create outcomes outside the nominal sample cloud.

W2. The W2 baseline follows the adversarial sample-perturbation approach of Sinha et al. (2018), adapted to the downstream decision problem. Given nominal scenarios from P_0^x, we introduce adversarial support points and move them directly in the output space to increase the downstream loss, while constraining the quadratic transport cost to be at most ρ. We use primal–dual updates to handle this constraint, so that the baseline uses a radius-constrained formulation comparable to

GDRO. The downstream decision is then updated against the perturbed scenarios. We use this finite-sample W2 baseline for both the multi-product newsvendor task and the nonconvex SocialGAN planning task. Unlike GDRO, the adversarial scenarios are free support points in the output space rather than samples produced by the chosen generator family.

GAS-DRO. For the newsvendor experiments, we include a conditional version of GAS-DRO (Wen & Yang, 2026). The adversarial model uses the same conditional decoder architecture as the nominal generator and is optimized at the active context. Its ambiguity set is controlled by the GAS-DRO reconstruction certificate, computed over conditional training pairs. We denote by J_0 the value of this reconstruction certificate for the fitted nominal generator on the training set, and set GAS-DRO radii as multiples of J_0 in the experiments below. We include this baseline only when an explicit decoder is available; we do not apply it to SocialGAN, whose released nominal model is an implicit trajectory sampler.

GDRO. We use the empirical context-local Sinkhorn formulation from Section 2. The nominal generator is frozen, the adversarial generator is initialized from the nominal architecture, and the inner player updates adversarial generator parameters subject to the Sinkhorn certificate at the active context. The decision update uses the average outer subgradient over the active inner adversaries. All Sinkhorn divergences are computed with GeomLoss (Feydy et al., 2019) using $p = 2$ and blur 0.05 across experiments. GeomLoss uses the ground cost $c(y, y') = \frac{1}{2}\|y - y'\|_2^2$ and regularization $\varepsilon = \texttt{blur}^2 = 0.0025$. All reported GDRO radii use this convention. Under the full squared-Euclidean cost used in our theory, the equivalent regularization is 0.005 and the radius is twice the reported value. Task-specific radii, sample sizes, and optimization budgets are listed in the configuration tables below.

Radius and optimization budget. For the synthetic newsvendor experiment, we use fixed radii for KL, W2, and GDRO. For M5, the risk score is the CVaR_{90} of nominally simulated newsvendor costs. For SocialGAN, it is the estimated collision probability of the nominal robot plan. Radii and GDRO iteration budgets are adjusted using fixed, method-specific mappings of these scores. The scores use observed contexts and nominal samples, not realized test outcomes. Configuration details are provided below and in the released code.

B.2 Synthetic Contextual Newsvendor

Motivation. The synthetic benchmark is designed to test a specific failure mode of nominal generative decision-making: a conditional generator may fit the bulk of the demand distribution while underrepresenting rare, high-cost demand tails. The data-generating process therefore contains frequent base contexts and rare contexts in which a tail component may activate. This lets us evaluate decisions against the true conditional law while still requiring every method to make decisions from the same learned nominal generator.

Data-generating process. We generate $p = 20$ product demands from contexts $x \in \mathbb{R}^5$, with

$$x \sim \mathcal{N}(0, I).$$

A context is rare if

$$\text{rare}(x) = \mathbf{1}\{a^\top x > \tau\},$$

where a is a fixed unit vector and τ is the 0.9 training quantile of $a^\top x$.

Demand is generated through a latent factor $h \in \mathbb{R}^2$. In the base component,

$$h \mid x \sim \mathcal{N}(Ax, \Sigma_h).$$

In rare contexts, a tail component is activated with probability $\pi_{\text{tail}} = 0.10$:

$$h \mid x, \text{tail} \sim \mathcal{N}(Ax + \Delta_h, \Sigma_{\text{tail}}), \qquad \Delta_h = (2.0, 0).$$

Frequent contexts always use the base component. For product j, the latent score is

$$s_j(x, h) = b_j + r_j^\top x + \beta_j h_1 + \gamma_j h_2 + a_j^{\text{sin}} \sin(v_j^\top h) + \epsilon_j, \qquad \epsilon_j \sim \mathcal{N}(0, \sigma_\epsilon^2),$$

and demand is

$$Y_j = \mu_0 + \mu_1 \,\text{softplus}(s_j(x, h)).$$

We use $\sigma_\epsilon = 0.15$, $\mu_0 = 10$, and $\mu_1 = 5$. Product parameters share group-level structure plus product-specific perturbations, inducing correlated nonlinear demands across products.

