Supplementary for *"Sinkhorn Distributionally Robust Optimization"*

EC.1. Detailed Experiment Setup

Unless stated otherwise, we solved the SAA, Wasserstein DRO, and KL-divergence DRO baseline models exactly using the off-the-shelf solver Mosek (ApS, 2021). Optimization hyperparameters, such as step size, maximum iterations, and number of levels, were tuned to minimize training error after 10 outer iterations. We use RT-MLMC subgradient estimator to solve the Sinkhorn DRO model. We employed the *warm starting* strategy during the iterative procedure: we set the initial guess of parameter θ at the beginning of outer iteration as the one obtained from the SAA approach. At other outer iterations, the initial guess of parameter θ is set to be the final obtained solution θ at the last outer iteration. The following subsections outline some special reformulations, optimization algorithms used to solve the baseline models.

EC.1.1. Setup for Newsvendor Problem and Running Time

To solve the 2-Wasserstein DRO model with radius ρ, we approximate the support of worst-case distribution using discrete grid points. Denote by $\mathcal{D}_n = \{x_1, \ldots, x_n\}$ the set of observed n samples and $\mathcal{G}_{200-n}$ the set of $200-n$ points evenly supported on the interval $[0, 10]$. Then the support of worst-case distribution is restricted to $\mathcal{D}_n \cup \mathcal{G}_{200-n} := \{\hat{z}_1, \ldots, \hat{z}_{200}\}$. The corresponding 2-Wasserstein DRO problem has the following linear programming reformulation:

$$\begin{aligned} \min_{\theta,\lambda,s} \quad & \lambda\rho + \frac{1}{n}\sum_{i=1}^n s_i \\ \text{s.t.} \quad & k\theta - u\min(\theta, \hat{z}_j) - \lambda(x_i - \hat{z}_j)^2 \le s_i, \quad \forall i \in [n], \forall j \in [200]. \end{aligned}$$

EC.1.2. Setup for Mean-risk Portfolio Optimization

From (Mohajerin Esfahani and Kuhn, 2017, Eq. (27)) we can see that the 1-Wasserstein DRO formulation with radius ρ for the portfolio optimization problem becomes

$$\begin{aligned} \min_{\theta,\tau,\lambda,s} \quad & \lambda\rho + \frac{1}{n}\sum_{i=1}^n s_i \\ \text{s.t.} \quad & \theta \in \Theta, \quad b_j\tau + a_j\langle\theta, \hat{z}_i\rangle \le s_i, i \in [n], j \in [H], \\ & \|a_j\theta\|_2 \le \lambda, j \in [H]. \end{aligned}$$

Also, we argue that the 2-Wasserstein DRO formulation with radius ρ for the portfolio optimization problem has a finite convex reformulation:

$$\begin{aligned} &\inf_{\theta\in\Theta,\tau} \sup_{\mathbb{P}:\, W_2(\mathbb{P},\widehat{\mathbb{P}}_n)\le\rho} \mathbb{E}_{\mathbb{P}}\big[\max_{j\in[H]} a_j\langle\theta, z\rangle + b_j\tau\big] \\ =& \inf_{\theta\in\Theta,\tau,\lambda\ge 0} \left\{ \lambda\rho^2 + \frac{1}{n}\sum_{i=1}^n \sup_{s_i} \left\{ \max_{j\in[H]} a_j\langle\theta, s_i\rangle + b_j\tau - \lambda\|s_i - \hat{z}_i\|_2^2 \right\} \right\}. \end{aligned}$$

In particular, the inner subproblem has the following reformulation:

$$\begin{aligned}
&\sup_{s_i}\left\{\max_{j\in[H]} a_j\langle\theta,s_i\rangle+b_j\tau-\lambda\|s_i-\hat{z}_i\|_2^2\right\}\\
=&\max_{j\in[H]}\ b_j\tau+\sup_{s_i}\left\{a_j\langle\theta,s_i\rangle-\lambda\|s_i-\hat{z}_i\|_2^2\right\}\\
=&\max_{j\in[H]}\ b_j\tau+\frac{a_j^2}{4\lambda}\|\theta\|_2^2+a_j\langle\theta,\hat{z}_i\rangle.
\end{aligned}$$

Hence, the 2-Wasserstein DRO can be reformulated as

$$\begin{aligned}
\min_{\theta,\tau,\lambda,s}\quad & \lambda\rho^2+\frac{1}{n}\sum_{i=1}^n s_i\\
\text{s.t.}\quad & \theta\in\Theta,\quad b_j\tau+a_j\langle\theta,\hat{z}_i\rangle+\frac{a_j^2}{4\lambda}\|\theta\|_2^2\le s_i,\quad i\in[n],j\in[H].
\end{aligned}$$

EC.1.3. Setup for Adversarial Multi-class Logistic Regression

The procedure for generating various adversarial perturbations is reported in the following:

(I) For a given classifer B and data sample $(x,\boldsymbol{y})$, the ℓ_p-norm ($p\in\{1,2\}$) adversarial attack based on projected gradient method (Madry et al., 2018) iterates as follows: $x_0\leftarrow x$ and

$$\begin{cases}
\Delta x^{k+1}\leftarrow \underset{\|\eta\|_p\le\xi}{\arg\max}\ \left\{\nabla_x h_B(x^k,\boldsymbol{y})^{\mathrm{T}}\eta\right\},\\
x^{k+1}\leftarrow \mathrm{Proj}_{\{x':\ \|x-x'\|_p\le\xi\}}\left\{x^k+\frac{\alpha}{\sqrt{k+1}}\Delta x^{k+1}\right\}.
\end{cases}$$

We perform the gradient update above for 15 steps with initial learning rate $\alpha=1$. When $p=1$, the radius of attack $\xi\in\{\texttt{0,3e-3,6e-3,9e-3,1.2e-2}\}\cdot\varrho$; and when $p=2$, the radius $\xi\in\{\texttt{0,8e-3,1.6e-2,2.4e-2,3.2e-2}\}\cdot\varrho$.

(II) For a given feature vector x, the perturbed feature using white Laplacian noise becomes $x+\xi\cdot\zeta$, where the random vector ζ follows the isotropic Laplace distribution with zero mean and unit variance. The ratio $\xi\in\{\texttt{0,2e-3,4e-3,6e-3,8e-3}\}\cdot\varrho$. Similarly, the perturbed feature using white Gaussian noise becomes $x+\xi\cdot\zeta$, with ζ being the isotropic Gaussian distribution with zero mean and unit variance. In this case, the ratio $\xi\in\{\texttt{0,5e-2,1e-1,1.5e-1,2e-1}\}\cdot\varrho$.

In this example, we use stochastic gradient methods to solve the SAA formulation and all penalized DRO formulations. We terminate the training of SAA or DRO models when the number of epoches, i.e., the number of times for processes each training sample, exceeds 30. In is worth mentioning that the Wasserstein DRO model with a fixed Lagrangian multiplier λ using samples $\{x_i,\boldsymbol{y}_i\}_{i=1}^n$ can be reformulated as

$$\min_B\ \frac{1}{n}\sum_{i=1}^n\left[\max_{x\in\mathbb{R}^d}\left\{h_B(x,\boldsymbol{y}_i)-\lambda c(x_i,x)\right\}\right]. \tag{EC.1}$$

EC.2. Additional Validation Experiments

EC.2.1. Comparison of Optimization Algorithms: Linear Regression

To examine the performance of different (sub)gradient estimators, we study the problem of distributionally robust linear regression (see the setup in Example 2). We take the nominal distribution $\widehat{\mathbb{P}}$ as the empirical one based on samples $\{(a_i, b_i)\}_{i=1}^n$. As a consequence, the inner objective function in (12) has the closed form expression:

$$F(\theta) = \frac{1}{n}\sum_{i=1}^n (a_i^{\mathrm{T}}\theta - b_i)^2 + \frac{\frac{1}{n}\sum_{i=1}^n (a_i^{\mathrm{T}}\theta - b_i)^2}{\frac{1}{2}\lambda\|\theta\|_2^{-2} - 1} - \frac{\lambda\epsilon}{2}\log\det\left(I - \frac{\theta\theta^{\mathrm{T}}}{\frac{1}{2}\lambda}\right), \quad \text{if } \|\theta\|_2^2 < \frac{\lambda}{2},$$

and otherwise $F(\theta) = \infty$. We take the constraint set $\Theta = \{\theta : \ \|\theta\|_2^2 \le 0.999 \cdot \frac{\lambda}{2}\}$. Similar to the setup in (Li et al., 2022, Section 5.1), we examine the performance using three LIBSVM regression real world datasets (Chang and Lin, 2011): housing, mg, and mpg.

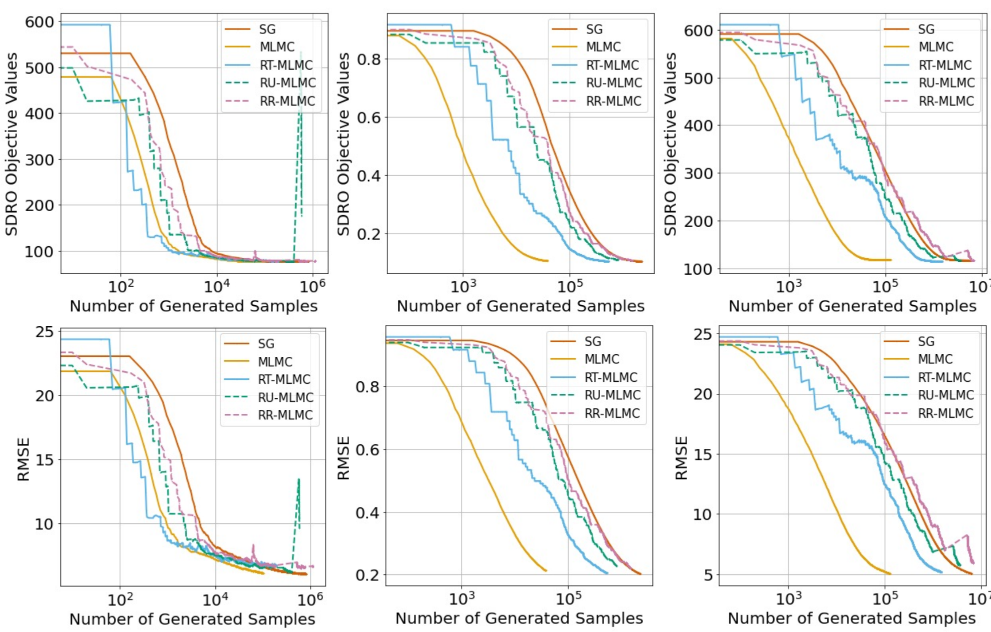

Figure EC.1 Comparison results of SG, (V-)MLMC, RT-MLMC, RU-MLMC, and RR-MLMC on robust linear regression problem in terms of sample complexities from $\widehat{\mathbb{P}}$ and $\mathbb{Q}_{x,\epsilon}$. From left to right, the figures correspond to three different regression datasets: (a) housing; (b) mg; and (c) mpg. From top to bottom, the figures correspond to plots of (a) Sinkhorn DRO objective values; and (b) RMSE of obtained solutions.

The quality of proposed gradient estimators is examined in a single BSMD step with specified hyper-parameters $(\lambda, \epsilon) = (10^3, 10^{-1})$. For baseline comparison, we examine the SG, RT-MLMC estimators together with the (V-)MLMC, RU-MLMC, and RR-MLMC estimators that have been proposed in (Hu et al., 2021). We have validated in Theorem 2 that both SG and RT-MLMC estimators have convergence guarantees for smooth and nonsmooth loss functions, whereas SG estimator has slower convergence rate. The

(V-)MLMC estimator only have convergence guarantees for smooth loss functions, and RU-MLMC/RR-MLMC estimators do not have convergence guarantees as their (sub)gradient second-order moments are unbounded.

For a given solution θ, we quantify its performance using the corresponding Sinkhorn DRO objective value. Besides, we report its root-mean-square error (RMSE) on training data. Thus, the smaller those two performance criteria are, the smaller the solution's optimization performance has. Fig. EC.1 shows the performance of various gradient estimators in terms of the number of generated samples from $\widehat{\mathbb{P}}$ and $\mathbb{Q}_{x,\epsilon}, x \in \mathrm{supp}\widehat{\mathbb{P}}$ based on these criteria. The results demonstrate that the SG scheme does not perform competitively, as expected from our theoretical analysis, which shows that SG has the worst complexity order. In contrast, using other four types of MLMC methods lead to faster convergence behavior. While the RU-MLMC and RR-MLMC schemes exhibit competitive performance, the optimization procedure shows some oscillations. One possible explanation is that the variance values of those gradient estimators are unbounded, making these two approaches unstable.

