Noname manuscript No.
(will be inserted by the editor)

Statistical Inference of Contextual Stochastic Optimization with Endogenous Uncertainty*

Junyu Cao · Rui Gao† · Zhen Yang

Abstract This paper considers contextual stochastic optimization with endogenous uncertainty, where random outcomes depend on both contextual information and decisions. We analyze the statistical properties of solutions from two prominent approaches: predict-then-optimize (PTO), which first predicts a model between outcomes, contexts, and decisions, and then optimizes the downstream objective; and estimate-then-optimize (ETO), which directly estimates the conditional expectation of the objective and optimizes it. Unlike many existing studies that assume independent and identically distributed observations and/or decision/context-independent noise, we consider a setting where historical observations form a general time series, allowing for arbitrary dependencies between current outcomes and past realizations, contexts, and decisions. For both approaches, we establish non-asymptotic performance guarantees using two criteria, approximation error and regret, deriving slow and fast convergence rates.

Keywords Contextual stochastic optimization · Decision-dependent uncertainty · Endogeneity · Predict-then-optimize

Mathematics Subject Classification (2020) 90C15

1 Introduction

Contextual stochastic optimization addresses the problem

$$\min_{x \in \mathcal{X}} \mathbb{E}_{\mathbb{P}_{\xi|\zeta}} \left[\Psi(x, \xi) \right]$$

where $x \in \mathcal{X}$ represents the decision variable, ξ denotes the random outcome realized after the decision is made, ζ symbolizes the contextual information available before decision-making, and Ψ is the objective

* A version of the manuscript was previously circulated as “Contextual Decision-making under Parametric Uncertainty and Data-driven Optimistic Optimization”.

Junyu Cao
E-mail: junyu.cao@mccombs.utexas.edu

Rui Gao† (corresponding author)
E-mail: rui.gao@mccombs.utexas.edu

Zhen Yang
E-mail: zhen.yang@mccombs.utexas.edu
Department of Information, Risk and Operatiosn Management, The University of Texas at Austin, Austin, Texas 78712

function dependent on both x and ξ. The distribution of ξ is influenced by the contextual information ζ. By leveraging the contextual information ζ, decision-makers can achieve more accurate and customized optimization outcomes.

In classic stochastic programming literature (e.g., [44]), uncertainties ζ and ξ are often assumed to be exogenous, namely, their distribution is independent of the decision x. However, endogenous uncertainty, where the uncertainty is influenced by the decision, frequently occurs in many applications. For instance, consider a revenue management problem, where the decision maker determines the pricing vector x for all products to maximize the expected revenue $\Psi(x, \xi) = x^\top \xi$. Here, the demand ξ depends not only on the product and customer features but also critically on the pricing decision x. Motivated by such problems, this paper considers endogenous uncertainty, and to reflect the dependence of the distribution of ξ on the decision x, we explicitly write it as $\mathbb{P}_{\xi|(\zeta,x)}$ and our problem of interest is

$$\min_{x \in \mathcal{X}} \ \mathbb{E}_{\mathbb{P}_{\xi|(\zeta,x)}} \left[\Psi(x, \xi) \right]. \tag{1}$$

In the simplest case where there is no context (other than the decision), the distribution $\mathbb{P}_{\xi|x}$ is decision-dependent, and (1) reduces to the stochastic optimization with decision-dependent uncertainty [20, 35, 34, 36].

Two popular solution approaches for solving (1) are *predict-then-optimize* (PTO) and *estimate-then-optimize* (ETO). PTO first estimates a prediction model between ξ and (ζ, x), then uses this model to solve (1). When Ψ is linear in ξ (e.g., $\Psi(x, \xi) = x^\top \xi$), thanks to the linearity of expectation, it suffices to estimate the conditional expectation of ξ given (ζ, x). In this case, it has been shown that PTO can achieve a fast convergence rate when there is no endogenous uncertainty. However, when Ψ is nonlinear in ξ or the noise exhibits heteroscedasticity with respect to the decisions, the model uncertainty in the noise distribution may be amplified in the downstream optimization task, raising concerns about the accuracy and robustness of the solutions. To address these concerns, the second approach, PTO, directly estimates the functional relationship between the conditional mean objective $\mathbb{E}_{\mathbb{P}_{\xi|(\zeta,x)}}[\Psi(x, \xi)]$ and (ζ, x), thereby reducing (1) to finding the optimal x using the estimated function [4]. These two approaches are widely adopted in practice, offer clear interpretability, and have been proven statistically effective in scenarios without model misspecification [22, 15]. Given the challenges in analyzing the statistical properties of these two approaches under endogenous uncertainty, we focus our study on PTO and ETO in this paper, while providing a brief discussion of other frameworks in Section 1.1.

In this paper, we are interested in assessing the statistical quality of the solution to (1). Endogeneity significantly impacts this assessment, and most existing analyses impose assumptions that rarely hold in its presence. For example, it is often assumed that one has i.i.d. observations. While this assumption is reasonable when this distribution is exogenous, it rarely holds under endogenous uncertainty. For example, in the pricing decision scenario, it is uncommon for the historical pricing decisions to be i.i.d. Another common assumption in the literature is that the noise in the regression model is independent of the context or decision, implying that the variability in outcomes remains constant regardless of the context and the decision. However, this assumption may not hold in practical scenarios. For example, the demand variation for a product might be larger during periods of promotional activities compared to periods without promotion.

We consider a general data-generating mechanism where the observations of $\{(\zeta_i, X_i, \xi_i)\}_{i=1}^n$ form a time series. The random outcome ξ_i can depend on its historical realizations $\{\xi_j\}_{j=1}^{i-1}$, as well as historical and current contexts and decisions $\{(\zeta_j, X_j)\}_{j=1}^{i}$, in an arbitrary manner. Notably, the noise in the regression model can be context- and decision-dependent (see Section 3 for details).

In Section 4, we quantify the parameter estimation error in the regression model. To deal with the dependent sequence, we employ a martingale argument based on self-normalized processes. This technique, previously used to establish the regret bounds for (generalized) linear bandits, is extended to more general parametric settings by coupling with a covering number argument.

Based on the statistical analysis of the estimator, in Section 5, we establish performance guarantees for both the PTO and ETO approaches. We use two criteria widely adopted in stochastic programming and learning theory: the approximation error, which quantifies the difference between the optimal value using the estimated model and that using the ground truth; the regret, which measures the objective difference between the estimated solution and the true optimal solution under the ground truth model. For both approaches, we establish a slow rate, proportional to the parameter estimation accuracy, and a fast rate, quadratic to the parameter estimation accuracy, under additional conditions. These results are illustrated in the context of a pricing problem.

1.1 Related Literature

Contextual stochastic optimization frameworks can be broadly categorized into three types: sequential learning and optimization (SLO), integrated learning and optimization (ILO), and decision rule optimization. Both PTO and ETO fall under the SLO category; their definitions are aligned with the tutorial [40]. Below we will focus on the statistical aspects of these problems, referring readers to [43] for an excellent recent survey on modeling and computational aspects.

Within the SLO framework, [4, 2] derived asymptotics for ETO using various non-parametric estimators. [2] provided finite-sample bounds when the objective function belongs to a reproducing kernel Hilbert space. [23, 24, 25] studied asymptotics and finite-sample guarantees for residual-based sample average approaches and distributionally robust optimization.

In the ILO framework, [12, 21, 32, 13] developed generalization bounds and risk bounds for the “Smart-Predict-then-Optimize” approach. [54] related their proposed robust estimates uncertainty set with confidence regions through maximum likelihood. [41] derived asymptotics and finite-sample guarantees for Integrated Conditional Estimation-Optimization. [45, 3, 6, 48, 9, 46] provided performance guarantees for approaches based on regularization and robust optimization.

For decision rule optimization, [5] provided asymptotics and finite sample guarantees for decision rules in a reproducing kernel Hilbert space. [18] demonstrated that the bias-corrected policy enjoys provably strong performance in the small-data, large-scale regime. [52] showed the consistency for the piecewise affine decision rule. [8] established performance guarantees for training a neural-network decision rule. In the context of feature-based newsvendor problems, [2] developed risk bounds for affine decision rules, [16] established consistency for decision rules based on operational statistics, and [51, 19] derived generalization bounds for Lipschitz regularized policies driven by Wasserstein distributionally robust optimization.

Comparing frameworks, [22] demonstrated that PTO could achieve faster regret convergence than Integrated Empirical Risk Minimization for linear objective functions. For nonlinear decision objectives, [15] proved that ETO could outperform ILO asymptotically in terms of stochastic dominance of regret when there is no model misspecification. Our fast convergence rate results on PTO and ETO are aligned with their insights.

All above analyses assume i.i.d. observations, except for a few exceptions. [4] analyzed scenarios when the data sequence is a certain mixing process and [46] considered finite-state Markov chains and auto-regressive process. [16] assumed that the new context is independent of previous outcomes conditional on the previous contexts in a newsvendor model. In contrast to these works, we consider a general time series setting allowing arbitrary dependence on the history. For approaches that involve an explicit noise component, most of them assume homogeneous noise, except for [24] that considered a decision-independent multiplicative noise heteroscedasticity. In contrast, our model allows arbitrary dependence of noise on past decisions and contexts.

Finally, our analysis is also related to the bandits literature (see, e.g., [28]), where the action depends on historical contexts so that the observations are adapted. A fair amount of works have been developed for linear bandits [11, 42, 30, 10], generalized linear models [17, 31, 27], kernelized contextual bandits [47], and neural contextual bandits [49, 53]. However, while the analysis in these works relies essentially on some (generalized) linear structure, we do not have hidden linear structure in our model and instead focus on the statistical inference for a general parametric class.