Table 5: Synthetic contextual newsvendor configuration.

Item	Value
Products	20
Context dimension	5
DGP latent dimension	2
Rare-context threshold	0.9 quantile of $a^\top x$
Tail probability in rare contexts	0.10
Tail shift	$(2.0, 0)$
Observation noise	0.15
Demand scale (μ_0, μ_1)	$(10, 5)$
Train / validation / test contexts	$2000/500/200$
Underage costs $c_{u,j}$	Uniform on $[15, 25]$
Overage costs $c_{o,j}$	Uniform on $[1, 3]$
Nominal generator	Conditional VAE
VAE latent dimension	4
VAE hidden width	128
cVAE epochs	300
VAE learning rate	10^{-3}
VAE KL weight	0.02
Nominal/KL samples per context	256
GDRO/W2/GAS-DRO samples per context	128
KL radius	0.1
GDRO radius	2.5
W2 radius	2.5
GAS-DRO radius	$1.2J_0$

Nominal generator. The nominal model is a conditional VAE trained on samples from the data-generating process. The context is provided to the encoder and decoder. Demand is standardized for VAE training, but generated samples are transformed back to demand units before the downstream decision and evaluation. The full synthetic newsvendor configuration is summarized in Table 5.

Downstream decision. The decision is the order vector $w \in \mathbb{R}^{20}_+$. The loss is the multi-product newsvendor cost

$$f(w, Y) = \sum_{j=1}^{20} c_{u,j}(Y_j - w_j)_+ + c_{o,j}(w_j - Y_j)_+.$$

For any weighted empirical scenario cloud, the optimal decision is the per-product weighted critical quantile with critical ratio $c_{u,j}/(c_{u,j} + c_{o,j})$. The oracle is computed using simulator samples from the true conditional law at the same context and is used only for evaluation.

B.3 M5 Retail Newsvendor

Dataset and product selection. The M5 experiment uses real retail demand from the M5 forecasting benchmark (Makridakis et al., 2022). Our goal in this experiment is not to compete on the full M5 forecasting task, but to obtain a stable real-data nominal generator for testing the downstream newsvendor decision mechanism. We therefore construct a contextual multi-product newsvendor task from historical sales, and select $p = 20$ products from the chosen M5 subset whose nonzero demand ratio is at least 0.95. This keeps the task multivariate while avoiding products dominated by zeros, which would make conditional scenario generation unstable and obscure the effect of the robust decision layer.

Downstream decision. The downstream decision is the same multi-product newsvendor decision as in Appendix B.2. Let $Y_t = (Y_{t,1}, \ldots, Y_{t,p})^\top$ denote the realized demand vector for $p = 20$ products on day t, and let $w_t = (w_{t,1}, \ldots, w_{t,p})^\top \in \mathbb{R}^p_+$ denote the order decision. The held-out