EC.2.2. Comparison of Optimization Algorithms: Portfolio Optimization

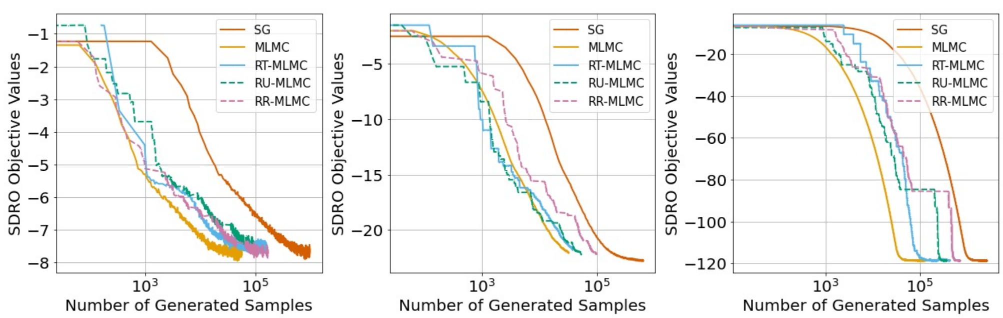

Figure EC.2 Comparison results of SG, (V-)MLMC, RT-MLMC, RU-MLMC, and RR-MLMC on portfolio optimization problem. From left to right, plots correspond to three different instances of $(n,d) \in \{(50,50),(100,100),(400,400)\}$.

In this subsection, we validate the competitive performance of RT-MLMC gradient estimator on the case where the loss is convex and nonsmooth, and we try to solve the 2-SDRO formulation. We consider the portfolio optimization problem, and specify instances $(n,d) = (50,50),(100,100),(400,400)$. We quantify the performance of obtained solution using the Sinkhorn DRO objective value. Since in this problem setup no analytical expression of the objective value is available, we estimate the objective value using (18) with hyper-parameters $L = 8$ and $n_L = 10^3$. Fig. EC.2 shows the performance in terms of the number of generated samples based on this criterion. The results demonstrate that even for nonsmooth loss function, those listed MLMC-based gradient estimators have better performance than the SG estimator. Besides, the proposed RT-MLMC and standard (V-)MLMC schemes have comparable performance, and in some cases (V-)MLMC estimator even has better performance. It is an open question that whether the (V-)MLMC estimator will have the similar performance guarantees as the RT-MLMC estimator for convex nonsmooth optimization, which can be a topic for future study.

EC.2.3. Comparison of Running Time for Different Baselines

The computational time for the newsvendor problem in Section 5.1 is reported in Table EC.1. We observe that the training time of 2-Wasserstein DRO model increases quickly as the sample size increases, while the training time of other DRO models increases mildly in the training sample size.

Table EC.1 Average computational time (in seconds) per problem instance for the newsvendor problem.

Model	Exponential			Gamma			Gaussian Mixture		
	$n=10$	$n=30$	$n=100$	$n=10$	$n=30$	$n=100$	$n=10$	$n=30$	$n=100$
SAA	4.11e-3	4.66e-3	4.67e-3	3.96e-3	4.57e-3	5.81e-3	3.82e-3	4.60e-3	4.79e-3
KL-DRO	6.92e-3	8.17e-3	1.15e-2	8.07e-3	8.24e-3	1.16e-2	7.77e-3	8.47e-3	1.12e-2
1-SDRO	8.77e-2	8.88e-2	1.03e-1	2.76e-2	3.40e-2	4.72e-2	2.90e-2	3.13e-2	4.50e-2
2-WDRO	1.68e00	5.67e00	2.71e01	1.72e00	5.63e00	2.77e01	1.51e00	5.47e00	2.84e01
2-SDRO	3.16e-2	3.77e-2	5.92e-2	2.64e-2	2.95e-2	5.02e-2	2.57e-2	3.10e-2	4.87e-2

The computational time for the portfolio optimization problem in Section 5.2 is reported in Table EC.2. We observe that the computational time of 1- or 2-SDRO model increases mildly as the problem input size increases. Also, SDRO models do not have the smallest computational time in general. The reason is that in this example, other DRO models have tractable finite-dimensional conic programming formulations so that off-the-shelf software can solve them efficiently. In contrast, Sinkhorn DRO models do not have special reformulation, but they can still be solved in a reasonable amount of time.

Table EC.2 Average computational time (in seconds) per problem instance for portfolio optimization problem.

(n,d) Values	SAA	KL-DRO	1-WDRO	1-SDRO	2-WDRO	2-SDRO
$(30,30)$	6.76e-03	1.42e-02	7.80e-03	4.91e-02	8.95e-03	5.00e-02
$(50,30)$	7.31e-03	1.84e-02	8.33e-03	1.87e-01	1.11e-02	5.88e-02
$(100,30)$	8.99e-03	2.95e-02	1.03e-02	2.78e-01	1.12e-02	6.00e-02
$(150,30)$	1.12e-02	4.14e-02	1.21e-02	2.80e-01	1.22e-02	6.95e-02
$(200,30)$	1.12e-02	5.66e-02	1.35e-02	2.99e-01	1.48e-02	7.67e-02
$(400,30)$	1.89e-02	6.45e-02	2.09e-02	2.99e-01	2.30e-02	1.62e-01
$(100,5)$	5.76e-03	1.46e-02	6.79e-03	1.05e-01	7.62e-03	5.40e-02
$(100,10)$	6.18e-03	1.70e-02	7.70e-03	1.08e-01	8.73e-03	5.55e-02
$(100,20)$	7.43e-03	1.82e-02	8.41e-03	1.12e-01	9.44e-03	5.58e-02
$(100,40)$	9.87e-03	3.25e-02	1.13e-02	1.16e-01	1.18e-02	5.70e-02
$(100,80)$	1.31e-02	6.48e-02	1.56e-02	1.19e-01	1.68e-02	5.72e-02
$(100,100)$	1.54e-02	7.00e-02	1.87e-02	1.22e-01	1.93e-02	5.73e-02

The computational time of adversarial multi-class classification problem in Section 5.3 is reported in Table EC.4, with the basic statistics of classification datasets presented in Table EC.3. The results indicate that Sinkhorn DRO models have shorter computational time than Wasserstein DRO models in general. Note that we solve all baseline methods with stochastic algorithms. For large-scale datasets optimizing the log-sum-exp type loss for Sinkhorn DRO seems to be more efficient than solving the minimax game formulation for Wasserstein DRO.

Table EC.3 Basic statistics of adversarial multi-class logistic regression datasets.

	MNIST	CIFAR-10	tinyImageNet	STL-10
Image Size (before pre-processing)	784	3072	12288	27648
Feature Dimension (after pre-processing)	512	512	512	512
# of classes	10	10	200	10
Training Size	50000	50000	90000	5000
Testing Size	10000	10000	10000	8000

Table EC.4 Average computational time (in seconds) per problem instance for adversarial multi-class logistic regression problem.

Dataset	SAA	KL-DRO	1-WDRO	1-SDRO	2-WDRO	2-SDRO
MNIST	37.2	60.1	154	94.1	166	84.0
CIFAR-10	31.6	51.7	133	98.3	140	80.6
tinyImageNet	58.1	102	248	153	259	143
STL-10	3.42	5.15	13.5	10.1	14.2	8.61

EC.2.4. Coefficient of Prescriptiveness for Different Parameter(s) Combination

In this subsection, we report the coefficient of prescriptiveness for different parameter(s) combination on instances which are omitted in the main content. Specifically,

- Fig. EC.3 and EC.4 correspond to the omitted experiment results in Section 5.1.
- Fig. EC.5 and EC.6 correspond to the omitted experiment results in Section 5.2.
- Fig. EC.7 corresponds to the omitted experiment results in Section 5.3.

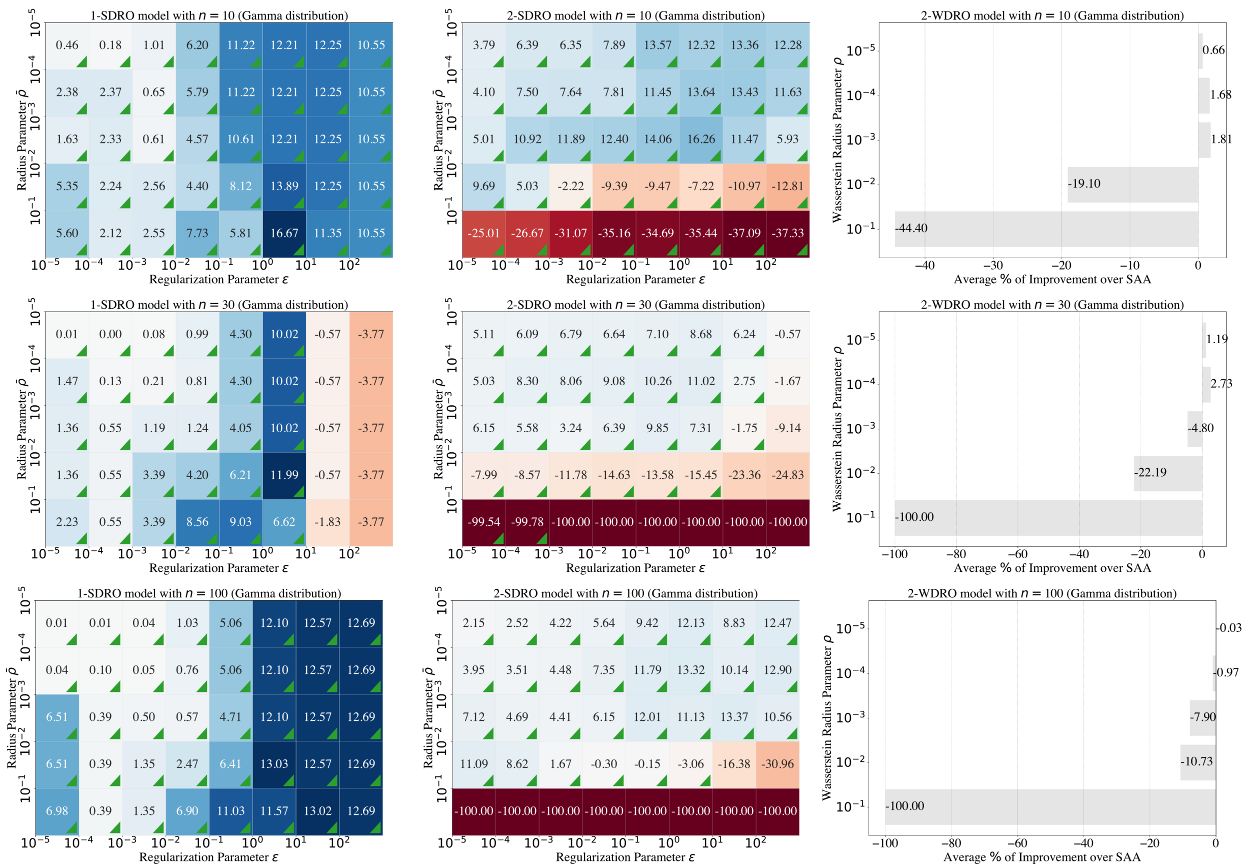

Figure EC.3 Experiment results of the newsvendor model for gamma data distribution. Details of these subplots follow the same setup from Figure 4.