2 Model Setup

2.1 Predict-Then-Optimize

Suppose the random outcome $\xi \in \Xi$, the context $\zeta \in \Omega$, and the decision $x \in \mathcal{X}$ have a relationship

$$\xi = \phi\big(f_{\theta^*}(\zeta, x)\big) + \epsilon, \tag{2}$$

where $\phi : \mathbb{R}^{d_\Xi} \to \mathbb{R}^{d_\Xi}$ is a link function, which is useful when ξ is linked with a transformation of f_{θ^*} (see Example 1 below); $f_{\theta^*} : \mathbb{R}^{d_\Omega + d_\mathcal{X}} \to \mathbb{R}^{d_\Xi}$ models the parametric relationship between the random outcome and the context; $\theta^* \in \mathbb{R}^d$ is an unknown parameter that can only be estimated from historical observations; and $\epsilon \in \mathbb{R}^{d_\Xi}$ denotes the noise vector. Our assumption on the noise (see Assumption 1 below) allows heteroskedasticity and arbitrary dependence on the context and decision.

Given n historical observations, the PTO approach first estimates the unknown model parameter, denoted by $\widehat{\theta}_n$, then optimizes the downstream objective building upon $\widehat{\theta}_n$. We focus on cases where Ψ is linear in ξ. This is crucial because when Ψ is nonlinear in ξ, the objective function depends not only on the conditional expectation of $\phi(f_{\theta^*}(\cdot))$ but also on the noise distribution. Consequently, estimating only the unknown parameter is insufficient for the downstream optimization, as relying solely on expectation information would introduce bias into the objective, especially in the presence of heteroskedasticity. Under the linear assumption, we can express $\Psi(x, \xi) = \psi(x)^\top \xi$. With the estimator $\widehat{\theta}_n$, the PTO approach solves the downstream optimization

$$\min_{x \in \mathcal{X}} \ \psi(x)^\top \phi\big(f_{\widehat{\theta}_n}(\zeta, x)\big). \tag{3}$$

This formulation extends existing literature [14, 33, 22] where ψ and ϕ are the identity maps.

The following example illustrates the application of our model in a single-product pricing problem.

Example 1 Let $x \in \mathbb{R}$ denote the pricing decision for a product, and let $\xi \in \{0, 1\}$ denote whether to purchase the product. Let $f_\theta(\zeta, x) = \theta^\top [1; \zeta; x]$ and $\phi(z) = \frac{1}{1+\exp(-z)}$, then the demand function $\phi(f_{\theta^*}(\zeta, x)) = \frac{1}{1+\exp(-(\theta^*)^\top [1;\zeta;x])}$ models the relationship between the purchasing probability and context ζ, known as the logit model. When the parameter associated with x is the price sensitivity coefficient. The objective function denotes the negative revenue $\Psi(x\xi) = -x\xi$. The unknown parameter can be estimated using estimation for generalized linear models. One can also consider other demand models, such as linear demand. In this case, the demand $\xi \in \mathbb{R}_+$ has an expectation $\phi(f_{\theta^*}(\zeta, x)) = (\theta^*)^\top [1; \zeta; x]$ with ϕ being the identity map. ♣

2.2 Estimate-Then-Optimize

The ETO approach directly works with the conditional expectation $\mathbb{E}_{\mathbb{P}_{\xi|\zeta}}[\Psi(x, \xi)]$. Set $y := \Psi(x, \xi)$. Suppose y can be represented as

$$y = \phi\big(f_{\theta^*}(\zeta, x)\big) + \epsilon, \tag{4}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a link function; $f_{\theta^*} : \mathbb{R}^{d_\Omega + d_\mathcal{X}} \to \mathcal{Y} \subset \mathbb{R}$ and $\phi(f_{\theta^*}(\zeta, x)) = \mathbb{E}_{\mathbb{P}_{\xi|\zeta}}[\Psi(x, \xi)]$; and ϵ denotes the noise. Unlike PTO, the ETO approach does not estimate the conditional mean of ξ but instead estimates the conditional mean $\mathbb{E}_{\mathbb{P}_{\xi|\zeta}}[\Psi(x, \xi)]$ directly. This approach offers three benefits: first, it circumvents the need for a separate estimation of ξ followed by optimization over x, potentially reducing the impact of statistical errors for downstream optimization. Second, it is well-suited for nonlinear objective functions

$\Psi(x, \cdot)$, as it avoids the need to estimate the noise distribution in (2). Third, it retains its applicability across scenarios where the explicit functional form of Ψ is unknown or when ξ is not directly observable.

Setting $y_i = \Psi(x, \xi_i)$ to be the objective value associated with the i-th observation. Given n historical observations, let $\widehat{\theta}_n$ be the estimator of the model parameter. With $\widehat{\theta}_n$, the ETO approach solves the optimization

$$\min_{x \in \mathcal{X}} \ \phi\big(f_{\widehat{\theta}_n}(\zeta, x)\big). \tag{5}$$

We exemplify this approach in the following newsvendor problem with pricing.

Example 2 Consider a retailer who faces the dual challenge of deciding the ordering quantity q and setting a selling price p for a product. Its random demand ξ is affected by the selling price p, the inventory quantity q, and other external factors such as market trends, seasonality, or promotional activities. The objective $\Psi((q, p), \xi)$ is the newsvendor cost:

$$\Psi((q, p), \xi) = -p \min(\xi, q) + cq - s(q - \min(\xi, q)),$$

where c is the unit ordering cost and s is the unit salvage value of unsold inventory. The retailer's objective is to find the optimal inventory quantity and selling price to minimize the expected cost

$$\min_{p,q \geq 0} \mathbb{E}_{\mathbb{P}_{\xi|\zeta}}[\Psi((q, p), \xi)].$$

Let the link function ϕ be the identity map and thereby $\mathbb{E}_{\mathbb{P}_{\xi|\zeta}}[\Psi((q, p), \xi)] = f_{\theta^*}(q, p, \zeta)$. ♣

3 Statistical Estimation

As previously noted, the historical observations are typically correlated due to endogeneity. We model these observations as a time series

$$\zeta_1 \rightsquigarrow X_1 \rightsquigarrow \xi_1 \rightsquigarrow \zeta_2 \rightsquigarrow X_2 \rightsquigarrow \xi_2 \rightsquigarrow \cdots \rightsquigarrow \zeta_n \rightsquigarrow X_n \rightsquigarrow \xi_n,$$

allowing for an arbitrary dependence structure. Note that we use a capital letter X to indicate that the decision can be history-dependent and thus random. For instance, consider the pricing example where a seller dynamically adjusts its pricing strategy. The i-th price X_i may depend on all historical prices and demands, and other historical and current contexts like inventory levels, consumer preferences, competitor pricing, etc. In our model, the i-th context ζ_i can depend on history $\{(\zeta_j, X_j, \xi_j)\}_{j=1}^{i-1}$ in any arbitrary manner; and the i-th demand ξ_i can depend on (ζ_i, X_i) without restriction. On the other hand, the i-th observation (ζ_i, X_i, ξ_i) does not depend on the subsequent observations $\{(\zeta_j, X_j, \xi_j)\}_{j=i+1}^{n}$. To formalize our data-generating mechanism, we introduce the σ-algebra

$$\mathcal{H}_i := \sigma(\zeta_1, X_1, \xi_1, \cdots, \zeta_{i-1}, X_{i-1}, \xi_{i-1}, \zeta_i, X_i),$$

which summarizes all information available just before the i-th random outcome ξ_i is observed.

We impose the following assumption on the noise in our regression models (2) and (4). Let ϵ_i be the noise associated with the i-th observation.

Assumption 1 *There exists $\sigma > 0$ such that for every $i = 1, \ldots, n$, every $\lambda > 0$ and every unit vector u, it holds that $\mathbb{E}[\exp(\lambda u^\top \epsilon_i) \mid \mathcal{H}_i] \leq \exp(\lambda^2 \sigma^2 / 2)$.*

This assumption states that the noise ϵ_i associated with the i-th demand ξ_i is sub-Gaussian, conditional on the historical information. Importantly, this assumption can accommodate scenarios where the noise component in the regression model is correlated not only with the current context but also with all historical decisions and contexts.

To accommodate the setups in both the PTO approach and the ETO approach, we set $y_i = \xi_i$ in the PTO approach. Suppose $\widehat{\theta}_n$ can be solved from the following equation

$$\sum_{i=1}^{n} \nabla f_\theta(\zeta_i, X_i)^\top \big(y_i - \phi(f_\theta(\zeta_i, X_i))\big) = 0, \tag{6}$$

where ∇f_θ denotes the Jacobian matrix of the vector-valued function f with respect to θ. This is a type of M-estimator based on the first derivative. When ϕ is the identity map, this corresponds to the least square estimate; when $f_\theta(\zeta) = \theta^\top \zeta$ and ϕ is the link function in the generalized linear model, this corresponds to the maximum quasi-likelihood estimate.

Example 3 Consider the revenue management problem in Example 1. Equation (6) corresponds to the first-order condition associated with the maximum quasi-likelihood estimation

$$\sum_{i=1}^{n} \left(y_i - \frac{1}{1 + \exp(-\theta^\top [1; \zeta_i; X_i])} \right) [1; \zeta_i; X_i] = 0,$$

and $\widehat{\theta}_n$ is the maximum quasi-likelihood estimator. ♣

Equation (6) includes a broad family of estimators for nonlinear models. In what follows, we construct a confidence region for the estimator under general structures between the random outcome, contexts, and decisions.

4 Confidence Region of the Estimated Parameters

In this section, we study the construction of confidence regions for the estimated parameter $\widehat{\theta}_n$. The result in this section will be the building block for our main results on the performance guarantees for the PTO and ETO approaches. We aim to establish a high-probability bound on $\|\widehat{\theta}_n - \theta^*\|_2$. To this end, we will extend the techniques, in a non-trivial way, from contextual linear bandit theory [29] to our broader parametric framework.

We begin by introducing several boundedness assumptions. Let $\nabla f_\theta(\zeta, x) \in \mathbb{R}^{d_Y \times d}$ denote the Jacobian matrix of f with respect to θ.