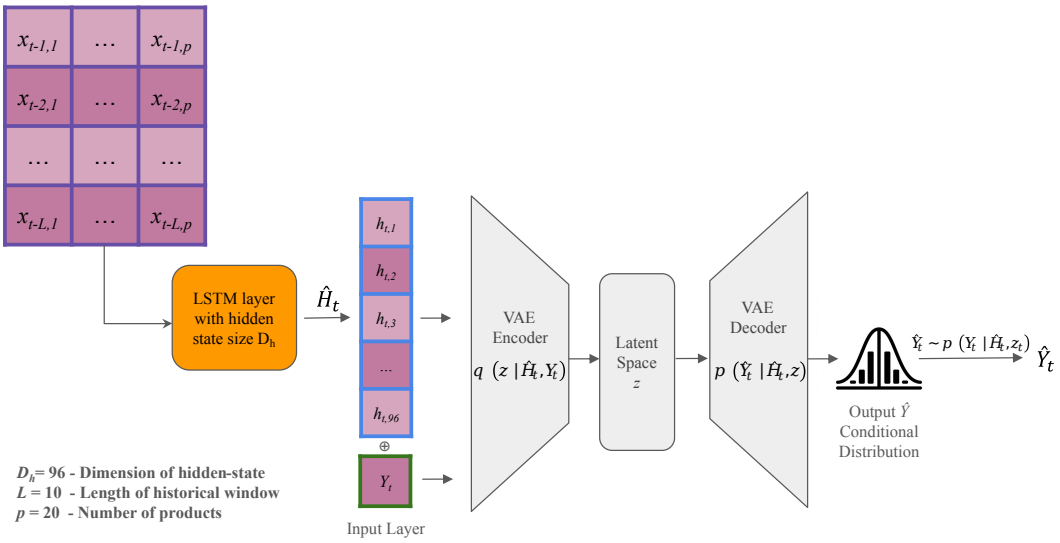

Figure 2: Nominal CVAE–LSTM generator for M5 demand scenarios. The LSTM encodes the historical demand window into $\widehat{H}_t$. During training, the encoder uses $(\widehat{H}_t, Y_t)$; during generation, latent variables are sampled from the conditional prior given $\widehat{H}_t$ and decoded into next-day demand scenarios.

realized loss is

$$f(w_t, Y_t) = \sum_{j=1}^{20} c_{u,j}(Y_{t,j} - w_{t,j})_+ + c_{o,j}(w_{t,j} - Y_{t,j})_+.$$

For any weighted empirical scenario cloud, the decision is computed by the per-product weighted critical quantile.

Nominal generator. The nominal model is a conditional VAE–LSTM demand generator. For each day t, the input context X_t is a historical demand window. A single-layer LSTM encodes this history into a temporal representation $\widehat{H}_t$. During training, the VAE encoder conditions on $(\widehat{H}_t, Y_t)$ and learns an approximate posterior over the latent variable. At generation time, the future demand Y_t is not observed; latent variables are sampled from the learned conditional prior given $\widehat{H}_t$, and the decoder produces next-day demand scenarios. The nominal CVAE–LSTM architecture is illustrated in Figure 2.

Adaptive radius and budget. For each M5 context, we draw 256 demand scenarios from the nominal generator and compute the nominal newsvendor decision. We evaluate its cost on the same scenarios and use the empirical CVaR_{90} as the risk score r, computed as the mean of costs at or above their empirical 90th percentile. A larger score indicates greater predicted tail cost under the nominal decision.

For each adaptive method m, we normalize the score using fixed low- and high-risk anchors:

$$\alpha_m(r) = \left[\frac{r - r_{\min,m}}{r_{\max,m} - r_{\min,m}} \right]_{[0,1]}.$$

KL, W2, and GDRO map this normalized score to their robustness radii using fixed, method-specific rules. GDRO also allocates larger outer and inner iteration budgets to higher-risk contexts. The other methods use fixed optimization budgets, and GAS-DRO uses a fixed radius, since its certificate is based on training-set reconstruction loss and preliminary experiments showed no improvement from adapting the radius.

The radius ranges are summarized in Table 6. The exact radius mappings and optimization budgets are provided in the released code.

Table 6: M5 retail newsvendor configuration.

Item	Value
Products	20
Historical window length	10 days
Input-context shape	10×20
Forecast horizon	1 day
Train / validation / test contexts	$1500/100/300$
Evaluated test contexts	200 (evenly selected)
Nominal generator	Conditional VAE–LSTM
LSTM layers	1
LSTM hidden-state dimension	96
VAE latent dimension	12
VAE hidden dimension	192
Maximum VAE epochs	600
Early-stopping patience	120
VAE batch size	64
VAE learning rate	5×10^{-4}
VAE weight decay	10^{-5}
VAE KL weight	10^{-3}
Prior-mean loss weight	1
Method seeds	$\{0, 1, 2, 3, 4\}$
Nominal/KL samples per context	256
GDRO/W2/GAS–DRO samples per context	128
KL radius	$[0.500, 0.506]$
W2 radius	$[1.0, 7.0]$
GAS–DRO radius	$1.2J_0$
GDRO radius	$[22.0, 40.0]$

B.4 SocialGAN Robot Navigation

Nominal model and data. We use the official pretrained SocialGAN pedestrian trajectory model and the ETH/UCY benchmark protocol from Gupta et al. (2018). The model observes 8 frames of pedestrian history and generates 12 future frames. We use the released SocialGAN predictor and its test split, without retraining the trajectory generator.