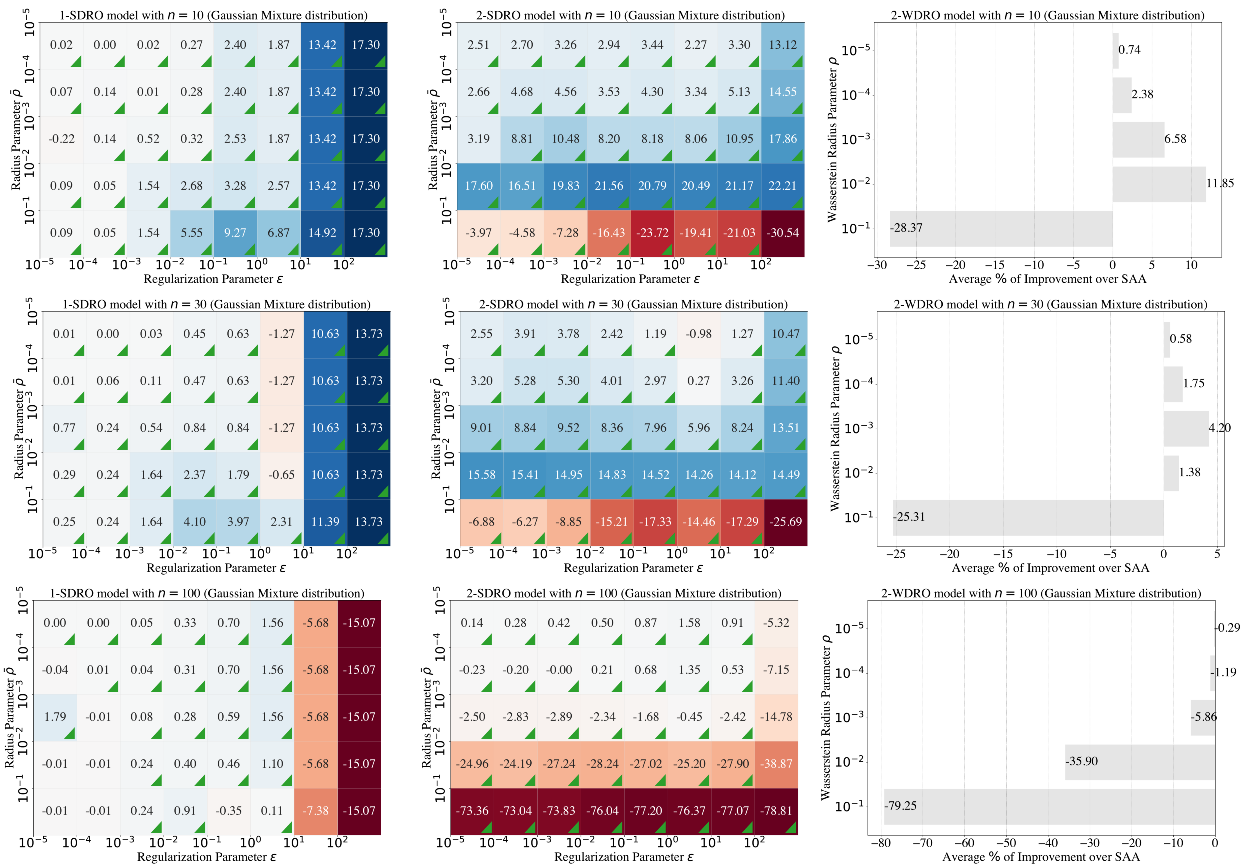

Figure EC.4 Experiment results of the newsvendor model for the mixture of truncated normal distributions. Details of these subplots follow the same setup from Figure 4.

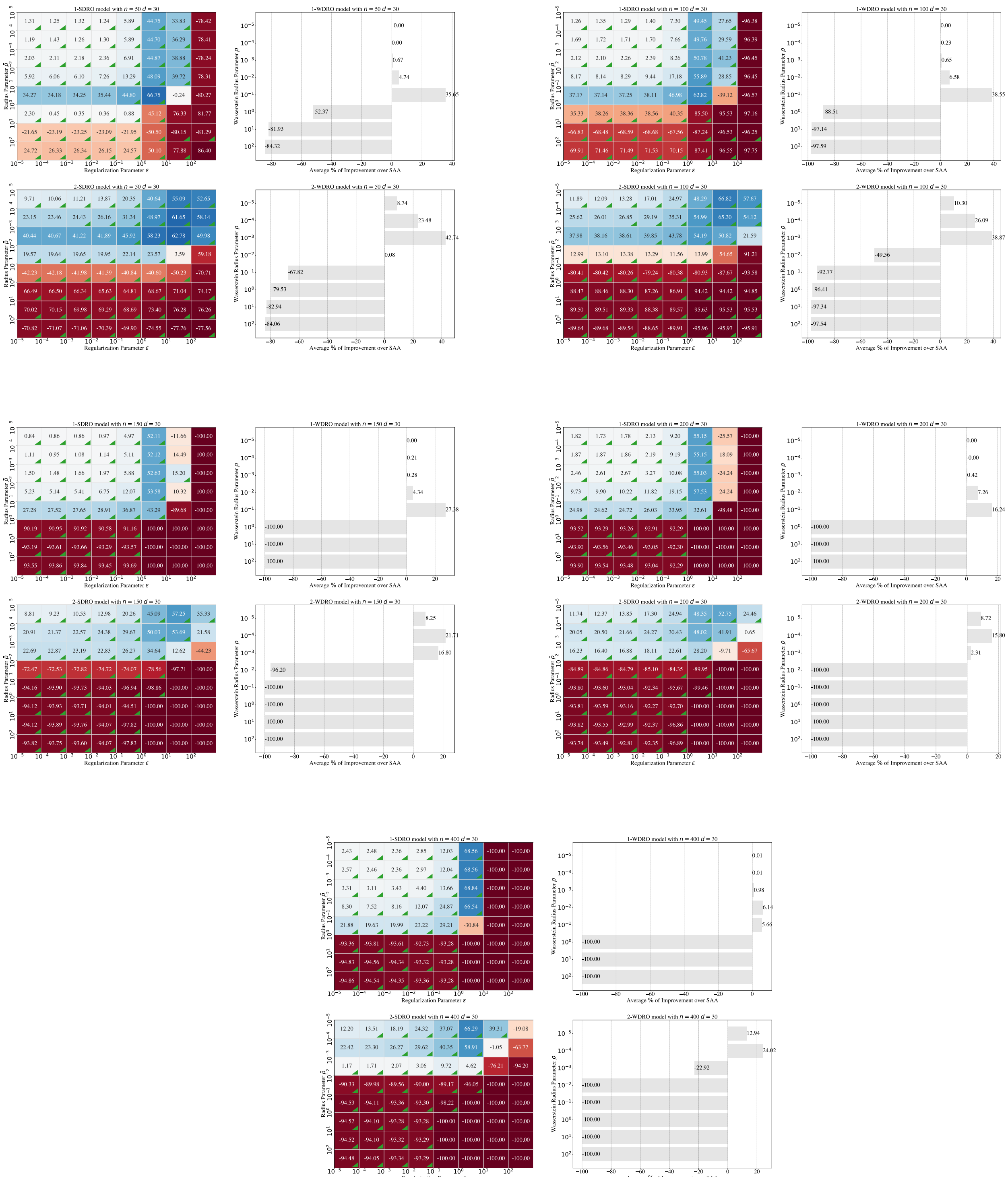

Figure EC.5 Additional experiment results of the portfolio optimization model for different data dimensions in heatmaps. Here we fix the data dimension $d = 30$ and for plots from left to right and from top to bottom, we vary the sample size $n \in \{50, 100, 150, 200, 400\}$. Details of these subplots follow the same setup from Fig. 6.

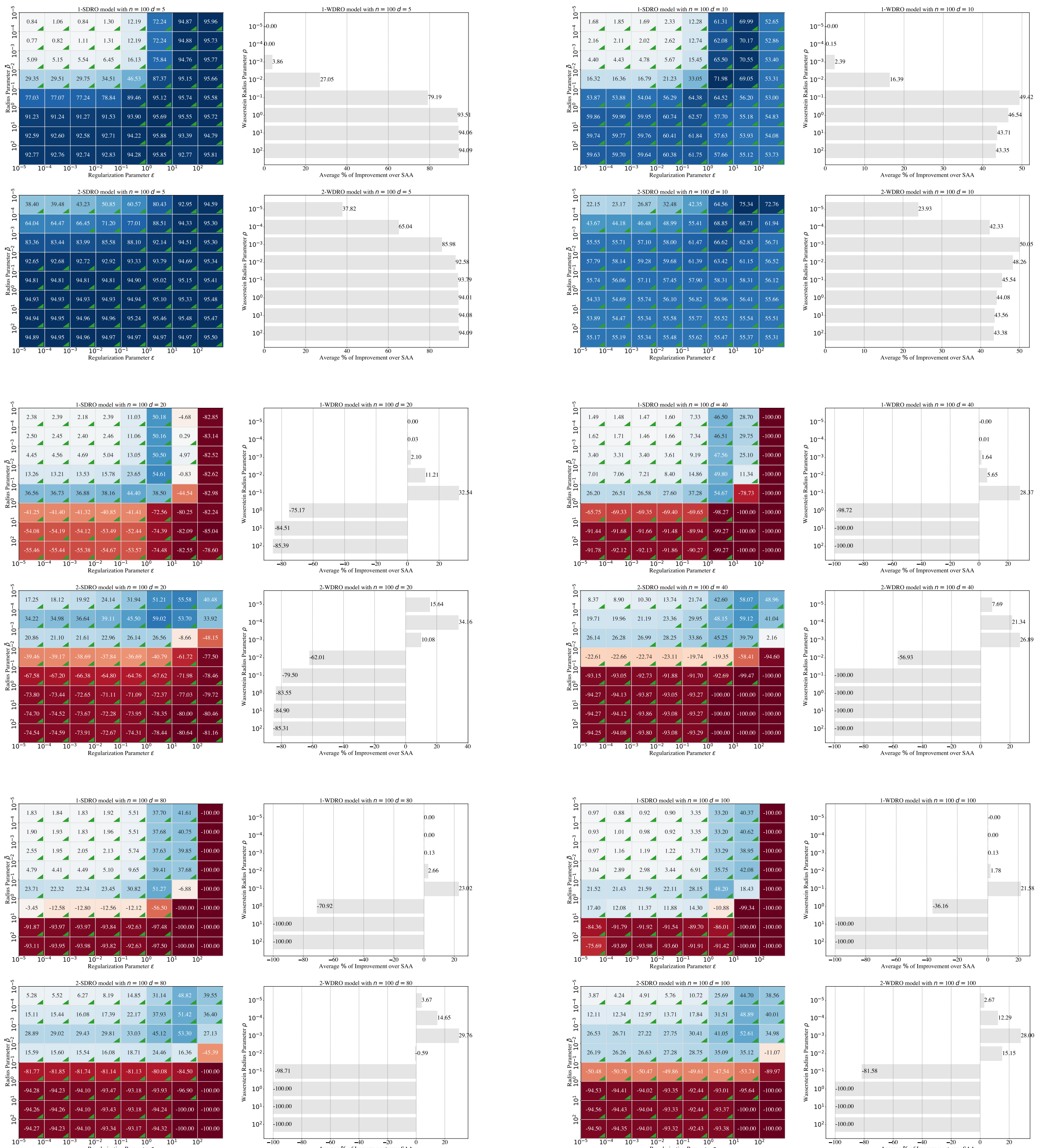

Figure EC.6 Additional experiment results of the portfolio optimization model for different data dimensions in heatmaps. Here we fix the sample size $n = 100$ and for plots from left to right and from top to bottom, we vary the data dimension $d \in \{5, 10, 20, 40, 80, 100\}$. Details of these subplots follow the same setup from Fig. 6.