Assumption 2 *The following holds:*

(i) $\kappa_f := \sup_{x\in\mathcal{X},\zeta\in\Omega,\theta\in\Theta} \|\nabla f_\theta(\zeta, x)\|_F < \infty$;
(ii) $\hbar_f := \sup_{x\in\mathcal{X},\zeta\in\Omega,\theta\in\Theta} \|\nabla^2 f_\theta(\zeta, x)\|_F < \infty$;
(iii) $\underline{\kappa}_\phi := \inf_{x\in\mathcal{X},\zeta\in\Omega,\theta\in\Theta} \phi'(f_\theta(\zeta, x)) > 0$;
(iv) $\beta_\Theta := \sup_{\theta\in\Theta} \|\theta\|_2 < \infty$;
(v) $\beta_\psi := \sup_{x\in\mathcal{X}} \|\psi(x)\|_2 < \infty$;

where $\|\cdot\|_F$ *denotes the Frobenius norm of a tensor, computed as the norm of the vectorization of the tensor.*

To deal with non-i.i.d. data, our analysis of the concentration bound extends the self-normalized processes [37, 38] developed for generalized linear regression with adapted data [1]. Specifically, let us define

$$Z_n(\theta) := \sum_{i=1}^{n} \nabla f_\theta(\zeta_i, X_i)^\top \epsilon_i, \qquad \widehat{Z}_n := Z_n(\widehat{\theta}_n),$$
$$V_n(\theta) := \sum_{i=1}^{n} \nabla f_\theta(\zeta_i, X_i)^\top \nabla f_\theta(\zeta_i, X_i), \qquad \widehat{V}_n := V_n(\widehat{\theta}_n).$$

Note that when $f_\theta(\cdot)$ is linear, V_n is related to the design matrix in linear regression.

The roadmap of our analysis is as follows. Detailed proofs will be postponed to Section 6.

(I) To establish a high-probability bound on $\|\widehat{\theta}_n - \theta^*\|_2$, we focus on analyzing the bound on $\|\widehat{\theta}_n - \theta^*\|_{\widehat{V}_n}$, where the matrix norm $\|a\|_{\widehat{V}_n} := \sqrt{a^\top \widehat{V}_n a}$. We will construct an exponential super-martingale as

$$\prod_{i=1}^{n} \exp\left(u v^\top \nabla f_\theta(\zeta_i, X_i)^\top \epsilon_i - \frac{u^2}{2} \cdot 2\sigma^2 \cdot \|v^\top \nabla f_\theta(\zeta_i, X_i)\|_2^2 \right),$$

where $v \in \mathbb{R}^d$ and $u \in \mathbb{R}$ are fixed. The concentration of this super-martingale renders a high-probability bound for $v^\top Z_n(\theta)$. Taking $v = V_n(\theta)^{-1/2} e_j$, $j = 1, \ldots, d$ and using union bound yields a high-confidence bound for $\|Z_n(\theta)\|^2_{V_n(\theta)^{-1}}$. This will be established in Lemma 1(I).

(II) To take into account the randomness of $\widehat{\theta}_n$, we adopt a covering number argument, using the fact that $\|Z_n(\theta)\|^2_{V_n(\theta)^{-1}}$ is a Lipschitz function of θ. This will establish a bound for $\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}}$; see Lemma 1(II).

(III) Finally, we show that $\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}}$ serves as an upper bound on a multiple of $\|\widehat{\theta}_n - \theta^*\|_{\widehat{V}_n}$ under proper conditions. The 2-norm distance $\|\widehat{\theta}_n - \theta^*\|_2$ can be further upper bounded based on the eigenvalues of $\widehat{V}_n$. See Theorem 1.

For the subsequent discussion, we assume that

$$\Lambda_n := \inf_{\theta \in \Theta} \lambda_{\min}(V_n(\theta)) > 0.$$

The results for the first two steps are established in Lemma 1.

Lemma 1 *Assume Assumptions 1 and 2(i) are in force. Let* $\delta \in (0, 1/2)$.

(I) *For every* $\theta \in \Theta$*, with probability at least* $1 - \delta$*, it holds that*

$$\|Z_n(\theta)\|^2_{V_n(\theta)^{-1}} \le 16 d \eta^2 \sigma^2 \log(n) \log(d/\delta),$$

where $\eta = \sqrt{3 + 2\log(1 + 2\kappa_f^2/\lambda_{\min}(V_n(\theta)))}$.

(II) *Assume additionally that Assumptions 2(ii)-(iv) are in force. Then with probability at least* $1 - 3\delta$*, it holds that*

$$\|\widehat{Z}_n\|^2_{\widehat{V}_n^{-1}} \le 16 d \eta^2 \sigma^2 \log(n) \mathfrak{C}_{\mathcal{F}},$$

where $\mathfrak{C}_{\mathcal{F}} = \log(d/\delta)$ *when* $\mathcal{F} = \left\{ \phi(\theta^\top [\zeta; x]) : \theta \in \Theta \right\}$ *and*

$$\mathfrak{C}_{\mathcal{F}} = 4d \log\left(n^3 d^3 d_{\mathcal{Y}}^2 \log(1/\delta) \varsigma / (\delta \Lambda_n) \right),$$

when $\mathcal{F}$ *is a general class* $\left\{ \phi(f_\theta(\cdot)) : \theta \in \Theta \right\}$*, where* $\varsigma = 96\sigma^2 \kappa_f (\hbar_f + 1) \beta_\Theta \left(1 + \frac{\hbar_f^2}{\Lambda_n}\right)$.

Note that for the (generalized) linear class, we can immediately get the result of part (II) from part (I), using the property $\|Z_n(\theta_1)\|_{V_n(\theta_1)^{-1}} = \|Z_n(\theta_2)\|_{V_n(\theta_2)^{-1}}$ for all θ_1 and θ_2. For a general class, this property no longer holds. As such, we have to deal with the randomness of the quantity $\|\widehat{Z}_n(\widehat{\theta}_n)\|_{\widehat{V}_n(\widehat{\theta}_n)^{-1}}$, through a more involved super-martingale argument combined with a covering number argument as described above. This introduces an extra $\sqrt{d}$ factor, resulting from the Lipschitz constant of the function $\theta \mapsto \|Z_n(\theta)\|_{V_n(\theta)^{-1}}$.

According to the roadmap, it remains to relate $\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}}$ to $\|\widehat{\theta}_n - \theta^*\|_{\widehat{V}_n}$. Utilizing the mean value theorem, we linearize the difference in function evaluations between parameters $\widehat{\theta}_n$ and θ^* as

$$\phi(f_{\widehat{\theta}_n}(\zeta,x)) - \phi(f_{\theta^*}(\zeta,x)) = g_{\bar{\theta}}(\zeta,x)(\widehat{\theta}_n - \theta^*),$$

where $g_{\bar{\theta}}(\zeta,x) := \nabla_\theta(\phi(f_\theta(\zeta,x)))\,|_{\theta=\bar{\theta}}$ and $\bar{\theta}$ is a convex combination of $\widehat{\theta}_n$ and θ^*. The term $\widehat{Z}_n$ is then given by

$$\widehat{Z}_n = \sum_{i=1}^{n} \nabla f_{\widehat{\theta}_n}(\zeta_i, X_i)^\top g_{\bar{\theta}_i}(\zeta_i, X_i)(\widehat{\theta}_n - \theta^*), \tag{7}$$

where $g_{\bar{\theta}}(\zeta,x) := \nabla_\theta \phi(f_\theta(\zeta,x))|_{\theta=\bar{\theta}}$ and $\bar{\theta}$ is a convex combination of $\widehat{\theta}_n$ and θ^*. Thus

$$\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}} \simeq \|(\widehat{\theta}_n - \theta^*)^\top \sum_{i=1}^{n} g_{\widehat{\theta}_n}(\zeta_i, X_i)\nabla f_{\widehat{\theta}_n}(\zeta_i, X_i)^\top\|_{\widehat{V}_n}.$$

To relate this with $\|\theta^* - \widehat{\theta}_n\|_{\widehat{V}_n}$, we impose the following assumption, which is closely related to the notion of one-point convexity in non-convex optimization [26], which indicates that proximity in function values implies proximity in parameters.

Assumption 3 *There exists $\alpha > 0$ such that for every $x \in \mathcal{X}$, $\zeta \in \Omega$ and $\theta \in \Theta$,*

$$\left(\nabla f_\theta(\zeta,x)^\top \left(\phi(f_\theta(\zeta,x)) - \phi(f_{\theta^*}(\zeta,x))\right)\right)^\top (\theta - \theta^*) \geq \alpha \|\nabla f_\theta(\zeta,x)(\theta - \theta^*)\|_2^2.$$

This assumption generalizes strong convexity to non-convex functions. Consider minimizing a statistical loss function $\ell(y, \phi(f_\theta(\zeta,x)))$. Then it can be verified that (6) holds when ℓ is the square loss, negative log-likelihood, or the cross-entropy loss. Note that a function $h : \Theta \to \mathbb{R}$ is said to satisfy restricted secant inequality [50] at θ^* if $\nabla h(\theta)^\top(\theta - \theta^*) \geq \tilde{\alpha}\|\theta - \theta^*\|$ for all $\theta \in \theta$, where $\tilde{\alpha} > 0$. If for every $\zeta \in \Omega$ and $y \in \mathcal{Y}$, $\ell(y, \phi(f_\theta(\zeta,x)))$ satisfies the restricted secant inequality at θ^*, then Assumption 3 holds. See Appendix 6.3 for details. This can occur even when $\ell(y, \phi(f_\theta(\zeta,x)))$ is nonconvex in θ. From an algorithmic point of view, this condition ensures the global optimality of gradient-based algorithms for non-convex optimization.

We are now ready to present the main result in this section.

Theorem 1 *Let $\delta \in (0, 1/2)$. Assume Assumptions 1, 2, 3 are in force. Then with probability at least $1 - 3\delta$,*

$$\|\theta^* - \widehat{\theta}_n\|_2 \leq \rho_n := \frac{4\sqrt{d}\eta\sigma\sqrt{\log(n)\mathfrak{C}_{\mathcal{F}}}}{\min(\underline{\kappa}_\phi, \alpha)\sqrt{\Lambda_n}},$$

where $\mathfrak{C}_{\mathcal{F}}$ is defined in Lemma 1.