Downstream decision. The robot navigation task is added by us on top of the SocialGAN prediction setting. At a test context x, the robot observes the same past pedestrian trajectories as SocialGAN. It then chooses a 12-step velocity plan

$$w = (u_1, \dots, u_{12}), \qquad u_t \in \mathbb{R}^2,$$

before the true pedestrian futures are revealed. The robot follows single-integrator dynamics from a sampled start point to a sampled goal, with each velocity projected onto a maximum-speed ball. Optimization uses generated pedestrian futures from the nominal or robust model; evaluation uses the true held-out pedestrian futures.

Planning objective and evaluation. For a pedestrian-future scenario $Y = \{y_{j,t}\}_{j=1,t=1}^{N,T}$, the robot decision is a T-step velocity plan

$$w = (u_1, \dots, u_T), \qquad u_t \in \mathbb{R}^2, \qquad T = 12.$$

Starting from s, the induced robot path is

$$r_t(w) = s + \sum_{\ell=1}^{t} u_\ell, \qquad \|u_t\|_2 \le v_{\max}.$$

The speed constraint is enforced by projecting each velocity vector onto the Euclidean ball of radius $v_{\max}$. Let g be the sampled goal and let $\bar{r}_t$ be the straight-line reference path from s to g. The

smooth loss optimized by all methods has the form

$$
\begin{aligned}
f(w,Y) = & \underbrace{\lambda_p \frac{1}{T}\sum_{t=1}^{T}\|r_t(w)-g\|_2^2 + \lambda_T\|r_T(w)-g\|_2^2}_{\text{progress and terminal arrival}} \\
& + \underbrace{\lambda_{\mathrm{tr}}\frac{1}{T}\sum_{t=1}^{T}\|r_t(w)-\bar{r}_t\|_2^2}_{\text{weak reference-path tracking}} \\
& + \underbrace{\lambda_u\frac{1}{T}\sum_{t=1}^{T}\|u_t\|_2^2 + \lambda_s\frac{1}{T-1}\sum_{t=2}^{T}\|u_t-u_{t-1}\|_2^2}_{\text{control effort and smoothness}} \\
& + \underbrace{\lambda_c\frac{1}{T}\sum_{t=1}^{T}\mathrm{softplus}\left(\frac{r_{\mathrm{safe}}-\widetilde{d}_t(w,Y)}{\tau_b}\right)^2}_{\text{differentiable collision barrier}}.
\end{aligned}
$$

Here

$$
\widetilde{d}_t(w,Y) = -\tau_d \log\sum_{j=1}^{N}\exp\left(-\frac{\|r_t(w)-y_{j,t}\|_2}{\tau_d}\right)
$$

is a soft-min approximation of the nearest-pedestrian distance at time t. The first line encourages the robot to make progress and arrive at the goal by the end of the horizon. The second line weakly discourages unnecessary detours. The third line penalizes large and rapidly changing velocities. The last line is a smooth safety barrier: it becomes large when the robot comes within the safety radius r_{safe} of a sampled pedestrian future. Nominal, W2, and GDRO optimize the average of $f(w,Y)$ over their corresponding scenario clouds, while KL optimizes a weighted average under adversarial sample weights.