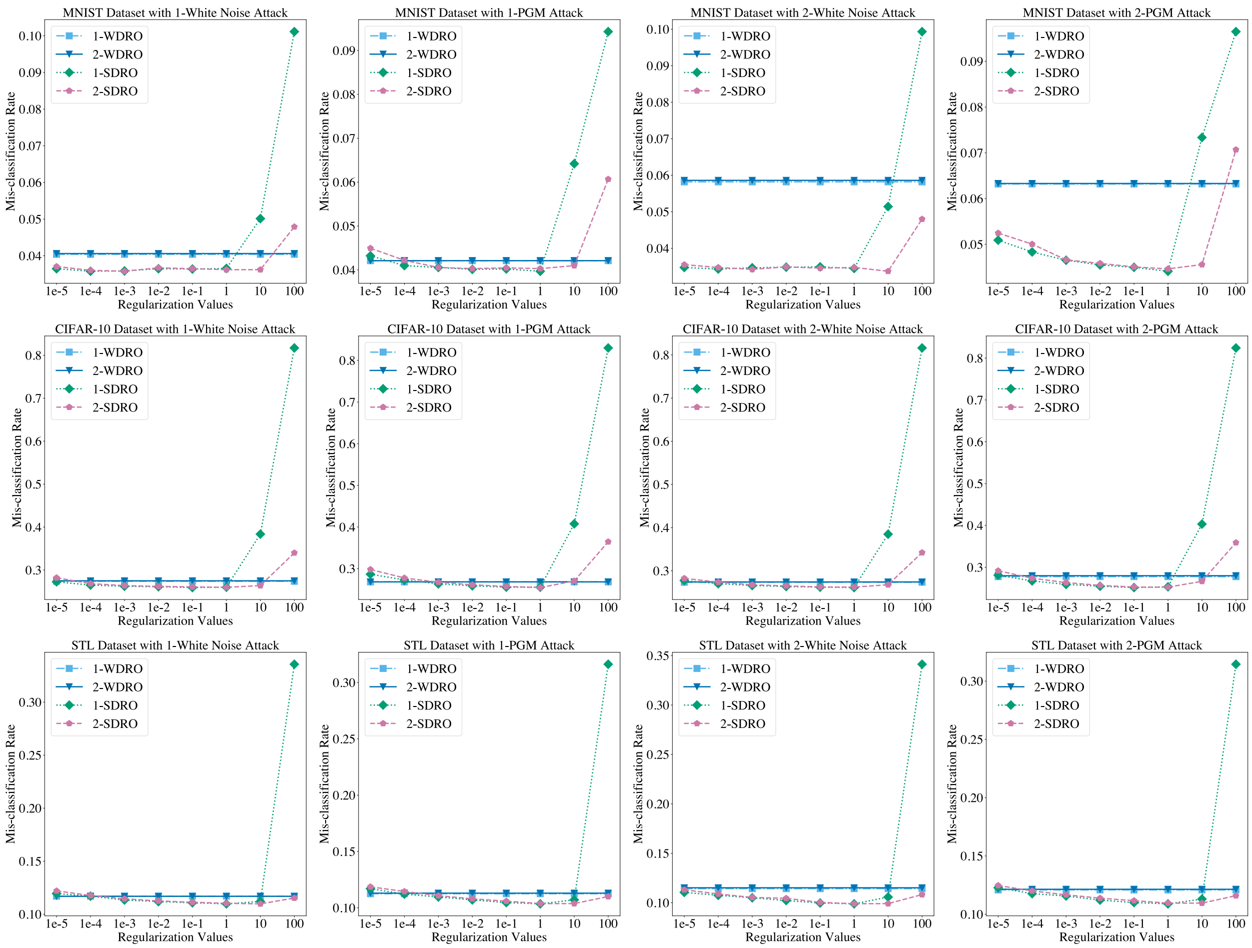

Figure EC.7 Additional experiment results of the adversarial classification problem for different datasets and different types of perturbations. Details of these subplots follow the same setup from Fig. 8.

EC.3. Sufficient Condition for Condition 1

PROPOSITION EC.1. *Condition 1 holds if there exists $p\geq 1$ so that the following conditions are satisfied:*

(I) *For any $x,y,z\in\mathcal{Z}$, $c(x,y)\geq 0$, and $(c(x,y))^{1/p}\leq (c(x,z))^{1/p}+(c(z,y))^{1/p}$.*

(II) *The nominal distribution $\widehat{\mathbb{P}}$ has a finite mean, denoted as $\overline{x}$. Moreover, $\nu\{z:\ 0\leq c(\overline{x},z)<\infty\}=1$ and $\Pr_{x\sim\widehat{\mathbb{P}}}\{c(x,\overline{x})<\infty\}=1$.*

(III) *Assumption 1(III) holds, and there exists $\lambda>0$ such that $\mathbb{E}_{z\sim\nu}\left[e^{f(z)/(\lambda\epsilon)}e^{-2^{1-p}c(\overline{x},z)/\epsilon}\right]<\infty$.*

We make some remarks for the sufficient conditions listed above. The first condition can be satisfied by taking the transport cost as the p-th power of the metric defined on $\mathcal{Z}$ for any $p\geq 1$. The second condition requires the nominal distribution $\widehat{\mathbb{P}}$ is finite almost surely, e.g., it can be a subguassian distribution with respect to the transport cost c. We first present an useful technical lemma before showing the proof of Proposition EC.1.

LEMMA EC.1. *Under the first condition of Proposition EC.1, for any $x\in\mathcal{Z}$, it holds that*

$$\mathbb{E}_{z\sim\nu}\left[e^{-c(x,z)/\epsilon}\right]\geq e^{-2^{p-1}c(x,\overline{x})/\epsilon}\mathbb{E}_{z\sim\nu}\left[e^{-2^{p-1}c(\overline{x},z)/\epsilon}\right].$$

Proof of Lemma EC.1. Based on the inequality $(a+b)^p\leq 2^{p-1}(a^p+b^p)$, we can see that

$$c(x,z)\leq (c(y,z)^{1/p}+c(z,y)^{1/p})^p\leq 2^{p-1}(c(y,z)+c(z,y)),\quad \forall x,y,z\in\mathcal{Z}.$$

Since $c(x,z)\leq 2^{p-1}(c(\overline{x},z)+c(x,\overline{x}))$, we can see that

$$\mathbb{E}_{z\sim\nu}\left[e^{-c(x,z)/\epsilon}\right]\geq \exp\left(-2^{p-1}c(x,\overline{x})/\epsilon\right)\mathbb{E}_{z\sim\nu}\left[e^{-2^{p-1}c(\overline{x},z)/\epsilon}\right].$$

The proof is completed. □

Proof of Proposition EC.1. One can see that for any $x\in\operatorname{supp}\widehat{\mathbb{P}}$, it holds that

$$\begin{aligned}
&\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right]=\mathbb{E}_{z\sim\nu}\left[e^{f(z)/(\lambda\epsilon)}\frac{e^{-c(x,z)/\epsilon}}{\mathbb{E}_{u\sim\nu}\left[e^{-c(x,u)/\epsilon}\right]}\right]\\
\leq&\mathbb{E}_{z\sim\nu}\left[e^{f(z)/(\lambda\epsilon)}\frac{e^{-c(x,z)/\epsilon}}{\mathbb{E}_{u\sim\nu}\left[e^{-2^{p-1}c(\overline{x},u)/\epsilon}\right]}\right]\leq\mathbb{E}_{z\sim\nu}\left[e^{f(z)/(\lambda\epsilon)}\frac{e^{-2^{1-p}c(\overline{x},z)/\epsilon}e^{c(x,\overline{x})/\epsilon}}{\mathbb{E}_{u\sim\nu}\left[e^{-2^{p-1}c(\overline{x},u)/\epsilon}\right]}\right]\\
=&\frac{e^{c(x,\overline{x})(1+2^{p-1})/\epsilon}}{\mathbb{E}_{u\sim\nu}\left[e^{-2^{p-1}c(\overline{x},u)/\epsilon}\right]}\mathbb{E}_{z\sim\nu}\left[e^{f(z)/(\lambda\epsilon)}e^{-2^{1-p}c(\overline{x},z)/\epsilon}\right],
\end{aligned}$$

where the first inequality is based on the lower bound in Lemma EC.1, the second inequality is based on the triangular inequality $c(x,z)\geq 2^{1-p}c(\overline{x},z)-c(x,\overline{x})$. Note that almost surely for all $x\in\operatorname{supp}\widehat{\mathbb{P}}$, $c(x,\overline{x})<\infty$. Moreover, $0<\mathbb{E}_{z\sim\nu}\left[e^{-2^{p-1}c(\overline{x},z)/\epsilon}\right]\leq\mathbb{E}_{z\sim\nu}\left[e^{-c(\overline{x},z)/\epsilon}\right]<\infty$, where the lower bound is because $c(\overline{x},z)<\infty$ almost surely for all z, the upper bound is because $c(\overline{x},z)\geq 0$ almost surely for all z. Based on these observations, we have that

$$\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right]\leq\frac{e^{c(x,\overline{x})(1+2^{p-1})/\epsilon}}{\mathbb{E}_{z\sim\nu}\left[e^{-2^{p-1}c(\overline{x},z)/\epsilon}\right]}\mathbb{E}_{z\sim\nu}\left[e^{f(z)/(\lambda\epsilon)}e^{-2^{1-p}c(\overline{x},z)/\epsilon}\right]<\infty$$

almost surely for all $x\sim\widehat{\mathbb{P}}$. □

EC.4. Proofs of Technical Results in Section 3.2 and 3.3

Proof of Remark 3. Recall the dual objective function in (1) is

$$v(\lambda;\epsilon)=\lambda\rho+\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\Big[\lambda\epsilon\,\log\mathbb{E}_{z\sim\nu}\Big[e^{(f(z)-\lambda c(x,z))/(\lambda\epsilon)}\Big]\Big].$$

We take limit for the second term in $v(\lambda;\epsilon)$ to obtain:

$$\begin{aligned}&\lim_{\epsilon\to 0}\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\Big[\lambda\epsilon\,\log\mathbb{E}_{z\sim\nu}\Big[e^{(f(z)-\lambda c(x,z))/(\lambda\epsilon)}\Big]\Big]=\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\lim_{\beta\to\infty}\frac{\lambda}{\beta}\log\mathbb{E}_{z\sim\nu}\Big[e^{(f(z)-\lambda c(x,z))\beta/\lambda}\Big]\right]\\
=&\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\lim_{\beta\to\infty}\lambda\nabla_\beta\log\mathbb{E}_{z\sim\nu}\Big[e^{(f(z)-\lambda c(x,z))\beta/\lambda}\Big]\right]=\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\lim_{\beta\to\infty}\frac{\mathbb{E}_{z\sim\nu}\Big[e^{(f(z)-\lambda c(x,z))\beta/\lambda}\big[f(z)-\lambda c(x,z)\big]\Big]}{\mathbb{E}_{z\sim\nu}\Big[e^{(f(z)-\lambda c(x,z))\beta/\lambda}\Big]}\right]\\
=&\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\Big[\sup_{z\in\operatorname{supp}\nu}\big\{f(z)-\lambda c(x,z)\big\}\Big].\end{aligned}$$

Particularly, when $\operatorname{supp}\nu=\mathcal{Z}$, it holds that

$$\sup_{z\in\operatorname{supp}\nu}\big\{f(z)-\lambda c(x,z)\big\}=\sup_{z\in\mathcal{Z}}\big\{f(z)-\lambda c(x,z)\big\}$$

and in this case the dual objective function of the Sinkhorn DRO problem converges into that of the Wasserstein DRO problem. □

Proof of Example 2. In this example, the dual objective becomes

$$V_{\mathrm{D}}=\inf_{\lambda\ge 0}\left\{\lambda\overline{\rho}+\mathbb{E}_{(a,b)\sim\widehat{\mathbb{P}}}\left[\lambda\epsilon\log\mathbb{E}_{a'\sim\mathcal{N}(a,\epsilon I_d)}\left[\exp\left(\frac{(\theta^{\mathrm{T}}a'-b)^2}{\lambda\epsilon}\right)\right]\right]\right\}.\tag{EC.2}$$

Specially, for any $a\in\mathbb{R}^d,b\in\mathbb{R},\theta\in\mathbb{R}^d$, it holds that

$$\begin{aligned}&\lambda\epsilon\log\left(\mathbb{E}_{a'\sim\mathcal{N}(a,\epsilon I_d)}\exp\left(\frac{(\theta^{\mathrm{T}}a'-b)^2}{\lambda\epsilon}\right)\right)=\lambda\epsilon\log\left(\mathbb{E}_{\Delta_a\sim\mathcal{N}(0,I_d)}\exp\left(\frac{\big[(\theta^{\mathrm{T}}a-b)+(\sqrt{\epsilon}\theta)^{\mathrm{T}}\Delta_a\big]^2}{\lambda\epsilon}\right)\right)\\
=&(\theta^{\mathrm{T}}a-b)^2+\lambda\epsilon\log\Bigg(\underbrace{\mathbb{E}_{\Delta_a\sim\mathcal{N}(0,I_d)}\exp\left(\frac{\epsilon(\theta^{\mathrm{T}}\Delta_a)^2-2(b-\theta^{\mathrm{T}}a)\sqrt{\epsilon}\theta^{\mathrm{T}}\Delta_a}{\lambda\epsilon}\right)}_{\text{(I)}}\Bigg).\end{aligned}$$

The term (I) can be simplified using the integral of exponential functions method:

$$\text{(I)}=\begin{cases}\det\left(I-\dfrac{2\theta\theta^{\mathrm{T}}}{\lambda}\right)^{-1/2}\exp\left(2\dfrac{(\theta^{\mathrm{T}}a-b)^2}{\lambda^2\epsilon}\theta^{\mathrm{T}}A^{-1}\theta\right), & \text{when } \|\theta\|_2^2<\dfrac{\lambda}{2},\\ \infty, & \text{otherwise,}\end{cases}$$

where the matrix $A=I-\frac{2\theta\theta^{\mathrm{T}}}{\lambda}$. Finally, we obtain that if $\|\theta\|_2^2<\frac{\lambda}{2}$,

$$\lambda\epsilon\log\left(\mathbb{E}_{a'\sim\mathcal{N}(a,\epsilon I_d)}\exp\left(\frac{(\theta^{\mathrm{T}}a'-b)^2}{\lambda\epsilon}\right)\right)=(\theta^{\mathrm{T}}a-b)^2+\frac{(\theta^{\mathrm{T}}a-b)^2}{\frac{1}{2}\lambda\|\theta\|_2^{-2}-1}-\frac{\lambda\epsilon}{2}\log\det\left(I-\frac{2\theta\theta^{\mathrm{T}}}{\lambda}\right).$$