According to Theorem 1, replacing δ with $\delta/3$, we can ensure that with probability at least $1 - \delta$, $\|\theta^* - \widehat{\theta}_n\|_2 \leq \rho_n$. This expression demonstrates that the precision of parameter estimates in our framework improves with increasing sample size, and the confidence interval for the error shrinks at a rate of $\tilde{O}(1/\sqrt{\Lambda_n})$. When the observations are i.i.d., for linear regression, the least square method yields an optimal rate $\tilde{O}(\sqrt{d/n})$, which is consistent with the rate established in Theorem 1 with $\Lambda_n = O(n)$ and $\mathfrak{C}_{\mathcal{F}} = \log(d/\delta)$. When the data is not i.i.d., the convergence rate depends on Λ_n, which reflects the informativeness of the data. The following example illustrates that the decision can be arbitrarily bad when $\Lambda_n = 0$.

Example 4 Consider the problem in Example 1. If there is no variation in the pricing decision observations x_i, then $\Lambda_n = 0$ and it would be impossible to learn the price sensitivity. In this case, the resulting pricing decision could be arbitrarily bad even with an infinite amount of data. ♣

Let us revisit Example 3 to illustrate the specific bound.

Example 5 In the setting of Example 3, with simple calculations, we obtain that $\kappa_f = \sqrt{1+\bar{\zeta}^2+\bar{x}^2}$ and $\hbar_f = 0$, where $\bar{\zeta} = \sup_{\zeta\in\Omega}\|\zeta\|_2$ and $\bar{x} = \sup_{x\in\mathcal{X}}\|x\|_2$. Note that $\phi'(z) = \frac{e^{-z}}{(1+e^{-z})^2}$ and $\theta^\top[1;\zeta;x] \le \|\theta\|_2\|[1;\zeta;x]\|_2 \le \sqrt{1+\bar{\zeta}^2+\bar{x}^2}\beta_\Theta$. Thus, we have $\underline{\kappa}_\phi = \frac{e^{-\sqrt{1+\bar{\zeta}^2+\bar{x}^2}\beta_\Theta}}{(1+e^{-\sqrt{1+\bar{\zeta}^2+\bar{x}^2}\beta_\Theta})^2}$. Assumption 3 holds with $\alpha = \underline{\kappa}_\phi$. With the above values of parameters, we have $\varsigma = 96\sigma^2\sqrt{1+\bar{\zeta}^2+\bar{x}^2}\beta_\Theta$ and then set

$$\rho_n = 4\sqrt{d}\eta\sigma\sqrt{\log(n)\mathfrak{C}_\mathcal{F}}e^{\sqrt{1+\bar{\zeta}^2+\bar{x}^2}\beta_\Theta}(1+e^{-\sqrt{1+\bar{\zeta}^2+\bar{x}^2}\beta_\Theta})^2/\sqrt{\Lambda_n},$$

where $\mathfrak{C}_\mathcal{F} = \log(d/\delta)$. Then it holds with probability at least $1-\delta$ that $\|\theta^* - \widehat{\theta}_n\|_2 \le \rho_n$. In this example, the upper bound of the 2-norm distance depends on the range of model parameters, and it shrinks in the order of $\tilde{O}(\sqrt{d/\Lambda_n})$. ♣

5 Performance Guarantees

In this section, we establish performance guarantees for the PTO and the ETO approaches, based on the confidence region developed in the previous section. We consider two performance criteria defined as follows.

Let $\Phi(x;\theta)$ be the objective value of a decision x when the model is parameterized by θ. More specifically, $\Phi(x;\theta) = \psi(x)^\top\phi(f_\theta(\zeta,x))$ in PTO, and $\Phi(x;\theta) = \phi\big(f_\theta(\zeta,x)\big)$ in ETO. Let

$$\Phi_*(x) = \Phi(x;\theta^*) = \mathbb{E}_{\mathbb{P}_{\xi|\zeta}}\big[\Psi(x,\xi)\big],$$

which represents the objective value of a decision x under the ground truth model (1), parameterized by θ^*, and let x^* be the corresponding optimal solution. Let Φ_n be the optimal value of (3) or (5),

$$\Phi_n = \min_{x\in\mathcal{X}} \Phi(x;\widehat{\theta}_n),$$

which represents the optimal objective value under the estimated model, parameterized by $\widehat{\theta}_n$, and let $\widehat{x}_n$ be the corresponding optimal solution. We define the *approximation error* as

$$\Phi_n - \Phi_*(x^*),$$

which measures the difference between the optimal value of the estimated problem (3) or (5) and the ground truth model (1). This quantity is often studied in the context of consistency in stochastic programming. Moreover, we define the *regret* as

$$\Phi_*(\widehat{x}_n) - \Phi_*(x^*),$$

which measures the difference under true objective value between the estimated solution $\widehat{x}_n$ and the ground truth solution x^*. This quantity is often studied in learning theory.

Recalling ρ_n defined in Theorem 1. Below, we develop an $O(\rho_n)$-convergence rate in Section 5.1 under standard smoothness conditions, and an improved $O(\rho_n^2)$-regret bound in Section 5.2 with additional strong convexity assumptions.

5.1 Slow Rate

The following Lipschitz assumption allows us to quantify how small changes in the parameter estimates translate to changes in the decision outcomes.

Assumption 4 *Suppose that there exists $L > 0$ such that for all $\theta_1, \theta_2 \in \Theta$,*

$$\|\phi(f_{\theta_1}(\zeta, x)) - \phi(f_{\theta_2}(\zeta, x))\|_2 \le L\|\theta_1 - \theta_2\|_2, \quad \forall \zeta \in \Omega,\ \forall x \in \mathcal{X}.$$

Theorem 2 (Performance of PTO) *Under Assumptions 1-4, with probability at least $1 - 3\delta$, the approximation error of the PTO solution is bounded by*

$$|\Phi_n - \Phi_*(x^*)| \le \beta_\psi L\rho_n.$$

Moreover, with probability at least $1 - 3\delta$, the regret of the PTO solution is bounded by

$$0 \le \Phi_*(\widehat{x}_n) - \Phi_*(x^*) \le 2\beta_\psi L\rho_n.$$

Proof First, we can bound the approximation error by

$$\Phi_n - \Phi_*(x^*) = \Phi_n - \Phi(x^*; \widehat{\theta}_n) + \Phi(x^*; \widehat{\theta}_n) - \Phi_*(x^*).$$

Since $\widehat{x}_n$ optimizes (5), we have

$$\Phi_n - \Phi(x^*; \widehat{\theta}_n) \le 0.$$

By Theorem 1 and Assumption 4, it follows that the second part

$$\begin{aligned} 0 &\le \Phi(x^*; \widehat{\theta}_n) - \Phi_*(x^*) \\ &\le \|\psi(x^*)\|_2 \|\phi(f_{\widehat{\theta}_n}(\zeta, x^*)) - \phi(f_{\theta^*}(\zeta, x^*))\|_2 \\ &\le \beta_\psi \cdot L\|\widehat{\theta}_n - \theta^*\|_2 \\ &\le \beta_\psi L\rho_n. \end{aligned}$$

On the other hand, to bound the approximation error from below, we similarly have

$$\begin{aligned} \Phi_n - \Phi_*(x^*) =& \Phi_n - \Phi(\widehat{x}_n; \theta^*) + \Phi(\widehat{x}_n; \theta^*) - \Phi_*(x^*) \\ \ge& -\sup_x \|\psi(x)\|_2 \cdot L\rho_n \\ =& -\beta_\psi L\rho_n, \end{aligned}$$

where the inequality holds because $\Phi(\widehat{x}_n; \theta^*) - \Phi_*(x^*) \ge 0$ due to the optimality of x^*. Thus, we conclude that

$$|\Phi_n - \Phi_*(x^*)| \le \beta_\psi L\rho_n.$$

Next, we bound the regret similarly by

$$\begin{aligned} &\Phi_*(\widehat{x}_n) - \Phi_*(x^*) \\ =&\Phi_*(\widehat{x}_n) - \Phi(\widehat{x}_n; \widehat{\theta}_n) + \Phi(\widehat{x}_n; \widehat{\theta}_n) - \Phi(x^*; \widehat{\theta}_n) + \Phi(x^*; \widehat{\theta}_n) - \Phi_*(x^*) \\ \le&\Phi_*(\widehat{x}_n) - \Phi(\widehat{x}_n; \widehat{\theta}_n) + \Phi(x^*; \widehat{\theta}_n) - \Phi_*(x^*) \\ \le&2\beta_\psi L\rho_n, \end{aligned}$$

which gives the desired result. □

Theorem 3 (Performance of ETO) *Under Assumptions 1-4, with probability at least* $1-3\delta$*, the approximation error of ETO solution is bounded by*

$$|\Phi_n - \Phi_*(x^*)| \leq L\rho_n.$$

Moreover, with probability at least $1-3\delta$*, the regret of the ETO solution is bounded by*

$$0 \leq \Phi_*(\widehat{x}_n) - \Phi_*(x^*) \leq 2L\rho_n.$$

Proof Similarly to the proof of Theorem 2, we first bound the approximation error by

$$\Phi_n - \Phi_*(x^*) = \Phi_n - \Phi(x^*;\widehat{\theta}_n) + \Phi(x^*;\widehat{\theta}_n) - \Phi_*(x^*) \leq \Phi(x^*;\widehat{\theta}_n) - \Phi_*(x^*).$$

By Theorem 1 and Assumption 4, it follows that

$$|\Phi(x^*;\widehat{\theta}_n) - \Phi_*(x^*)| \leq L\rho_n,$$

so we have

$$\Phi_n - \Phi_*(x^*) \leq L\rho_n.$$

On the other hand, to bound the approximation error from below, we similarly have

$$\Phi_n - \Phi_*(x^*) = \Phi_n - \Phi(\widehat{x}_n;\theta^*) + \Phi(\widehat{x}_n;\theta^*) - \Phi_*(x^*) \geq -L\rho_n.$$

Next, we bound the regret by

$$\begin{aligned}
&\Phi_*(\widehat{x}_n) - \Phi_*(x^*)\\
=&\Phi_*(\widehat{x}_n) - \Phi(\widehat{x}_n;\widehat{\theta}_n) + \Phi(\widehat{x}_n;\widehat{\theta}_n) - \Phi(x^*;\widehat{\theta}_n) + \Phi(x^*;\widehat{\theta}_n) - \Phi_*(x^*)\\
\leq&2L\rho_n,
\end{aligned}$$

which completes the proof. □

Under the additional assumption on the Lipschitz continuity of the objective with respect to the parameter, both Theorem 2 and Theorem 3 demonstrate that the approximate error and the regret of the PTO and the ETO approaches are $O(\rho_n)$, which is $O(\sqrt{d\mathfrak{C}_{\mathcal{F}}/\Lambda_n})$.