For final evaluation, we use the true held-out pedestrian future and a hard, non-differentiable score that prioritizes safety and arrival:

$$
\begin{aligned}
\mathrm{Eval}(w,Y^{\mathrm{true}}) = & C_{\mathrm{col}}\mathbf{1}\{d_{\min} < r_{\mathrm{col}}\} + C_{\mathrm{near}}\mathbf{1}\{d_{\min} < r_{\mathrm{near}}\} \\
& + C_{\mathrm{fail}}\mathbf{1}\{t_{\mathrm{arr}} = -1\} + c_{\mathrm{time}}\widetilde{t}_{\mathrm{arr}} + c_{\mathrm{len}}L_{\mathrm{path}} + c_{\mathrm{dev}}D_{\mathrm{ref}}.
\end{aligned}
$$

Here $d_{\min}$ is the minimum robot–pedestrian distance before arrival, t_{arr} is the first frame at which $\|r_t(w)-g\|_2 \le r_{\mathrm{arr}}$, $\widetilde{t}_{\mathrm{arr}} = t_{\mathrm{arr}}$ if the robot arrives and $T+1$ otherwise, L_{path} is path length until arrival, and D_{ref} is average deviation from the straight-line reference. Once the robot reaches the goal, later pedestrian motion is not counted as a collision. This hard score is used only for reporting; optimization uses the smooth loss above.

Fixed task set. We uniformly select 50 contexts from the SocialGAN test split. For each selected context, we generate 4 start–goal queries, giving 200 fixed robot tasks. Queries are selected using only nominal SocialGAN samples: the straight-line robot path must have nominal collision probability in $[0.60, 0.90]$ under a screening radius 0.30, and the start and goal locations must be initially clear of pedestrians. The evaluation collision radius is 0.50. Thus the benchmark focuses on nontrivial planning instances without using true future trajectories during task construction.

Adaptive radius and budget. For each query, we first solve the nominal robot plan and estimate its nominal collision probability $\widehat{p}$ using generated futures and the evaluation collision radius. The GDRO and W2 radii are chosen by the clipped linear rule

$$
\rho(\widehat{p}) = \rho_{\min} + (\rho_{\max}-\rho_{\min})\left[\frac{\widehat{p}-p_{\min}}{p_{\max}-p_{\min}}\right]_{[0,1]},
$$

with $\rho_{\min} = 0.5$, $\rho_{\max} = 3.0$, $p_{\min} = 0.05$, and $p_{\max} = 0.50$. The KL radius is adapted by the same rule from 0.1 to 0.7. The GDRO optimization budget is also adapted using $\widehat{p}$: low-risk queries use the smaller budget in Table 8, while high-risk queries use the larger budget. All adaptive quantities use only the observed context and nominal generated futures, never the realized test future.

Worst-case trajectory inspection. We inspect an additional fixed SocialGAN task with nontrivial nominal risk. For this task, the nominal-plan collision probability is $\widehat{p} = 0.406$, giving the adaptive radius $\rho = 2.479$. We measure trajectory irregularity by the second difference

$$a_{j,t} = \|y_{j,t+1} - 2y_{j,t} + y_{j,t-1}\|_2 ,$$

which measures changes in velocity across consecutive prediction steps.

Figure 3(a)–(b) compares the adversarial trajectories produced by W_2 and GDRO. The solid orange and green curves show the critical pedestrian in each method's highest-loss joint scenario. Gray dashed curves are the paired nominal trajectories, blue dashed curves are the realized trajectories, and black curves are the resulting robot decisions. Faint curves show additional adversarial samples for the same pedestrian. Because the W_2 adversary optimizes trajectory coordinates directly, some of its samples change direction sharply over one or two prediction steps despite having smooth paired nominal trajectories. Their acceleration 95th percentile is 1.152, compared with 0.053 for GDRO and 0.037 for the nominal generator. GDRO instead generates its adversarial samples through a perturbed SocialGAN model and retains smoother trajectory structure.

Figure 3(c) evaluates the two robot decisions on the realized pedestrian future. The W_2 decision reaches a minimum distance of 0.45, below the collision threshold of 0.5, and does not reach the goal. The GDRO decision maintains a minimum distance of 1.38 and reaches the goal at step 9.