Substituting this expression into (EC.2) gives the desired result. □

Proof of Corollary 1. We now introduce the epi-graphical variables $s_i, i = 1, \ldots, n$ to reformulate V_{D} as

$$V_{\mathrm{D}} = \begin{cases} \inf_{\lambda \ge 0, s_i} & \lambda\overline{\rho} + \frac{1}{n}\sum_{i=1}^n s_i \\ \text{s.t.} & \lambda\epsilon \log \mathbb{E}_{z\sim\mathbb{Q}_{\hat{x}_i,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right] \le s_i, \forall i \end{cases}$$

For fixed i, the i-th constraint can be reformulated as

$$\begin{aligned} &\left\{\exp\left(\frac{s_i}{\lambda\epsilon}\right) \ge \mathbb{E}_{z\sim\mathbb{Q}_{\hat{x}_i,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right]\right\} = \left\{1 \ge \mathbb{E}_{z\sim\mathbb{Q}_{\hat{x}_i,\epsilon}}\left[e^{(f(z)-s_i)/(\lambda\epsilon)}\right]\right\} \\ =&\left\{\lambda\epsilon \ge \mathbb{E}_{z\sim\mathbb{Q}_{\hat{x}_i,\epsilon}}\left[\lambda\epsilon e^{(f(z)-s_i)/(\lambda\epsilon)}\right]\right\} \\ =&\left\{\lambda\epsilon \ge \sum_{\ell=1}^{L_{\max}} \mathbb{Q}_{\hat{x}_i,\epsilon}(z_\ell) a_{i,\ell}\right\} \bigcap \left\{a_{i,\ell} \ge \lambda\epsilon\exp\left(\frac{f(z_\ell)-s_i}{\lambda\epsilon}\right), \forall \ell\right\}, \end{aligned}$$

where the second constraint set can be formulated as $(\lambda\epsilon, a_{i,\ell}, f(z_\ell) - s_i) \in \mathcal{K}_{\exp}$. Substituting this expression into V_{D} completes the proof. □

EC.5. Proofs of Technical Results in Section 3.4

We rely on the following technical lemma to derive our strong duality result.

LEMMA EC.2. *((Hu and Hong, 2012, Section 2.1) or (Shapiro, 2017)) For fixed τ and a reference measure $\nu \in \mathcal{M}(\mathcal{Z})$, consider the optimization problem*

$$v(\tau) = \sup_{\mathbb{P}\in\mathcal{P}(\mathcal{Z})} \left\{\mathbb{E}_{z\sim\mathbb{P}}\left[f(z) - \tau\log\left(\frac{\mathrm{d}\mathbb{P}(z)}{\mathrm{d}\nu(z)}\right)\right]\right\}. \tag{EC.3}$$

Suppose there exists a probability measure $\mathbb{Q} \in \mathcal{P}(\mathcal{Z})$ such that $\mathbb{Q} \ll \nu$.

(I) *When $\tau = 0$,*

$$v(0) = \operatorname*{ess\,sup}_{\nu}(f) \triangleq \inf\{t \in \mathbb{R} : \nu\{f(z) > t\} = 0\}.$$

(II) *When $\tau > 0$ and*

$$\mathbb{E}_{z\sim\nu}\left[e^{f(z)/\tau}\right] < \infty,$$

it holds that

$$v(\tau) = \tau\log\left(\mathbb{E}_{z\sim\nu}\left[e^{f(z)/\tau}\right]\right),$$

and $\lim_{\tau\downarrow 0} v(\tau) = v(0)$. The optimal solution in (EC.3) has the expression

$$\mathrm{d}\mathbb{P}(z) = \frac{e^{f(z)/\tau}}{\mathbb{E}_{u\sim\nu}\left[e^{f(u)/\tau}\right]}\mathrm{d}\nu(z).$$

(III) *When $\tau > 0$ and*

$$\mathbb{E}_{z\sim\nu}\left[e^{f(z)/\tau}\right] = \infty,$$

we have that $v(\tau) = \infty$.

LEMMA EC.3 (**Measurability of** $v_x(\lambda)$**).** *Assume Assumptions 1(I), 1(II), 1(III) hold. For fixed* $\lambda \geq 0$, *define the function* $v_x(\lambda)\colon \operatorname{supp}\widehat{\mathbb{P}} \to \mathbb{R}\cup\{+\infty\}$ *as*

$$v_x(\lambda) = \sup_{\gamma_x\in\mathcal{P}(\mathcal{Z})}\left\{\mathbb{E}_{z\sim\gamma_x}\left[f(z)-\lambda c(x,z)-\lambda\epsilon\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\nu(z)}\right)\right]\right\}.$$

The function $v_x(\lambda)$ *is measurable with respect to* $x\sim\widehat{\mathbb{P}}$ *regardless of the choice of* $\lambda\geq 0$.

Proof of Lemma EC.3. When $\lambda = 0$, by Lemma EC.2, it holds that

$$v_x(\lambda) = \operatorname*{ess\,sup}_{\nu}(f),$$

which is a constant independent of x, which is clearly measurable. When $\lambda > 0$ and satisfies Condition 1, by Lemma EC.2, it holds that

$$v_x(\lambda) = \lambda\epsilon\log\mathbb{E}_{z\sim\nu}\left[e^{(f(z)-\lambda c(x,z))/(\lambda\epsilon)}\right] < \infty.$$

As loss function f and cost function c are both measurable, by conditioning Lemma (Kallenberg, 1997, Lemma 2.11), $v_x(\lambda)$ is measurable. When $\lambda > 0$ such that the event

$$E = \left\{x:\ \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right] = \infty\right\} = \left\{x:\ \mathbb{E}_{z\sim\nu}\left[e^{(f(z)-\lambda c(x,z))/(\lambda\epsilon)}\right] = \infty\right\}$$

satisfies $\widehat{\mathbb{P}}(E) > 0$, by Lemma EC.2, it holds that

$$v_x(\lambda) = \begin{cases}\lambda\epsilon\log\mathbb{E}_{z\sim\nu}\left[e^{(f(z)-\lambda c(x,z))/(\lambda\epsilon)}\right] < \infty, & \text{if } x\in E^c,\\ \infty, & \text{if } x\in E.\end{cases}$$

For fixed $\alpha\in\mathbb{R}$, the level set

$$\{x:\ v_x(\lambda)\geq\alpha\} = \{x\in E^c:\ v_x(\lambda)\geq\alpha\}\cup E = \left\{x\in E^c:\ \lambda\epsilon\log\mathbb{E}_{z\sim\nu}\left[e^{(f(z)-\lambda c(x,z))/(\lambda\epsilon)}\right]\geq\alpha\right\}\cup E,$$

which is clearly a measurable set, and therefore $v_x(\lambda)$ is measurable. The proof is completed. □

Proof of Lemma 2. Recall from (6) that

$$V = \sup_{\{\gamma_x\}_{x\in\operatorname{supp}\widehat{\mathbb{P}}}\subset\mathcal{P}(\mathcal{Z})}\left\{\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_x}[f(z)]:\ \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_x}\left[c(x,z)+\epsilon\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\nu(z)}\right)\right]\leq\rho\right\}.$$

Based on the change-of-measure identity $\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\nu(z)}\right) = \log\left(\frac{\mathrm{d}\mathbb{Q}_{x,\epsilon}(z)}{\mathrm{d}\nu(z)}\right)+\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\mathbb{Q}_{x,\epsilon}(z)}\right)$ and the expression of $\mathbb{Q}_{x,\epsilon}$, the constraint can be reformulated as

$$\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_x}\left[c(x,z)+\epsilon\log\left(\frac{e^{-c(x,z)/\epsilon}}{\int e^{-c(x,u)/\epsilon}\,\mathrm{d}\nu(u)}\right)+\epsilon\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\mathbb{Q}_{x,\epsilon}(z)}\right)\right]\leq\rho.$$

Combining the first two terms within the expectation term and substituting the expression of $\overline{\rho}$, it is equivalent to

$$\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_x}\left[\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\mathbb{Q}_{x,\epsilon}(z)}\right)\right]\le\overline{\rho}.$$

In summary, the primal problem (Primal) can be reformulated as a generalized KL-divergence DRO problem

$$V=\sup_{\{\gamma_x\}_{x\in\operatorname{supp}\widehat{\mathbb{P}}}\subset\mathcal{P}(\mathcal{Z})}\left\{\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_x}[f(z)]:\ \epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_x}\left[\log\left(\frac{\mathrm{d}\gamma_x(z)}{\mathrm{d}\mathbb{Q}_{x,\epsilon}(z)}\right)\right]\le\overline{\rho}\right\}. \qquad \square$$

In the remaining of this subsection, we provide the full proof of Theorem 1. We first show that the dual minimizer exists.

LEMMA EC.4 (**Existence of Dual Minimizer**). *Suppose $\overline{\rho}>0$ and Condition 1 is satisfied, then the dual minimizer λ^* exists, which either equals to 0 or satisfies Condition 1.*

Proof of Lemma EC.4. We first show that $\lambda^*<\infty$. Denote by $v(\lambda)$ the objective function for the dual problem:

$$v(\lambda)=\lambda\overline{\rho}+\lambda\epsilon\,\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\Big[\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\big[e^{f(z)/(\lambda\epsilon)}\big]\Big].$$

The integrability condition for the dominated convergence theorem is satisfied, which implies

$$\begin{aligned}&\lim_{\lambda\to\infty}\lambda\epsilon\,\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\Big[\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\big[e^{f(z)/(\lambda\epsilon)}\big]\Big]=\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\lim_{\beta\to0}\frac{\epsilon}{\beta}\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\big[e^{\beta f(z)/\epsilon}\big]\right]\\=&\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\lim_{\beta\to0}\epsilon\nabla_\beta\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\big[e^{\beta f(z)/\epsilon}\big]\right]=\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\lim_{\beta\to0}\frac{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\big[f(z)e^{\beta f(z)/\epsilon}\big]}{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\big[e^{\beta f(z)/\epsilon}\big]}\right]\\=&\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[f(z)],\end{aligned}$$

where the first equality follows from the change-of-variable technique with $\beta=1/\lambda$, the second equality follows from the definition of derivative, the third and the last equality follows from the dominated convergence theorem. As a consequence, as long as $\overline{\rho}>0$, we have $\lim_{\lambda\to\infty}v(\lambda)=\infty$. We can take λ satisfying Condition 1 and then $v(\lambda)<\infty$. This, toegther with the fact that $v(\cdot)$ is continuous, guarantees the existence of the dual minimizer. Hence $\lambda^*<\infty$, which implies that either $\lambda^*=0$ or λ^* satisfies Condition 1. $\square$

Next, we establish first-order optimality condition for cases $\lambda^*>0$ or $\lambda^*=0$, corresponding to whether the Sinkhorn distance constraint in (Primal) is binding or not. Lemma EC.5 below presents a necessary and sufficient condition for the dual minimizer $\lambda^*=0$, corresponding to the case where the Sinkhorn distance constraint in (Primal) is not binding.

LEMMA EC.5 (**Necessary and Sufficient Condition for $\lambda^*=0$**). *Suppose $\overline{\rho}>0$ and Condition 1 is satisfied, then the dual minimizer $\lambda^*=0$ if and only if all the following conditions hold:*

(I) $\operatorname{ess\,sup}_\nu\ f\triangleq\inf\{t:\ \nu\{f(z)>t\}=0\}<\infty$.

(II) $\overline{\rho}' = \overline{\rho} + \epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)]\right] \geq 0$, *where* $A := \{z:\ f(z) = \operatorname{ess\,sup}_\nu\ f\}$.

Recall that we have the convention that the dual objective evaluated at $\lambda = 0$ equals $\operatorname{ess\,sup}_\nu\ f$. Thus Condition (I) ensures that the dual objective function evaluated at the minimizer is finite. When the minimizer $\lambda^* = 0$, the Sinkhorn ball should be large enough to contain at least one distribution with objective value $\operatorname{ess\,sup}_\nu\ f$, and Condition (II) characterizes the lower bound of $\overline{\rho}$.