5.2 Fast rate

In this subsection, we show that much faster rates of PTO and ETO can be achieved under certain strong convexity conditions.

The following theorem provides a faster convergence rate of both methods when $\Phi(x;\theta)$ exhibits certain nice structures.

Theorem 4 *Assume* $\Phi(\cdot;\theta^*)$ *is* α_Φ*-strongly convex and* $\hbar_\Phi$*-smooth in* x*, and* $\Phi(x;\theta)$ *has* L_1*-Lipschitz gradient with respect to* x *for all* $\theta \in \Theta$*. Suppose* $\widehat{x}_n$ *is an interior point of* $\mathcal{X}$*. Then, under the same setting of Theorem 1, for both PTO and ETO, we have the following regret bound*

$$\Phi_*(\widehat{x}_n) - \Phi_*(x^*) \leq \frac{2\hbar_\Phi L_1^2}{\alpha_\Phi^2}\rho_n^2.$$

Proof The first order optimality of $\widehat{x}_n$ and x^* under parameter $\widehat{\theta}_n$ and θ^*, respectively, reads

$$\nabla_x \Phi(\widehat{x}_n; \widehat{\theta}_n) = 0, \quad \text{and} \quad \nabla_x \Phi(x^*; \theta^*) = 0.$$

By the α_Φ-strong convexity of $\Phi(x; \theta^*)$ with respect to x, we have that

$$\begin{aligned} \frac{\alpha_\Phi}{2} \|\widehat{x}_n - x^*\|_2 &\leq \left\|\nabla_x \Phi(\widehat{x}_n; \theta^*) - \nabla_x \Phi(x^*; \theta^*)\right\|_2 \\ &= \left\|\nabla_x \Phi(\widehat{x}_n; \theta^*)\right\|_2. \end{aligned}$$

Since $\Phi(x; \theta)$ has Lipschitz gradient with respect to x, we have that

$$\|\nabla_x \Phi(\widehat{x}_n; \theta^*)\|_2 = \|\nabla_x \Phi(\widehat{x}_n; \theta^*) - \nabla_x \Phi(\widehat{x}_n; \widehat{\theta}_n)\|_2 \leq L_1 \|\theta^* - \widehat{\theta}_n\|_2.$$

It follows that

$$\|\widehat{x}_n - x^*\|_2 \leq \frac{2L_1}{\alpha_\Phi} \|\widehat{\theta}_n - \theta^*\|_2.$$

Using the smoothness of $\Phi(\cdot; \theta^*)$, we conclude that

$$\Phi_*(\widehat{x}_n) - \Phi_*(x^*) \leq \frac{\hbar_\Phi}{2} \|\widehat{x}_n - x^*\|_2^2 \leq \frac{2\hbar_\Phi L_1^2}{\alpha_\Phi^2} \|\widehat{\theta}_n - \theta^*\|_2^2 \leq \frac{2\hbar_\Phi L_1^2}{\alpha_\Phi^2} \rho_n^2.$$

□

Theorem 4 demonstrates a faster $O(\rho_n^2)$ convergence rate for both PTO and ETO, given additional conditions involving strong convexity of the objective function in decisions and the Lipschitz continuity of its gradient. The intuition is that these additional conditions ensure the smoothness of the optimal solution with respect to the parameters. We demonstrate this result in the next example.

Example 6 Consider the linear demand model in Example 1. The objective function, representing the negative revenue, is given by $\Phi(x; \theta) = -x\theta^\top [1; \zeta; x] = -(\theta_0 + \theta_1^\top \zeta)x - \theta_2 x^2$, where the price sensitivity $\theta_2 \in [\theta_{\min}, \theta_{\max}]$ with $\theta_{\min}, \theta_{\max} < 0$. Suppose θ, ζ, x are all bounded. Let us verify the conditions in Theorem 4. Since $\nabla_x^2 \Phi(x; \theta^*) = -2\theta_2^*$, the true objective function is strongly convex and smooth in x. Furthermore, since $\nabla_x \Phi(x; \theta) = -(\theta_0 + \theta_1^\top \zeta) - 2\theta_2 x$, $\Phi(x; \theta)$ is L_1-Lipschitz where $L_1 = \sup_{x \in \mathcal{X}, \zeta \in \Omega, \theta \in \Theta} |\theta_0 + \theta_1^\top \zeta + 2\theta_2 x|$. For an estimator $\widehat{\theta}$, the best price is $\widehat{x} = -(\widehat{\theta}_0 + \widehat{\theta}_1^\top \zeta)/(2\widehat{\theta}_2)$. Suppose the possible price range is sufficiently large, i.e., $\sup_{x \in \mathcal{X}} x < \sup_{\theta \in \Theta} -(\theta_0 + \theta_1^\top \zeta)/(2\theta_2)$, then $\widehat{x}$ is always an interior point of $\mathcal{X}$. ♣

In [22], it is shown that in a decision-independent i.i.d. setting with a linear objective and a polytope feasible region for the decision, the PTO approach can achieve a faster convergence rate if the problem instances do not exhibit arbitrarily bad near-dual-degeneracy. In their setting, the objective function is linear in the decision, and their assumption of non-degeneracy implies a superlinear dependence of the regret on the parameter estimation error. Comparatively, in our decision-dependent setting, the objective function is no longer linear in the decision; for instance, in Example 6, the revenue function is quadratic in the pricing decision. Instead, the strong convexity and Lipschitz gradient conditions imply a quadratic dependence of the regret on the parameter estimation error. Nevertheless, the underlying insights are similar and this insight also applies to our described PTO and ETO approaches.

6 Proofs for Section 4

6.1 Proof of Lemma 1(I)

This proof relies on the following result for self-normalized random variables.

Lemma 2 (Corollary 2.2 in [37]) *Suppose random variables W_1 and W_2 satisfies for all $u \in \mathbb{R}$,*

$$\mathbb{E}\left[\exp\left(uW_1 - \frac{u^2}{2}W_2^2\right)\right] \le 1.$$

Then for any $c \ge \sqrt{2}$ and $\epsilon > 0$,

$$\mathbb{P}\left\{|W_1| \ge c\sqrt{(W_2^2+\epsilon)\Big(1+\frac{1}{2}\log\big(\frac{W_2^2}{\epsilon}+1\big)\Big)}\right\} \le \exp(-c^2/2).$$

Equipped with the above concentration inequality, let us start the proof by introducing some notations. Fix θ. Let $\nu \in \mathbb{R}^d$ whose value will be specified later. Define

$$\begin{aligned} W_1 &:= \nu^\top Z_n(\theta), \\ W_2 &:= \sqrt{2}\sigma\|\nu\|_{V_n(\theta)}. \end{aligned}$$

Recall $Z_n(\theta) = \sum_{i=1}^n \nabla f_\theta(\zeta_i, X_i)^\top \epsilon_i$ and $V_n(\theta) = \sum_{i=1}^n \nabla f_\theta(\zeta_i, X_i)^\top \nabla f_\theta(\zeta_i, X_i)$. For any $u \ge 0$, we have

$$\begin{aligned} uW_1 - \frac{u^2}{2}W_2^2 =& u\nu^\top Z_n(\theta) - \frac{u^2\nu^\top V_n(\theta)\nu}{2} \\ =& \sum_{i=1}^n \left(u\nu^\top \nabla f_\theta(\zeta_i, X_i)^\top \epsilon_i - \frac{u^2}{2}\cdot 2\sigma^2 \cdot \nu^\top \nabla f_\theta(\zeta_i, X_i)^\top f_\theta(\zeta_i, X_i)\nu\right). \end{aligned}$$

Define $D_i = u\nu^\top \nabla f_\theta(\zeta_i, X_i)^\top \epsilon_i - \frac{u^2}{2}\cdot 2\sigma^2 \cdot \nu^\top \nabla f_\theta(\zeta_i, X_i)^\top f_\theta(\zeta_i, X_i)\nu$. Note that conditioning on $\mathcal{H}_{i-1}$ and θ, the randomness of D_i comes from ϵ_i only. Thus,

$$\begin{aligned} &\mathbb{E}\left[\exp\left(u\nu^\top \nabla f_\theta(\zeta_i, x_i)^\top \epsilon_i - \frac{u^2}{2}\cdot 2\sigma^2 \cdot \nu^\top \nabla f_\theta(\zeta_i, x_i)^\top \nabla f_\theta(\zeta_i, x_i) w\right) \mid \mathcal{H}_i\right] \\ &\le \mathbb{E}\left[\exp\left(u\nu^\top \nabla f_\theta(\zeta_i, x_i)^\top \epsilon_i\right) \mid \mathcal{H}_i\right] \exp\left(-u^2\sigma^2\nu^\top \nabla f_\theta(\zeta_i, X_i)^\top f_\theta(\zeta_i, X_i)\nu\right) \\ &\le \exp\left(-\frac{1}{2}u^2\sigma^2\nu^\top \nabla f_\theta(\zeta_i, X_i)^\top \nabla f_\theta(\zeta_i, X_i)\nu - u^2\sigma^2\nu^\top \nabla f_\theta(\zeta_i, X_i)^\top \nabla f_\theta(\zeta_i, X_i)\nu\right) \\ &\le 1. \end{aligned}$$

It follows that

$$\mathbb{E}[\exp(D_i) \mid \mathcal{H}_i] \le 1.$$

Using the tower property of conditional expectations and the inequality above, we obtain that

$$\mathbb{E}\Big[\prod_{i=1}^n \exp(D_i)\Big] = \mathbb{E}\Big[\prod_{i=1}^{n-1} \exp(D_i)\mathbb{E}[\exp(D_n \mid \mathcal{H}_n]\Big] \le \mathbb{E}\Big[\prod_{i=1}^n \exp(D_i)\Big].$$