One may think that the irregular W_2 trajectories are caused simply by a large ambiguity radius. Figure 3(d) tests this explanation by varying the prescribed radius while keeping the remaining settings fixed. Even at $\rho = 0.02$, the acceleration p95 is $1.0\times$ nominal, whereas the maximum remains $19.7\times$ nominal. Thus, most adversarial samples are close to the nominal motion, while a small number still exhibit abrupt and kinematically extreme changes. This behavior reflects the average-transport geometry of the empirical W_2 ball: the transport budget can be concentrated on a few support points when most points move very little. The green band reports the fixed GDRO reference, and the vertical dotted line marks the W_2 radius $\rho = 2.479$ used for this task.

Mini-batch-size sensitivity. To isolate sensitivity to the stochastic Sinkhorn mini-batch size, we hold the 200 tasks, random seed, nominal plans, adaptive radii and budgets, loss, and optimization settings fixed, and vary only M. Table 7 shows that GDRO remains stable for $M \in \{8, 16, 32, 64\}$. The five-seed comparison in Table 4 uses $M = 8$; this ablation uses one fixed seed to isolate the effect of M.

Table 7: Sensitivity of GDRO to the number M of generated trajectories per stochastic Sinkhorn evaluation on the 200 fixed SocialGAN robot-navigation tasks. All settings other than M are unchanged.

M	Collisions	Near misses	Arrivals	Eval. loss ↓	Min. dist. ↑	Path length
8	5/200	43/200	198/200	21.41	0.959	5.876
16	7/200	36/200	199/200	21.72	0.961	5.964
32	6/200	43/200	199/200	22.01	0.957	5.994
64	6/200	45/200	200/200	22.05	0.953	6.004