Proof of Lemma EC.5. Suppose the dual minimizer $\lambda^* = 0$, then taking the limit of the dual objective function gives

$$\lim_{\lambda\to 0} v(\lambda) = \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\big[H^u(x)\big] < \infty,$$

where $H^u(x) := \inf\{t:\ \mathbb{Q}_{x,\epsilon}\{f(z) > t\} = 0\} \triangleq \operatorname*{ess\,sup}_{\mathbb{Q}_{x,\epsilon}} f$. For notational simplicity we take $H^u = \operatorname*{ess\,sup}_{\nu} f$. One can check that $H^u(x) \equiv H^u$ for any $x \in \operatorname{supp}\widehat{\mathbb{P}}$: for any t so that $\mathbb{Q}_{x,\epsilon}\{f(z) > t\} = 0$, we have that

$$\mathbb{E}_{z\sim\nu}\left[1\{f(z) > t\}e^{-c(x,z)/\epsilon}\right] = 0,$$

which, together with the fact that $\nu\{c(x,z) < \infty\} = 1$ for fixed x, implies

$$\mathbb{E}_{z\sim\nu}[1\{f(z) > t\}] = 0.$$

On the contrary, for any t so that $\nu\{f(z) > t\} = 0$, we have that

$$0 \leq \mathbb{E}_{z\sim\nu}\left[1\{f(z) > t\}e^{-c(x,z)/\epsilon}\right] \leq \mathbb{E}_{z\sim\nu}[1\{f(z) > t\}] = 0,$$

where the second inequality is because that $\nu\{c(x,z) \geq 0\} = 1$. As a consequence, $\mathbb{Q}_{x,\epsilon}\{f(z) > t\} = 0$. Hence we can assert that $H^u(x) = H^u$ for all $x \in \operatorname{supp}\widehat{\mathbb{P}}$, which implies

$$\lim_{\lambda\to 0} v(\lambda) = H^u < \infty.$$

Then we show that almost surely for all x,

$$\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] > 0, \quad \text{where } A = \{z:\ f(z) = H^u\}.$$

Denote by D the collection of samples x so that $\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] = 0$. Assume the condition above does not hold, which means that $\widehat{\mathbb{P}}\{D\} > 0$. For any $\tau > 0$ and $x \in D$, there exists $H^l(x) < H^u$ such that

$$0 < \mathfrak{h}_x := \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_{B(x)}(z)] \leq \tau, \quad \text{where } B(x) = \{z:\ H^l(x) \leq f(z) \leq H^u\}.$$

Define $H^{\text{gap}}(x) = H^u - H^l(x)$, $\mathfrak{h}_x^c = 1 - \mathfrak{h}_x$. Then we find that for $x \in D$,

$$\begin{aligned} v_x(\lambda) &= \lambda\epsilon\log\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}1_{B(x)}(z)\right] + \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}1_{B(x)^c}(z)\right]\right) \\ &\leq H^u + \lambda\epsilon\log\left(\mathfrak{h}_x + e^{-H^{\text{gap}}(x)/(\lambda\epsilon)}\mathfrak{h}_x^c\right). \end{aligned}$$

Since $\widehat{\mathbb{P}}\{D\} > 0$, the dual objective function for $\lambda > 0$ is upper bounded as

$$\begin{aligned} v(\lambda) &= \lambda\overline{\rho} + \mathbb{E}_{x\sim\widehat{\mathbb{P}}}[v_x(\lambda)] \\ &\le H^u + \lambda\overline{\rho} + \lambda\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathfrak{h}_x + e^{-H^{\text{gap}}(x)/(\lambda\epsilon)}\mathfrak{h}_x^c\right)1_D(x)\right]. \end{aligned}$$

We can see that

$$\lim_{\lambda\to 0}\ \lambda\overline{\rho} + \lambda\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathfrak{h}_x + e^{-H^{\text{gap}}(x)/(\lambda\epsilon)}\mathfrak{h}_x^c\right)1_D(x)\right] = 0,$$

and

$$\begin{aligned} &\lim_{\lambda\to 0}\ \nabla\left[\lambda\overline{\rho} + \lambda\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathfrak{h}_x + e^{-H^{\text{gap}}(x)/(\lambda\epsilon)}\mathfrak{h}_x^c\right)1_D(x)\right]\right] \\ =&\overline{\rho} + \epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathfrak{h}_x\right)1_D(x)\right] \le \overline{\rho} + \epsilon\log(\tau)\widehat{\mathbb{P}}\{D\} \le -\overline{\rho} < 0, \end{aligned}$$

where the second inequality is by taking the constant $\tau = \exp\left(-\frac{2\overline{\rho}}{\epsilon\widehat{\mathbb{P}}\{D\}}\right)$. Hence, there exists $\overline{\lambda} > 0$ such that

$$v(\overline{\lambda}) \le H^u + \overline{\lambda}\overline{\rho} + \overline{\lambda}\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathfrak{h}_x + e^{-H^{\text{gap}}(x)/(\overline{\lambda}\epsilon)}\mathfrak{h}_x^c\right)1_D(x)\right] < v(0),$$

which contradicts to the optimality of $\lambda^* = 0$. As a result, almost surely for all x, we have that

$$\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] > 0.$$

To show the second condition, we re-write the dual objective function for $\lambda > 0$ as

$$v(\lambda) = \lambda\overline{\rho} + \lambda\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)\right]\right)\right] + H^u.$$

The gradient of $v(\lambda)$ becomes

$$\begin{aligned} \nabla v(\lambda) = &\overline{\rho} + \epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)\right]\right)\right] \\ &+ \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\frac{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)(H^u - f(z))/(\lambda)\right]}{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)\right]}\right]. \end{aligned}$$

We can see that $\lim_{\lambda\to\infty}\nabla v(\lambda) = \overline{\rho}$. Take

$$v_{1,x}(\lambda) = \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)\right].$$

Then $\lim_{\lambda\to 0} v_{1,x}(\lambda) = 0$ and $v_{1,x}(\lambda) \ge 0$. Take

$$v_{2,x}(\lambda) = \frac{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)(H^u - f(z))/(\lambda)\right]}{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{[f(z)-H^u]/(\lambda\epsilon)}1_{A^c}(z)\right]}.$$

Then $\lim_{\lambda\to 0} v_{2,x}(\lambda) = 0$ and $v_{2,x}(\lambda) \ge 0$. It follows that

$$\lim_{\lambda\to 0}\nabla v(\lambda) = \overline{\rho} + \epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)]\right] = \overline{\rho}'.$$

Hence, if the last condition is violated, based on the mean value theorem, we can find $\overline{\lambda} > 0$ so that $\nabla v(\overline{\lambda}) = 0$, which contradicts to the optimality of $\lambda^* = 0$.

Now we show the converse direction. For any $\lambda > 0$, we find that

$$\nabla v(\lambda) = \overline{\rho} + \epsilon \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + v_{1,x}(\lambda)\right)\right] + \mathbb{E}_{x\sim\widehat{\mathbb{P}}}[v_{2,x}(\lambda)].$$

For fixed x, when $\mathbb{E}_{\mathbb{Q}_{x,\epsilon}}[1_A] = 1$, we can see that $v_{1,x}(\lambda) = v_{2,x}(\lambda) = 0$, then

$$\overline{\rho} + \epsilon\left[\log\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + v_{1,x}(\lambda)\right)\right] + v_{2,x}(\lambda) = \overline{\rho} > 0.$$

When $\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] \in (0,1)$, we can see that $v_{1,x}(\lambda) > 0, v_{2,x}(\lambda) > 0$. Then

$$\overline{\rho} + \epsilon\left[\log\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)] + v_{1,x}(\lambda)\right)\right] + v_{2,x}(\lambda) > \overline{\rho} + \epsilon\log(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[1_A(z)]) = \overline{\rho}' \geq 0.$$

Therefore, $\nabla v(\lambda) > 0$ for any $\lambda > 0$. By the convexity of $v(\lambda)$, the dual minimizer $\lambda^* = 0$. □

Proof of Lemma 3. Recall that $v(\lambda)$ denotes the objective function for the dual problem. The optimality condition can be derived by taking $\nabla_\lambda\, v(\lambda)\mid_{\lambda=\lambda^*} = 0$. To show the uniqueness of λ^*, we find that

$$\begin{aligned}&\nabla^2_\lambda v(\lambda)\\ =&\frac{1}{\lambda^3\epsilon}\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[e^{f(z)/(\lambda\epsilon)}]\right)^{-2}\cdot\left(\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[e^{f(z)/(\lambda\epsilon)}f^2(z)]\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[e^{f(z)/(\lambda\epsilon)}] - \left\{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[e^{f(z)/(\lambda\epsilon)}f(z)]\right\}^2\right)\right].\end{aligned}$$

It can be shown by the Cauchy-Schwarz inequality that $\nabla^2_\lambda v(\lambda) \geq 0$ for any $\lambda > 0$, and the equality holds if and only if $f(\cdot)$ is a constant. If it is the case, the dual objective $v(\lambda)$ has the unique minimizer $\lambda^* = 0$, which contradicts to our assumption. Hence, strict convexity holds for the dual objective and it implies the uniqess of λ^*. □

Proof of Theorem 1. Recall the feasibility result in Theorem 1(I) can be easily shown by considering the reformulation of V in Lemma 2 and the non-negativity of KL-divergence. When $\overline{\rho} = 0$, one can see that

$$\begin{aligned}V_{\mathrm{D}} &= \inf_{\lambda\geq 0}\left\{\lambda\epsilon\,\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right]\right]\right\}\\ &\leq \lim_{\lambda\to\infty}\ \lambda\epsilon\,\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f(z)/(\lambda\epsilon)}\right]\right] = \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}[f(z)] = V.\end{aligned}$$

Therefore, the strong duality result holds in this case. Theorem 1(IV) can be shown by Lemma EC.5. It remains to show the strong duality result for $\overline{\rho} > 0$, which can be further separated to two cases: Condition 1 holds or not.

- When Condition 1 holds, by Lemma EC.4, the dual minimizer λ^* exists. The proof for $\lambda^* > 0$ can be found in main context. When $\lambda^* = 0$, the optimality condition in Lemma EC.5 holds. We construct the primal (approximate) solution $\mathbb{P}_* = \mathrm{Proj}_{2\#}\gamma_*$, where γ_* satisfies

$$\mathrm{d}\gamma_*(x,z) = \mathrm{d}\gamma_*^x(z)\,\mathrm{d}\widehat{\mathbb{P}}(x),\quad \text{where } \mathrm{d}\gamma_*^x(y) = \begin{cases}0, & \text{if } z\notin A,\\ \frac{e^{-c(x,z)/\epsilon}\,\mathrm{d}\nu(z)}{\mathbb{E}_{u\sim\nu}\left[e^{-c(x,u)/\epsilon}1_A\right]}, & \text{if } z\in A.\end{cases}$$

We can verify easily that the primal solution is feasible based on the optimality condition $\overline{\rho}' \geq 0$ in Lemma EC.5. Moreover, we can check that the primal optimal value is lower bounded by the dual optimal value:

$$V \geq \mathbb{E}_{(x,z)\sim\gamma_*}[f(z)] = \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_*^x}[f(z)] = \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\mathbb{E}_{z\sim\gamma_*^x}\left[\operatorname*{ess\,sup}_{\nu} f\right] = \operatorname*{ess\,sup}_{\nu} f = V_{\mathrm{D}},$$

where the second equality is because that $z \in A$ so that $f(z) = \operatorname*{ess\,sup}_{\nu} f$. This, together with the weak duality result, completes the proof in this part.

- When Condition 1 does not hold, we consider a sequence of real numbers $\{R_j\}_j$ such that $R_j \to \infty$ and take the objective function $f_j(z) = f(z)1\{f(z) \leq R_j\}$. Hence, there exists $\lambda > 0$ satisfying $\Pr_{x\sim\widehat{\mathbb{P}}}\left\{x : \mathbb{E}_{\mathbb{Q}_{x,\epsilon}}\left[e^{f_j(z)/(\lambda\epsilon)}\right] = \infty\right\} = 0$. According to the necessary condition in Lemma EC.5, the corresponding dual minimizer $\lambda_j^* > 0$ for sufficiently large index j. Then we can apply the duality result in the first part of Theorem 1(III) to show that for sufficiently large j, it holds that

$$\sup_{\mathbb{P}\in\mathbb{B}_{\rho,\epsilon}(\widehat{\mathbb{P}})} \{\mathbb{E}_{z\sim\mathbb{P}}[f_j(z)]\} \geq \lambda_j^*\overline{\rho} + \lambda_j^*\epsilon\mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\log \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_j(z)/(\lambda\epsilon)}\right]\right].$$

Taking $j \to \infty$ both sides implies that $V = \infty$. □

EC.6. Proof of Theorem 2 in Section 4.2.1

In this section, we omit the dependence of λ when defining objective or subgradient terms, e.g., we write $F(\theta)$ for $F(\theta;\lambda)$. We first present some preliminaries that can be useful for developing the proof result in Section 4.2.1. As any two norms on a finite-dimensional vector space are equivalent, we impose the following assumption throughout Section 4 without loss of generality:

ASSUMPTION EC.1. *There exists $\mathfrak{c}$ and $\mathfrak{d}$ such that $\mathfrak{c}\|\cdot\|_2 \leq \|\cdot\| \leq \mathfrak{d}\|\cdot\|_2$.*

By Assumption EC.1, we obtain the bound regarding the dual norm $\|\cdot\|_*$:

$$\mathfrak{d}^{-1}\cdot\|\cdot\|_2 \leq \|\cdot\|_* \leq \mathfrak{c}^{-1}\cdot\|\cdot\|_2.$$

The complexity result of our proposed gradient estimators is summarized below.