Applying this inequality recursively yields

$$\mathbb{E}\Big[\exp(uW_1 - \frac{u^2}{2}W_2^2)\Big] \le \mathbb{E}\Big[\prod_{i=1}^n \exp(D_i)\Big] \le \cdots \le \mathbb{E}[\exp(D_1)] \le 1.$$

The derivation above verifies the condition required by Lemma 2. Set $\epsilon = 2\sigma^2\lambda_{\min}(V_n(\theta))\|\nu\|_2^2$ in Lemma 2, then for any $0 < \delta \leq 1/e$ and $n \geq 1$, with probability $1-\delta$,

$$\begin{aligned}&|\nu^\top Z_n(\theta)|\\ \leq&\sqrt{2}\sqrt{(2\sigma^2\|\nu\|_{V_n(\theta)}^2 + 2\sigma^2\lambda_{\min}(V_n(\theta))\|\nu\|_2^2)\Big(1+\tfrac{1}{2}\log\big(1+\tfrac{\|\nu\|_{V_n(\theta)}^2}{\lambda_{\min}(V_n(\theta))\|\nu\|_2^2}\big)\Big)}\\ &\cdot\sqrt{2\log(1/\delta)}.\end{aligned} \tag{8}$$

Note that for $n \geq \max(d,2)$, $\lambda_{\min}(V_n(\theta))\|\nu\|_2^2 \leq \|\nu\|_{V_n(\theta)}^2 \leq n\|\nu\|_2^2\kappa_f^2$, we have $\|\nu\|_{V_n(\theta)}^2 + \lambda_{\min}(V_n(\theta))\|\nu\|_2^2 \leq 2\|\nu\|_{V_n(\theta)}^2$ and $1+\frac{1}{2}\log\left(1+\frac{\|\nu\|_{V_n(\theta)}^2}{\lambda_{\min}(V_n(\theta))}\right) \leq 1+\frac{1}{2}\log(1+\frac{n\kappa_f^2}{\lambda_{\min}(V_n(\theta))}) \leq \eta^2\log(n)/2$ where $\eta = \sqrt{3+2\log(1+2\kappa_f^2/\lambda_{\min}(V_n(\theta)))}$. Therefore,

$$\begin{aligned}|\nu^\top Z_n(\theta)| &\leq 2\sigma\sqrt{2\log(1/\delta)\cdot 2\|\nu\|_{V_n(\theta)}^2\Big(1+\tfrac{1}{2}\log\big(1+\tfrac{n\kappa_f^2}{\lambda_{\min}(V_n(\theta))}\big)\Big)}\\ &\leq 4\sigma\eta\|\nu\|_{V_n(\theta)}\sqrt{\log\tfrac{1}{\delta}\log n}.\end{aligned} \tag{9}$$

Now we specify the value of ν. Let $\nu = V_n(\theta)^{-1/2}e_j$. Observe that

$$\begin{aligned}&\|Z_n(\theta)\|_{V_n(\theta)^{-1}}^2 = Z_n(\theta)^\top V_n(\theta)^{-1}Z_n(\theta) = Z_n(\theta)^\top V_n(\theta)^{-1/2}IV_n(\theta)^{-1/2}Z_n(\theta)\\ =&\sum_{j=1}^d Z_n(\theta)^\top V_n(\theta)^{-1/2}e_je_j^\top V_n(\theta)^{-1/2}Z_n(\theta),\end{aligned}$$

where $\{e_j\}_{j=1}^d$ denotes the standard orthonormal basis in $\mathbb{R}^d$. Thus, for any constant $c > 0$ it holds that

$$\begin{aligned}\mathbb{P}\left\{\|Z_n(\theta)\|_{V_n(\theta)^{-1}}^2 \geq dc^2\right\} &= \mathbb{P}\left\{\sum_{j=1}^d Z_n(\theta)^\top V_n(\theta)^{-1/2}e_je_j^\top V_n(\theta)^{-1/2}Z_n(\theta) \geq dc^2\right\}\\ &\leq \sum_{j=1}^d \mathbb{P}\left\{Z_n(\theta)^\top V_n(\theta)^{-1/2}e_je_j^\top V_n(\theta)^{-1/2}Z_n(\theta) \geq c^2\right\}\\ &\leq \sum_{j=1}^d \mathbb{P}\left\{\|Z_n(\theta)^\top V_n(\theta)^{-1/2}e_j\|_2 \geq c\right\}.\end{aligned}$$

Set $c = 4\eta\sigma\sqrt{\log n\log(d/\delta)}$, $j = 1,\ldots,d$, in the above inequality, we conclude that

$$\begin{aligned}&\mathbb{P}\{\|Z_n(\theta)\|_{V_n(\theta)^{-1}}^2 \geq dc^2\}\\ =&\mathbb{P}\{\|Z_n(\theta)\|_{V_n(\theta)^{-1}}^2 \geq 16d\eta^2\sigma^2\log(n)\log(d/\delta)\}\\ \leq&\mathbb{E}[\mathbf{1}(\|Z_n(\theta)\|_{V_n(\theta)^{-1}}^2 \geq 16d\eta^2\sigma^2\log(n)\log(d/\delta))]\\ \leq&\sum_{j=1}^d \mathbb{E}\left[\mathbf{1}\left(|Z_n(\theta)^\top V_n(\theta)^{-1/2}e_j| \geq c\right)\right]\\ =&\sum_{j=1}^d \mathbb{P}\left\{|Z_n(\theta)^\top V_n(\theta)^{-1/2}e_j| \geq c\right\} \leq \delta.\end{aligned}$$

The proof is completed. □

6.2 Proof of Lemma 1(II)

Before the proof, we introduce the following tail bounds for the maximum of sub-Gaussian random vector variables.

Lemma 3 (Maximum of sub-Gaussian variables) *Let $\{\varepsilon_i\}_{1\le i\le n}$ be a sequence of zero-mean sub-Gaussian random vector variables with parameter σ. Then for every $\delta > 0$, it holds with probability $1-\delta$ that*

$$\max_{1\le i\le n} \|\varepsilon_i\|_2 \le 2\sqrt{2}\sigma\sqrt{\log(1/\delta) + d_{\mathcal{Y}}\log 6 + \log n}.$$

Proof Let $\mathbb{B}_1^{d_{\mathcal{Y}}}$ be a ball with radius 1. Let $\mathcal{N}$ be a 1/2-net of $\mathbb{B}_1^{d_{\mathcal{Y}}}$ with respect to the Euclidean norm that satisfies $|\mathcal{N}| \le 6^{d_{\mathcal{Y}}}$. Next, observe that for every $y \in \mathbb{B}_1^{d_{\mathcal{Y}}}$, there exists $a \in \mathcal{N}$ and b such that $\|v\|_2 \le 1/2$ and $y = a + b$. Therefore,

$$\max_{y\in\mathbb{B}_1^{d_{\mathcal{Y}}}} y^\top\varepsilon \le \max_{a\in\mathcal{N}} a^\top\varepsilon + \max_{b\in\frac{1}{2}\mathbb{B}_1^{d_{\mathcal{Y}}}} b^\top\varepsilon.$$

By the equality

$$\max_{b\in\frac{1}{2}\mathbb{B}_1^{d_{\mathcal{Y}}}} b^\top\varepsilon = \frac{1}{2}\max_{c\in\frac{1}{2}\mathbb{B}_1^{d_{\mathcal{Y}}}} c^\top\varepsilon,$$

we have

$$\max_{z\in\mathbb{B}_1^{d_{\mathcal{Y}}}} y^\top\varepsilon \le 2\max_{z'\in\mathcal{N}} z'^\top\varepsilon.$$

Therefore, by maximal inequality of sub-Gaussian variables [7, Theorem 5.2], we get

$$\mathbb{E}\left[\max_{y\in\mathbb{B}_1^{d_{\mathcal{Y}}}} y^\top\varepsilon\right] \le 2\mathbb{E}\left[\max_{y'\in\mathcal{N}} y'^\top\varepsilon\right] \le 2\mathbb{E}\left[\max_{y'\in\mathcal{N}} \frac{y'^\top\varepsilon}{\|y'\|_2}\right]. \tag{10}$$

The bound with high probability follows because

$$\begin{aligned}
\mathbb{P}\left(\max_{y\in\mathbb{B}_1^{d_{\mathcal{Y}}}} y^\top\varepsilon > t\right) &\le \mathbb{P}\left(2\max_{y'\in\mathcal{N}} y'^\top\varepsilon > t\right)\\
&\le \sum_{y'\in\mathcal{N}} \mathbb{P}\left(\frac{y'^\top\varepsilon}{\|y'\|_2} > \frac{t}{2}\right)\\
&\le |\mathcal{N}| e^{-\frac{t^2}{8\sigma^2}}\\
&\le 6^{d_{\mathcal{Y}}} e^{-\frac{t^2}{8\sigma^2}}.
\end{aligned}$$

Now we have n Gaussian random variables in total. Since ε_i's are sub-Gaussian, for each $i = 1, 2, \ldots, n$, we have $\max_{z\in\mathbb{B}_1^{d_{\mathcal{Y}}}} z^\top\varepsilon_i = \|\varepsilon_i\|_2$. Then, due to that for any $y \in \mathbb{B}_1^{d_{\mathcal{Y}}}$, $y^\top\varepsilon_i$ is a sub-Gaussian variable with the parameter σ, it follows that

$$\mathbb{P}\left\{\max_{1\le i\le n} \|\varepsilon_i\|_2 \ge t\right\} \le \sum_{i=1}^{n-1} \mathbb{P}\left\{\max_{z\in\mathbb{B}_1^{d_{\mathcal{Y}}}} y^\top\varepsilon_i \ge t\right\} \le n6^{d_{\mathcal{Y}}} e^{-\frac{t^2}{8\sigma^2}}.$$

Taking $t = 2\sqrt{2}\sigma\sqrt{\log(1/\delta) + d_{\mathcal{Y}}\log 6 + \log n}$, it holds with probability $1-\delta$ that

$$\max_{1\le i\le n} \|\varepsilon_i\|_2 \le 2\sqrt{2}\sigma\sqrt{\log(1/\delta) + d_{\mathcal{Y}}\log 6 + \log n}.$$