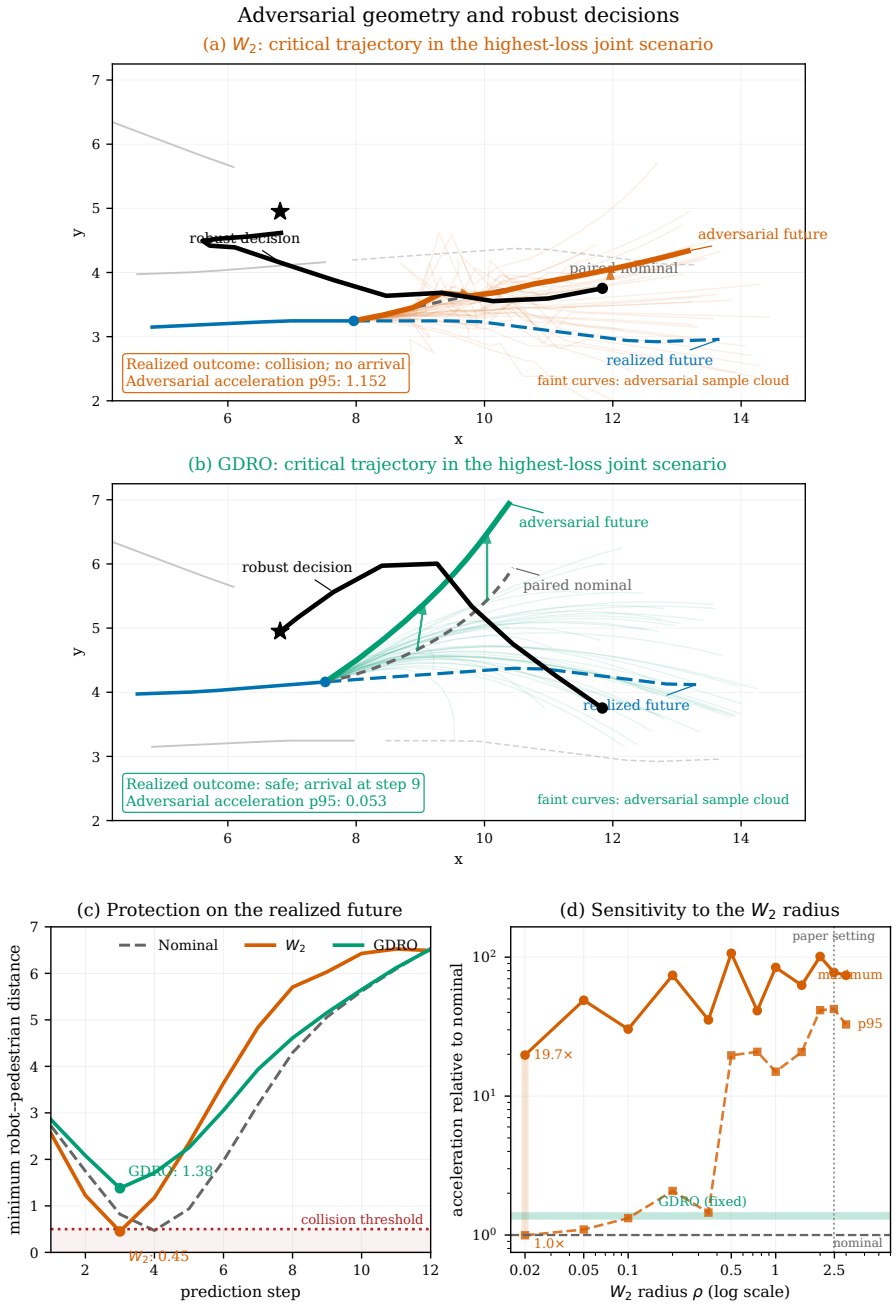

Figure 3: Worst-case inspection on an additional fixed SocialGAN task. Panels (a)–(b) show adversarial trajectories and the corresponding robot decisions for W_2 and GDRO. Panel (c) reports the minimum robot–pedestrian distance on the realized future. Panel (d) reports the normalized p95 and maximum acceleration as the prescribed W_2 radius varies. The green band is the fixed GDRO reference, and the vertical dotted line marks the radius used for this task.

Table 8: SocialGAN robot navigation configuration.

Item	Value
Nominal model	Official SocialGAN (Gupta et al., 2018)
Dataset protocol	ETH/UCY SocialGAN test split
Observed / predicted frames	$8/12$
Selected test contexts	50 uniformly selected contexts
Queries per context	4
Total robot tasks	200
Robot decision horizon T	12
Robot decision w	12 two-dimensional velocity controls
Robot dynamics	$r_t = s + \sum_{\ell=1}^{t} u_\ell$
Robot speed limit $v_{\max}$	$2\times$ reference pedestrian speed
Velocity constraint	Projection to $\|u_t\|_2 \le v_{\max}$
Straight-line screening probability	$[0.60, 0.90]$
Screening collision radius	0.30
Evaluation collision radius r_{col}	0.50
Near-miss radius r_{near}	0.80
Arrival radius r_{arr}	0.25
Nominal SocialGAN samples per context	64
Nominal plan steps	600
GDRO mini-batch size M	8 trajectories per stochastic Sinkhorn evaluation
W2 support samples per context	512
GDRO/W2 radius range ρ	$[0.5, 3.0]$
KL radius range	$[0.1, 0.7]$
Risk anchors $(p_{\min}, p_{\max})$	$(0.05, 0.50)$
Low-risk GDRO budget	150 outer, 3 dual cycles, 5 adversary steps
High-risk GDRO budget	500 outer, 5 dual cycles, 8 adversary steps
Controls updates per outer step	5
Sinkhorn blur	0.05
Loss scale in adversarial update	25.0
Progress weight λ_p	1.2
Terminal arrival weight λ_T	35.0
Reference tracking weight λ_{tr}	0.05
Control effort weight λ_u	0.15
Control smoothness weight λ_s	0.75
Collision barrier weight λ_c	80.0
Safety radius in smooth loss r_{safe}	0.55
Barrier temperature τ_b	0.10
Soft-min temperature τ_d	0.10
Decision learning rate	0.035
Adversarial generator learning rate	10^{-6}
Initial dual value / GDRO dual learning rate	$5.0/10.0$
Gradient clipping	1.0
W2 support-point learning rate	0.015
W2 dual learning rate	2.0
W2 outer / support / control steps	$180/15/5$
Evaluation collision penalty C_{col}	100.0
Evaluation near-miss penalty C_{near}	20.0
Evaluation failure penalty C_{fail}	50.0
Evaluation time / path / deviation weights $(c_{\mathrm{time}}, c_{\mathrm{len}}, c_{\mathrm{dev}})$	$1.0/1.0/0.2$