REMARK EC.1 (COMPLEXITY OF GRADIENT ESTIMATORS). To generate the SG estimator $v^{\mathrm{SG}}(\theta)$, one needs to generate one sample from $\widehat{\mathbb{P}}$ and 2^L samples from $\mathbb{Q}_{x,\epsilon}$ for some $x \in \operatorname{supp}\widehat{\mathbb{P}}$. To generate the RT-MLMC estimator $v^{\text{RT-MLMC}}(\theta)$, one needs to generate one sample from $\widehat{\mathbb{P}}$ and the required (expected) number of samples from $\mathbb{Q}_{x,\epsilon}$ for some $x \in \operatorname{supp}\widehat{\mathbb{P}}$ equals

$$\frac{L}{2-2^{-L}} = \mathcal{O}(L). \qquad \clubsuit$$

Next, we present some basic properties regarding the approximation function $F^\ell(\theta)$ defined in (14) in Lemma EC.6, which can be used to show Theorem 2. Recall that we defined the constant $K_{\lambda,\epsilon,B} = B/(\lambda\epsilon)$.

LEMMA EC.6. (I) *Under Assumption 2(III), it holds that*

$$\left|F^{\ell}(\theta)-F(\theta)\right| \le \lambda\epsilon \exp\left(2K_{\lambda,\epsilon,B}\right)\cdot 2^{-(\ell+1)}, \qquad \forall\theta\in\Theta.$$

(II) *Under Assumption 2(III) and 2(II), it holds that*

$$\left\|\nabla F^{\ell}(\theta)-\nabla F(\theta)\right\|_2^2 \le L_f^2 \exp\left(4K_{\lambda,\epsilon,B}\right)\cdot 2^{-\ell}, \qquad \forall\theta\in\Theta.$$

(III) *Under Assumption 2(II), it holds that*

$$\mathbb{E}\left[\left\|g^{\ell}(\theta,\zeta^{\ell})\right\|_2^2\right] \le L_f^2, \qquad \forall\theta\in\Theta.$$

Additionally when Assumption 2(III) holds, it holds that

$$\mathbb{E}\left[\left\|G^{\ell}(\theta,\zeta^{\ell})\right\|_2^2\right] \le L_f^2 \exp\left(4K_{\lambda,\epsilon,B}\right)\cdot 2^{-\ell}, \qquad \forall\theta\in\Theta.$$

Proof of Lemma EC.6. Recall that (12) is a special CSO problem in (13), by taking $H^1(\cdot)=\lambda\epsilon\log(\cdot)$ and $H^2(\cdot,z)=\exp(f_\cdot(z)/(\lambda\epsilon))$. Under the assumptions stated in Lemma EC.6, it can be shown that $H^2(\cdot,z)$ is $\exp(K_{\lambda,\epsilon,B})$-uniformly bounded, $\exp(K_{\lambda,\epsilon,B})L_f/(\lambda\epsilon)$-Lipschitz continuous. The function $H^1(\cdot)$ has the domain set $[1,\exp(K_{\lambda,\epsilon,B})]$, and is therefore $\lambda\epsilon$-Lipschitz continuous and $\lambda\epsilon$-smooth. Thus, the desired results hold by applying (Hu et al., 2020a, Lemma 3.1) and (Hu et al., 2021, Proposition 4.1).

EC.6.1. Proof of Theorem 2

We first study the convergence guarantees for solving a generic nonsmooth convex optimization $\min_{\theta\in\Theta} F(\theta)$. Let $\overline{F}(\theta)$ denote its approximation, with the approximation bias Δ_F satisfying

$$|\overline{F}(\theta)-F(\theta)|\le\Delta_F, \quad \forall\theta\in\Theta.$$

Denote by $\nabla\overline{F}(\theta)$ a subgradient of $\overline{F}$ at θ. Suppose for a given θ, the subgradient estimate of $F(\theta)$, denoted as $v(\theta)$, satisfies

$$\mathbb{E}[v(\theta)]=\nabla\overline{F}(\theta), \quad \mathbb{E}\left[\|v(\theta)\|_*^2\right]\le M_*^2.$$

Let $\overline{\theta}^* \in \arg\min_{\theta\in\Theta} \overline{F}(\theta)$ and $\theta^*\in\arg\min_{\theta\in\Theta} F(\theta)$. We then establish the following result.

LEMMA EC.7 **(BSMD for Nonsmooth Convex Optimization).** *Under the assumptions stated above and with the initial guess $\theta_0\in\Theta$, consider the BSMD algorithm that generates the following iteration:*

$$\theta_{t+1}=\mathrm{Prox}_{\theta_t}\big(hv(\theta_t)\big), \quad \theta_0\in\Theta, \quad t=0,\ldots,T-1,$$

where the stepsize parameter $h=\sqrt{\frac{2\kappa D_\omega(\theta_0,\overline{\theta}^*)}{TM_*^2}}$. Let the estimated optimal solution generated by BSMD algorithm be $\widehat{\theta}=\frac{1}{T}\sum_{t=1}^T\theta_t$. Then, the suboptimality gap satisfies:

$$\mathbb{E}\big[F(\widehat{\theta})-F(\theta^*)\big]\le 2\Delta_F+M_*\sqrt{\frac{2D_\omega(\theta_0,\overline{\theta}^*)}{\kappa T}}.$$

REMARK EC.2. If the approximation bias is zero (i.e., $\Delta_F = 0$), the BSMD algorithm reduces to the standard SMD studied in (Nemirovski et al., 2009). By (Nemirovski et al., 2009, Section 2.3), the suboptimality gap in Lemma EC.7 is bounded by $M_* \sqrt{\frac{2D_\omega(\theta_0,\overline{\theta}^*)}{\kappa T}}$. For the case where $\Delta_F > 0$, the proof of Lemma EC.7 follows from the decomposition argument similar to (Hu et al., 2021, Eq. (9)). However, our result generalizes to the BSMD algorithm with (potentially) nonsmooth loss functions, whereas (Hu et al., 2021) focuses only on the SGD algorithm for unconstrained optimization with smooth loss functions. ♣

Now we are ready to show complexity results for BSMD using SG and RT-MLMC estimators. Both estimators rely on the same approximation function $F^L(\theta)$ defined in (14). By Lemma EC.6(I), $\Delta_F = \lambda\epsilon \exp(2K_{\lambda,\epsilon,B}) \cdot 2^{-(L+1)}$. We now analyze each estimator separately.

SG. It can be shown from the first part of Lemma EC.6(III) that $\mathbb{E}[\|v^{\mathrm{SG}}(\theta)\|_*^2] \le \left(M_*^{\mathrm{SG}}\right)^2 := \mathfrak{c}^{-2}L_f^2$. To obtain δ-optimal solution for SG estimator, by Lemma EC.7, it suffices to ensure

$$2\Delta_F \le \frac{\delta}{2}, \quad M_*^{\mathrm{SG}}\sqrt{\frac{2D_\omega(\theta_0,\overline{\theta}^*)}{\kappa T}} \le \frac{\delta}{2}.$$

To satisfy these conditions, we specify the following hyper-parameters:

$$L = \left\lceil \frac{1}{\log 2}\left[\log\frac{2\lambda\epsilon\exp(2K_{\lambda,\epsilon,B})}{\delta}\right]\right\rceil, \quad T = \left\lceil \frac{8L_f^2 D_\omega(\theta_0,\overline{\theta}^*)}{\kappa\mathfrak{c}^2\delta^2}\right\rceil, \quad h = \sqrt{\frac{2\kappa\mathfrak{c}^2 D_\omega(\theta_0,\overline{\theta}^*)}{TL_f^2}}.$$

RT-MLMC. By the second part of Lemma EC.6(III) and basic calculation, we find

$$\begin{aligned}\mathbb{E}\left[\left\|v^{\mathrm{RT-MLMC}}(\theta)\right\|_*^2\right] &\le \mathfrak{c}^{-2}\mathbb{E}\left[\left\|v^{\mathrm{RT-MLMC}}(\theta)\right\|_2^2\right] = \sum_{\ell=0}^{L}\frac{1}{p_\ell}\mathbb{E}\left[\left\|G^\ell(\theta,\zeta_1^\ell)\right\|_2^2\right]\\ &\le \left(M_*^{\mathrm{RT-MLMC}}\right)^2 := 2(L+1)L_f^2\exp(4K_{\lambda,\epsilon,B}).\end{aligned}$$

Similar to the case of SG, we ensure

$$2\Delta_F \le \frac{\delta}{2}, \quad M_*^{\mathrm{RT-MLMC}}\sqrt{\frac{2D_\omega(\theta_0,\overline{\theta}^*)}{\kappa T}} \le \frac{\delta}{2}.$$

To satisfy these conditions, we select the following hyper-parameters:

$$L = \left\lceil \frac{1}{\log 2}\left[\log\frac{2\lambda\epsilon\exp(2K_{\lambda,\epsilon,B})}{\delta}\right]\right\rceil, \quad T = \left\lceil \frac{16(L+1)L_f^2 D_\omega(\theta_0,\overline{\theta}^*)\exp(4K_{\lambda,\epsilon,B})}{\kappa\mathfrak{c}^2\delta^2}\right\rceil,$$

$$h = \sqrt{\frac{2\kappa D_\omega(\theta_0,\overline{\theta}^*)}{T\left(M_*^{\mathrm{RT-MLMC}}\right)^2}}.$$

By Remark EC.1, when running BSMD with SG estimator, the sample complexity from $\widehat{\mathbb{P}}$ equals $\mathcal{O}(T)$ and that from $\mathbb{Q}_{x,\epsilon}$ equals $\mathcal{O}(T2^L)$; when running BSMD with RT-MLMC estimator, the sample complexity from $\widehat{\mathbb{P}}$ equals $\mathcal{O}(T)$ and that from $\mathbb{Q}_{x,\epsilon}$ equals $\mathcal{O}(TL)$. Substituting the expressions of T, L gives the desired result.

EC.7. Proofs of Technical Results in Section 4.2.2

We first provide two technical lemmas that can be useful to show the main results in Section 4.2.2.

LEMMA EC.8. *Under Assumption 2(III), it holds that* $\mathbb{E}[(A^\ell(\theta,\zeta^\ell;\lambda))^2] \le \lambda^2\epsilon^2\exp(2K_{\lambda,\epsilon,B})\cdot 2^{-\ell}$.

Proof of Lemma EC.8. The proof follows the similar procedure from (Hu et al., 2021, Proposition 4.1).

LEMMA EC.9 **(Complexity of RT-MLMC-based Objective Estimator).** *Let error probability* $\alpha\in(0,1)$ *and accuracy level* $\delta>0$. *Assume Assumption 2(III) holds and specify*

$$L=\left\lceil \frac{1}{\log 2}\left[\log\frac{\lambda\epsilon\exp(2K_{\lambda,\epsilon,B})}{\delta}\right]\right\rceil,\qquad m'=\mathcal{O}(1)\frac{\lambda^2\epsilon^2\exp(2K_{\lambda,\epsilon,B})(L+1)}{\delta^2}\cdot\log\frac{2}{\alpha}. \tag{EC.4}$$

Then, the RT-MLMC estimator (18) has an accuracy error δ *with probability at least* $1-\alpha$. *Its sample complexity from* $\widehat{\mathbb{P}}$ *equals* $\mathcal{O}(m')=\widetilde{\mathcal{O}}(\lambda^2\epsilon^2K_{\lambda,\epsilon,B}\exp(2K_{\lambda,\epsilon,B})\cdot\delta^{-2})$ *and that from* $\mathbb{Q}_{x,\epsilon}$ *equals* $\mathcal{O}(m'\cdot L)=\widetilde{\mathcal{O}}(\lambda^2\epsilon^2K_{\lambda,\epsilon,B}^2\exp(2K_{\lambda,\epsilon,B})\cdot\delta^{-2})$. *Here* $\widetilde{\mathcal{O}}(\cdot)$ *hides constants linearly depending on* $(\log\frac{\lambda\epsilon}{\delta})^2$ *and* $\log\frac{1}{\alpha}$.