This completes the proof. □

With Lemma 3, we now prove Lemma 1, using a covering number argument. Define $H(\theta) = \|Z_n(\theta)\|^2_{V_n(\theta)^{-1}}$, set $\bar{\epsilon} = 2\sqrt{2}\sigma\sqrt{\log(1/\delta) + d_{\mathcal{Y}} \log 6 + \log n}$, and define events $\mathcal{B}_n = \left\{\{\epsilon_i\}_{i=1}^n : \max_{1\le i\le n} \ \|\epsilon_i\|_2 \le \bar{\epsilon}\right\}$. Lemma 3 states that event $\mathcal{B}_n$ holds with probability at least $1-\delta$. We first construct ω−net where $\omega = \log n$. For any θ_1 and θ_2,

$$\begin{aligned}
&|H(\theta_1) - H(\theta_2)| \\
=& \left| \|Z_n(\theta_1)\|^2_{V_n(\theta_1)^{-1}} - \|Z_n(\theta_2)\|^2_{V_n(\theta_2)^{-1}} \right| \\
=& \left| \|Z_n(\theta_1)\|^2_{V_n(\theta_1)^{-1}} - \|Z_n(\theta_1)\|^2_{V_n(\theta_2)^{-1}} + \|Z_n(\theta_1)\|^2_{V_n(\theta_2)^{-1}} - \|Z_n(\theta_2)\|^2_{V_n(\theta_2)^{-1}} \right| \\
\le& \left| \|Z_n(\theta_1)\|^2_{V_n(\theta_1)^{-1}} - \|Z_n(\theta_1)\|^2_{V_n(\theta_2)^{-1}} \right| + \left| \|Z_n(\theta_1)\|^2_{V_n(\theta_2)^{-1}} - \|Z_n(\theta_2)\|^2_{V_n(\theta_2)^{-1}} \right|.
\end{aligned}$$

For ease of notation, let $u_1 = Z_n(\theta_1)$, $u_2 = Z_n(\theta_2)$, $V_1 = V_n(\theta_1)$, and $V_2 = V_n(\theta_2)$. We first note that

$$\begin{aligned}
&\left| \|u_1\|^2_{V_2^{-1}} - \|u_2\|^2_{V_2^{-1}} \right| \mathbf{1}(\mathcal{B}_n) \\
=&(u_1 - u_2)^\top V_2^{-1}(u_1 + u_2)\mathbf{1}(\mathcal{B}_n) \\
\le&\|u_1 - u_2\|_2 \mathbf{1}(\mathcal{B}_n) \max_{\theta\in\Theta} \ 2\|V_2^{-1} Z_n(\theta)\|_2 \\
=&\| \sum_{i=1}^n \nabla f_{\theta_1}(\zeta_i, X_i)^\top \epsilon_i - \sum_{i=1}^n \nabla f_{\theta_2}(\zeta_i, X_i)^\top \epsilon_i \|_2 \mathbf{1}(\mathcal{B}_n) \max_{\theta\in\Theta} \ 2\|V_2^{-1} Z_n(\theta)\|_2 \\
\le& \sum_{i=1}^n \|\nabla f_{\theta_1}(\zeta_i, X_i)^\top \epsilon_i - \nabla f_{\theta_2}(\zeta_i, X_i)^\top\|_2 \|\epsilon_i\|_2 \mathbf{1}(\mathcal{B}_n) \max_{\theta\in\Theta} \ 2\|V_2^{-1} Z_n(\theta)\|_2 \\
\le& \sum_{i=1}^n \|\nabla f_{\theta_1}(\zeta_i, X_i) - \nabla f_{\theta_2}(\zeta_i, X_i)\|_F \bar{\epsilon} \mathbf{1}(\mathcal{B}_n) \max_{\theta\in\Theta} \ 2\|V_2^{-1} Z_n(\theta)\|_2 \\
\le& \sum_{i=1}^n \|\nabla f_{\theta_1}(\zeta_i, X_i) - \nabla f_{\theta_2}(\zeta_i, X_i)\|_F \bar{\epsilon} \cdot \frac{2n\kappa_f \bar{\epsilon}}{\Lambda_n},
\end{aligned}$$

where the last inequality holds because $\|Z_n(\theta)\|\mathbf{1}(\mathcal{B}_n) \le n\kappa_f \bar{\epsilon}$.

Note that $\nabla f_\theta(\zeta_i, X_i)$ is a matrix with the dimension $d_{\mathcal{Y}} \times d$. Define $\nabla f_\theta(\zeta_i, X_i) = [\nabla f_\theta(\zeta_i, X_i)_{jk}]_{1\le j\le d_{\mathcal{Y}}, 1\le k\le d}$. For $j = 1, \cdots, d_{\mathcal{Y}}$ and $k = 1, \cdots, d$, by Mean Value Theorem, there exists a $\bar{\theta}_{jk}$ which is a convex combination of θ_1 and θ_2 such that

$$|\nabla f_{\theta_1}(\zeta_i, X_i)_{jk} - \nabla f_{\theta_2}(\zeta_i, X_i)_{jk}| = |(\nabla(\nabla f_{\bar{\theta}_{jk}}(\zeta_i, X_i)_{jk})^\top (\theta_1 - \theta_2)| \le \hbar_f \|\theta_1 - \theta_2\|_2.$$

Then, we have

$$\|\nabla f_{\theta_1}(\zeta_i, X_i) - \nabla f_{\theta_2}(\zeta_i, X_i)\|_F \le \hbar_f \sqrt{d d_{\mathcal{Y}}} \|\theta_1 - \theta_2\|_2,$$

for $j = 1, \cdots, d_{\mathcal{Y}}$ and $k = 1, \cdots, d$. Hence, it follows that

$$\begin{aligned}
&\left| \|u_1\|^2_{V_2^{-1}} - \|u_2\|^2_{V_2^{-1}} \right| \mathbf{1}(\mathcal{B}_n) \\
\le& \sum_{i=1}^n \|\nabla f_{\theta_1}(\zeta_i, X_i) - \nabla f_{\theta_2}(\zeta_i, X_i)\|_F \bar{\epsilon} \cdot \frac{2n\kappa_f \bar{\epsilon}}{\Lambda_n} \\
\le& n\hbar_f \sqrt{d d_{\mathcal{Y}}} \|\theta_1 - \theta_2\|_2 \bar{\epsilon} \cdot \frac{2n\kappa_f \bar{\epsilon}}{\Lambda_n} \\
=& \frac{2\kappa_f \hbar_f (n\bar{\epsilon})^2 \sqrt{d d_{\mathcal{Y}}}}{\Lambda_n} \|\theta_1 - \theta_2\|_2,
\end{aligned}$$

Now we analyze $\left|\|u_1\|^2_{V_1^{-1}} - \|u_1\|^2_{V_2^{-1}}\right|$. Let $W = V_n(\theta)^{-1}$. By chain rule, we have

$$\begin{aligned}\partial_{\theta_j}(\|u_1\|^2_{V_n(\theta)^{-1}})\mathbf{1}(\mathcal{B}_n) &= \operatorname{tr}\left(\nabla_W(\|u_1\|^2_W)\mid_{W=V_n(\theta)^{-1}}\cdot\partial_{\theta_j}(V_n(\theta)^{-1})\right)\mathbf{1}(\mathcal{B}_n)\\ &= \operatorname{tr}\left(u_1u_1^\top\partial_{\theta_j}(V_n(\theta)^{-1})\right)\mathbf{1}(\mathcal{B}_n),\end{aligned}$$

where the first equality holds due to Equation (137) in [39]. According to Equation (59) in [39] that

$$\partial_{\theta_j}V_n(\theta)^{-1} = -V_n(\theta)^{-1}\partial_{\theta_j}V_n(\theta)V_n(\theta)^{-1}, \quad j = 1,\ldots,d,$$

we have

$$\begin{aligned}&\left|\partial_{\theta_j}(\|u_1\|^2_{V_n(\theta)^{-1}})\right|\mathbf{1}(\mathcal{B}_n)\\ =&\left|\operatorname{tr}\left(u_1u_1^\top\partial_{\theta_j}(V_n(\theta)^{-1})\right)\right|\mathbf{1}(\mathcal{B}_n)\\ =&\left|\operatorname{tr}\left(u_1u_1^\top V_n(\theta)^{-1}\partial_{\theta_j}V_n(\theta)V_n(\theta)^{-1}\right)\right|\mathbf{1}(\mathcal{B}_n)\\ \le&\|u_1\|_2^2\|V_n(\theta)^{-1}\|_2^2\left\|\partial_{\theta_j}\Big(\sum_{i=1}^n\nabla f_\theta(\zeta_i,X_i)^\top\nabla f_\theta(\zeta_i,X_i)\Big)\right\|_2\mathbf{1}(\mathcal{B}_n)\\ \le&\frac{\|u_1\|_2^2}{\lambda_{\min}(V_n(\theta))^2}\left\|\partial_{\theta_j}\Big(\sum_{i=1}^n\nabla f_\theta(\zeta_i,X_i)^\top\nabla f_\theta(\zeta_i,X_i)\Big)\right\|_2.\end{aligned}$$