Proof of Lemma EC.9. We first specify L as in (EC.4) such that $|F^L(\theta;\lambda)-F(\theta;\lambda)|\le\frac{\delta}{2}$. The RT-MLMC estimator (18) satisfies that

$$\mathbb{E}[\widehat{F}(\theta;\lambda)]=F^L(\theta;\lambda),$$
$$\mathbb{V}\mathrm{ar}\left(\widehat{F}(\theta;\lambda)\right)\le\frac{1}{m'}\sum_{\ell=0}^{L}\frac{1}{p_\ell}\mathbb{E}[(A^\ell(\theta,\zeta^\ell))^2]\le\frac{2}{m'}\lambda^2\epsilon^2\exp(2K_{\lambda,\epsilon,B})\cdot(L+1).$$

Consequently, there exists $\delta'>0$ such that

$$\begin{aligned}&\Pr\left\{|F(\theta;\lambda)-\widehat{F}(\theta;\lambda)|>\delta\right\}\le\Pr\left\{|F^L(\theta;\lambda)-\widehat{F}(\theta;\lambda)|>\frac{\delta}{2}\right\}\\ \le&2\exp\left(-\frac{\delta^2}{4(\delta'+2)\mathbb{V}\mathrm{ar}\left(\widehat{F}(\theta;\lambda)\right)}\right)\le 2\exp\left(-\frac{\delta^2m'}{8(\delta'+2)\lambda^2\epsilon^2\exp(2K_{\lambda,\epsilon,B})(L+1)}\right),\end{aligned}$$

where the second inequality is based on the Cramer's large deviation theorem (Kleywegt et al., 2002), and the last inequality is by the upper bound on $\mathbb{V}\mathrm{ar}\left(\widehat{F}(\theta;\lambda)\right)$. To make the desired coverage probability, we take m' as in (EC.4). The complexity results are derived by standard calculation similar to Remark EC.1. □

In the following, we provide the proof of Proposition 1.

Proof of Proposition 1. Denote by $\theta^*=\arg\min_{\theta\in\Theta}\ F(\theta;\lambda)$. The goal is to choose hyper-parameters such that

$$\Pr\left\{\left|\min_{i\in[m]}\ \widehat{F}(\widehat{\theta}_i;\lambda)-F(\theta^*;\lambda)\right|\le\delta\right\}\ge 1-\eta.$$

On the one hand,

$$\min_{i\in[m]}\ \widehat{F}(\widehat{\theta}_i;\lambda)-F(\theta^*;\lambda)\le\min_{i\in[m]}\ F(\widehat{\theta}_i;\lambda)-F(\theta^*;\lambda)+\max_{i\in[m]}\ |F(\widehat{\theta}_i;\lambda)-\widehat{F}(\widehat{\theta}_i;\lambda)|.$$

On the other hand,

$$\begin{aligned}&F(\theta^*;\lambda)-\min_{i\in[m]}\ \widehat{F}(\widehat{\theta}_i;\lambda)\\ \le&F(\theta^*;\lambda)-\min_{i\in[m]}\ F(\widehat{\theta}_i;\lambda)+\max_{i\in[m]}\ |F(\widehat{\theta}_i;\lambda)-\widehat{F}(\widehat{\theta}_i;\lambda)|\\ \le&\max_{i\in[m]}\ |F(\widehat{\theta}_i;\lambda)-\widehat{F}(\widehat{\theta}_i;\lambda)|.\end{aligned}$$

Based on those two inequalities, it suffices to choose hyper-parameters such that

$$\Pr\left\{\max_{i\in[m]}\ |F(\widehat{\theta}_i;\lambda)-\widehat{F}(\widehat{\theta}_i;\lambda)|\le\frac{\delta}{2}\right\}\ge 1-\frac{\eta}{2} \tag{EC.5}$$

and

$$\Pr\left\{\min_{i\in[m]}\ F(\widehat{\theta}_i;\lambda)-F(\theta^*;\lambda)\le\frac{\delta}{2}\right\}\ge 1-\frac{\eta}{2}. \tag{EC.6}$$

To ensure the relation (EC.5), it suffices to apply Lemma EC.9 with error probability $\frac{\eta}{2m}$ and accuracy level $\delta/2$. It implies that the sample complexity from $\widehat{\mathbb{P}}$ at Step 3 of Algorithm 2 for each independent repetition is $\widetilde{\mathcal{O}}(\lambda^2\epsilon^2K_{\lambda,\epsilon,B}\exp(2K_{\lambda,\epsilon,B})\cdot\delta^{-2})$, and that from $\mathbb{Q}_{x,\epsilon}$ is $\widetilde{\mathcal{O}}(\lambda^2\epsilon^2K_{\lambda,\epsilon,B}^2\exp(2K_{\lambda,\epsilon,B})\cdot\delta^{-2})$. To ensure the relation (EC.6), it suffices to take

$$\Pr\left\{F(\widehat{\theta}_i;\lambda)-F(\theta^*;\lambda)\le\frac{\delta}{2}\right\}\ge 1-\left(\frac{\eta}{2}\right)^{1/m},\quad \forall i\in[m].$$

By Markov's inequality, it suffices to ensure

$$\mathbb{E}[F(\widehat{\theta}_i;\lambda)-F(\theta^*;\lambda)]\le\frac{\delta}{2}\left(\frac{\eta}{2}\right)^{1/m},\quad \forall i\in[m]. \tag{EC.7}$$

By Theorem 2(II) with accuracy level $\frac{\delta}{2}\left(\frac{\eta}{2}\right)^{1/m}$, the sample complexity from $\widehat{\mathbb{P}}$ at Step 2 of Algorithm 2 for each independent repetition is $\widetilde{\mathcal{O}}(K_{\lambda,\epsilon,B}\exp(4K_{\lambda,\epsilon,B})\cdot\delta^{-2}\eta^{-2/m})$, and that from $\mathbb{Q}_{x,\epsilon}$ is $\widetilde{\mathcal{O}}(K_{\lambda,\epsilon,B}^2\exp(4K_{\lambda,\epsilon,B})\cdot\delta^{-2}\eta^{-2/m})$. Therefore, the sample complexity from $\widehat{\mathbb{P}}$ of Algorithm 2 is

$$\begin{aligned}&m\cdot\left[\widetilde{\mathcal{O}}(K_{\lambda,\epsilon,B}\exp(4K_{\lambda,\epsilon,B})\cdot\delta^{-2}\eta^{-2/m})+\widetilde{\mathcal{O}}(\lambda^2\epsilon^2K_{\lambda,\epsilon,B}\exp(2K_{\lambda,\epsilon,B})\cdot\delta^{-2})\right]\\ =&\widetilde{\mathcal{O}}(H_{\lambda,\epsilon,B}K_{\lambda,\epsilon,B}\exp(2K_{\lambda,\epsilon,B})\delta^{-2}\cdot m(1+\eta^{-2/m}))\end{aligned}$$

and that from $\mathbb{Q}_{x,\epsilon}$ is

$$\begin{aligned}&m\cdot\left[\widetilde{\mathcal{O}}(K_{\lambda,\epsilon,B}^2\exp(4K_{\lambda,\epsilon,B})\cdot\delta^{-2}\eta^{-2/m})+\widetilde{\mathcal{O}}(\lambda^2\epsilon^2K_{\lambda,\epsilon,B}^2\exp(2K_{\lambda,\epsilon,B})\cdot\delta^{-2})\right]\\ =&\widetilde{\mathcal{O}}(H_{\lambda,\epsilon,B}K_{\lambda,\epsilon,B}^2\exp(2K_{\lambda,\epsilon,B})\delta^{-2}\cdot m(1+\eta^{-2/m})).\end{aligned}$$

In the above deviation, we defined the constant $H_{\lambda,\epsilon,B}=\max(\exp(2K_{\lambda,\epsilon,B}),\lambda^2\epsilon^2)$. Hence, it suffices to specify m such that $\widetilde{\mathcal{O}}(m(1+\eta^{-2/m}))$ is minimized. One valid choice is $m=\lceil\log_2\frac{2}{\eta}\rceil$, which leads to the desired complexity bounds. □

Finally, we show the proof of Theorem 3. A key technique is the following complexity result on bisection search with inexact oracles.

LEMMA EC.10 **(Complexity for Noisy Bisection).** *Let the accuracy level $\delta > 0$, and $\Psi: \mathbb{R} \to \mathbb{R}$ be a L_Ψ-Lipschitz continuous and convex function defined on the interval $[\lambda_l, \lambda_u]$. Assume there exists an oracle $\widehat{\Psi}: \mathbb{R} \to \mathbb{R}$ such that $|\widehat{\Psi}(\lambda) - \Psi(\lambda)| \le \delta, \forall \lambda$. Let us run Algorithm 3 for $T' = \lceil \log_2\left(\frac{L_\Psi(\lambda_u - \lambda_l)}{\delta}\right)\rceil$ iterations, then with at most $3 + 2T'$ calls to $\widehat{\Psi}$, Algorithm 3 outputs $\widehat{\lambda}$ so that*

$$\Psi(\widehat{\lambda}) - \min_{\lambda \in [\lambda_l, \lambda_u]} \Psi(\lambda) \le 4\delta.$$

Proof of Lemma EC.10. The proof is straightforward by following (Cohen et al., 2016, Lemma 33) □

Proof of Theorem 3. It can be verified that Ψ is a convex function with a subgradient

$$\frac{\partial}{\partial \lambda}\Psi(\lambda) = \overline{\rho} + \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\epsilon \log \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_\lambda^*}(z)/(\lambda\epsilon)}\right]\right] - \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\frac{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_\lambda^*}(z)/(\lambda\epsilon)} f_{\theta_\lambda^*}(z)\right]}{\lambda \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_\lambda^*}(z)/(\lambda\epsilon)}\right]}\right],$$

where $\theta_\lambda^* \in \arg\min_{\theta\in\Theta} F(\theta;\lambda)$. By Assumption 2 and $\lambda \in [\lambda_l, \lambda_u]$, this subgradient vector is bounded:

$$\left|\frac{\partial}{\partial \lambda}\Psi(\lambda)\right| \le L_\Psi := \overline{\rho} + \frac{B}{\lambda_l}\left[1 + \exp(K_{\lambda_l,\epsilon,B})\right].$$

In summary, $\Psi(\lambda)$ is a L_λ-Lipschitz and convex function defined on $[\lambda_l, \lambda_u]$. Applying Lemma EC.10 with accuracy level $\delta/4$ together with the union bound, we are able to find the optimal multiplier up to accuracy δ with probability at least $1 - \eta$ by calling the oracle $\widehat{\Psi}$ for $3 + 2\left\lceil \log_2\left(\frac{4L_\Psi(\lambda_u - \lambda_l)}{\delta}\right)\right\rceil$ times. □

LEMMA EC.11. *Under Assumption 2, the optimal multiplier λ^* to (D) satisfies $\lambda^* \le \frac{B}{\rho}$.*

Proof. It can be verified that

$$\begin{aligned}
0 &= \frac{\partial}{\partial \lambda}\Psi(\lambda)\Big|_{\lambda=\lambda^*} \\
&= \overline{\rho} + \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\epsilon \log \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_{\lambda^*}^*}(z)/(\lambda\epsilon)}\right]\right] - \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\frac{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_{\lambda^*}^*}(z)/(\lambda\epsilon)} f_{\theta_{\lambda^*}^*}(z)\right]}{\lambda^* \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_{\lambda^*}^*}(z)/(\lambda\epsilon)}\right]}\right] \\
&\ge \overline{\rho} - \mathbb{E}_{x\sim\widehat{\mathbb{P}}}\left[\frac{\mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_{\lambda^*}^*}(z)/(\lambda\epsilon)} f_{\theta_{\lambda^*}^*}(z)\right]}{\lambda^* \mathbb{E}_{z\sim\mathbb{Q}_{x,\epsilon}}\left[e^{f_{\theta_{\lambda^*}^*}(z)/(\lambda\epsilon)}\right]}\right] \\
&\ge \overline{\rho} - \frac{B}{\lambda^*},
\end{aligned}$$

where the two inequalities is based on the fact that $0 \le f_\theta(z) \le B$. The desired result holds directly. □