It implies that

$$\|\nabla_\theta(\|u_1\|^2_{V_n(\theta)^{-1}})\|_2\mathbf{1}(\mathcal{B}_n) \le \frac{\|u_1\|_2^2}{\lambda_{\min}(V_n(\theta))^2}\cdot\left\|\nabla_\theta\Big((\sum_{i=1}^n\nabla f_\theta(\zeta_i,X_i)\nabla f_\theta(\zeta_i,X_i)^\top)\Big)\right\|_2,$$

where $\nabla_\theta(V_n(\theta)^{-1})$ is a $d\times d\times d$ tensor and the expansion along the last dimension has the form $\nabla_\theta(V_n(\theta)^{-1}) = (\partial_{\theta_1}(V_n(\theta)^{-1}),\ldots,\partial_{\theta_d}(V_n(\theta)^{-1}))$. Let $\nabla f_\theta(\zeta_i,X_i) = [\nabla f_\theta(\zeta_i,X_i)_{jk}]_{1\le j\le d_{\mathcal{Y}},1\le k\le d}$. Note that

$$\begin{aligned}\nabla_\theta(\nabla f_\theta(\zeta_i,X_i)^\top\nabla f_\theta(\zeta_i,X_i))_{k,l} &= \nabla_\theta(\sum_{\ell=1}^{d_{\mathcal{Y}}}\nabla f_\theta(\zeta_i,X_i)_{\ell j}\nabla f_\theta(\zeta_i,X_i)_{\ell k})\\ =\sum_{\ell=1}^{d_{\mathcal{Y}}}&\left(\nabla^2 f_\theta(\zeta_i,X_i)_{\ell j}\nabla f_\theta(\zeta_i,X_i)_{\ell k} + \nabla f_\theta(\zeta_i,X_i)_{\ell j}\nabla^2 f_\theta(\zeta_i,X_i)_{\ell k}\right),\end{aligned}$$

which implies that

$$\left\|\nabla_\theta\Big((\sum_{i=1}^n\nabla f_\theta(\zeta_i,X_i)^\top\nabla f_\theta(\zeta_i,X_i))\Big)\right\|_2 \le 2nd^2d_{\mathcal{Y}}\kappa_f\hbar_f.$$

Thus, we can bound $\left\|\nabla_\theta(\|u_1\|^2_{V_n(\theta)^{-1}})\right\|_2$ by

$$\left\|\nabla_\theta(\|u_1\|^2_{V_n(\theta)^{-1}})\right\|_2\mathbf{1}(\mathcal{B}_n) \le 2\frac{(n\bar{\epsilon}\kappa_f)^2\cdot nd^2d_{\mathcal{Y}}\kappa_f\hbar_f}{\Lambda_n^2} = \frac{2\bar{\epsilon}^2\kappa_f^3\hbar_f n^3d^2d_{\mathcal{Y}}}{\Lambda_n^2}.$$

Then,

$$\begin{aligned}|H(\theta_1)-H(\theta_2)|\mathbf{1}(\mathcal{B}_n) &\le \left(\frac{2\kappa_f\hbar_f n^2\bar{\epsilon}^2\sqrt{dd_{\mathcal{Y}}}}{\Lambda_n} + \frac{2\bar{\epsilon}^2\hbar_f^3\kappa_f n^3d^2d_{\mathcal{Y}}}{\Lambda_n^2}\right)\|\theta_1-\theta_2\|_2\\ &= L_0\|\theta_1-\theta_2\|_2,\end{aligned}$$

where $L_0 = \frac{2\kappa_f \hbar_f n^2 \bar{\epsilon}^2 \sqrt{d d_{\mathcal{Y}}}}{\Lambda_n} + \frac{2\hbar_f^3 \kappa_f \bar{\epsilon}^2 n^3 d^2 d_{\mathcal{Y}}}{\Lambda_n^2} = \frac{2\kappa_f \hbar_f n^2 \bar{\epsilon}^2 \sqrt{d d_{\mathcal{Y}}}}{\Lambda_n} \left(1 + \frac{\hbar_f^2 d \sqrt{d d_{\mathcal{Y}}} n}{\Lambda_n}\right)$.

From Lemma 1(I), we conclude that for any $\gamma > 0$,

$$\mathbb{P}\{H(\theta)\mathbf{1}(\mathcal{B}_n) > dc\gamma\} \le d\exp(-\gamma),$$

where $c = 16\eta^2\sigma^2\log(n)$. Define $\gamma = 2d\log(L_0\beta_\Theta/\omega) + \log(d/\delta)$. Let Θ_ω be the $\omega-$covering set for Θ regarding function $H(\cdot)$, i.e., for any $\theta \in \Theta$, there exists $\theta' \in \Theta_\omega$ such that $\|H(\theta) - H(\theta')\|_2 \le \omega$. Define $N(\omega; H, \|\cdot\|_2)$ as the $\omega-$covering number. Therefore,

$$\begin{aligned}
\mathbb{P}\{H(\widehat{\theta}_n)\mathbf{1}(\mathcal{B}_n) > 2dc\gamma\} &\le \mathbb{P}(\exists\theta, H(\theta)\mathbf{1}(\mathcal{B}_n) > 2dc\gamma) \\
&\le \mathbb{P}(\exists\theta \in \Theta_\omega, H(\theta)\mathbf{1}(\mathcal{B}_n) > dc\gamma) \\
&\le 2N(\omega; H, \|\cdot\|_2) d\exp(-\gamma) \\
&\le 2\left(\frac{L_0\beta_\Theta}{\omega}\right)^d d\exp(-\gamma) \\
&= 2\exp(d\log(L_0\beta_\Theta/\omega) + \log d - \gamma) \\
&\le 2\exp(-\log(1/\delta)) \le 2\delta,
\end{aligned}$$

where the second inequality comes from the fact $\omega \le dc\gamma$ since $n \ge 1/4\eta\sigma\sqrt{d\gamma}$.

Due to that $\mathcal{B}_n$ holds with probability at least $1-\delta$, by the union bound, we conclude that

$$\begin{aligned}
&\mathbb{P}\{H(\widehat{\theta}_n) > 2dc\gamma\} \\
=&\mathbb{P}\{H(\widehat{\theta}_n) > 2dc\gamma \text{ and } \mathbf{1}(\mathcal{B}_n) = 1\} + \mathbb{P}\{H(\widehat{\theta}_n) > 2dc\gamma \text{ and } \mathbf{1}(\mathcal{B}_n) = 0\} \le 3\delta.
\end{aligned}$$

It implies that

$$\begin{aligned}
&3\delta \\
\ge&\mathbb{P}(H(\widehat{\theta}_n) > 2dc\gamma) \\
=&\mathbb{P}\Big\{H(\widehat{\theta}_n) > 2dc\Big(2d\log\Big(\tfrac{2\kappa_f\hbar_f n^2\bar{\epsilon}^2\sqrt{dd_{\mathcal{Y}}}\beta_\Theta}{\omega\Lambda_n}\cdot\Big(1+\tfrac{\hbar_f^2 d\sqrt{dd_{\mathcal{Y}}}n}{\Lambda_n}\Big)\Big) + \log(d/\delta)\Big)\Big\} \\
\ge&\mathbb{P}\Big\{H(\widehat{\theta}_n) > 2dc\Big(2d\log\Big(\tfrac{2\kappa_f\hbar_f n^3\bar{\epsilon}^2 d^2 d_{\mathcal{Y}}\beta_\Theta}{\log(n)\Lambda_n}\cdot\Big(1+\tfrac{\hbar_f^2 d\sqrt{dd_{\mathcal{Y}}}}{\Lambda_n}\Big)\Big) + \log(d/\delta)\Big)\Big\} \\
=&\mathbb{P}\Big\{H(\widehat{\theta}_n) > 4d^2c\Big(\log\Big(\tfrac{16\sigma^2\big(\log(1/\delta)+d_{\mathcal{Y}}\log 6+\log(n)\big)\kappa_f\hbar_f n^3 d^2 d_{\mathcal{Y}}\beta_\Theta}{\log(n)\Lambda_n}\Big(1+\tfrac{\hbar_f^2}{\Lambda_n}\Big)\Big) \\
&\quad + \log(d/\delta)\Big)\Big\} \\
\ge&\mathbb{P}\Big\{H(\widehat{\theta}_n) > 4d^2c\Big(\log\Big(\tfrac{16\sigma^2\big(6\log(1/\delta)\big)\kappa_f\hbar_f\log(n)n^3 d^2 d_{\mathcal{Y}}^2\beta_\Theta}{\log(n)\Lambda_n}\Big(1+\tfrac{\hbar_f^2}{\Lambda_n}\Big)\Big) + \log(d/\delta)\Big)\Big\} \\
=&\mathbb{P}\Big\{H(\widehat{\theta}_n) > 4d^2c\Big(\log\Big(\tfrac{16\sigma^2\big(6\log(1/\delta)\big)\kappa_f\hbar_f n^3 d^3 d_{\mathcal{Y}}^2\beta_\Theta}{\Lambda_n\delta}\Big(1+\tfrac{\hbar_f^2}{\Lambda_n}\Big)\Big)\Big)\Big\} \\
\ge&\mathbb{P}\Big\{H(\widehat{\theta}_n) > 64\eta^2\sigma^2 d^2\log(n)\log(n^3 d^3 d_{\mathcal{Y}}^2\log(1/\delta)\varsigma/(\delta\Lambda_n))\Big\},
\end{aligned}$$

where $\varsigma = 96\sigma^2\kappa_f(\hbar_f+1)\beta_\Theta\big(1+\frac{\hbar_f^2}{\Lambda_n}\big)\big)$ and the third last inequality comes from the fact that $\log(1/\delta) + d_{\mathcal{Y}}\log 6 + \log(n) \le 6\log(1/\delta)d_{\mathcal{Y}}\log(n)$ when $\delta \le 1/2$ and $n > 1$.

In the special scenario where the function class is linear or generalized linear, $f_\theta(\zeta, x) = \theta^\top(\zeta, x)$. Then, $Z_n(\theta) = \sum_{i=1}^n (\zeta_i, X_i)$ and $V_n(\theta) = \sum_{i=1}^n (\zeta_i, X_i)(\zeta_i, X_i)^\top$, so $H(\theta_1) = H(\theta_2)$ for any θ_1 and θ_2. Thus, we immediately conclude from Lemma 1(I) that

$$\begin{aligned}&\mathbb{P}\big(\|\widehat{Z}_n\|^2_{\widehat{V}_n^{-1}} \geq dc^2\big)\\ =&\mathbb{P}\big\{\|Z_n(\theta)\|^2_{V_n(\theta)^{-1}} \geq dc^2\big\}\\ =&\mathbb{P}\big\{\|Z_n(\theta)\|^2_{V_n(\theta)^{-1}} \geq 16d\eta^2\sigma^2\log(n)\log(d/\delta)\big\} \leq \delta.\end{aligned}$$

Define $\mathfrak{C}_{\mathcal{F}} = \log(d/\delta)$ when the function class is linear or generalized linear. For the general function class,

$$\mathfrak{C}_{\mathcal{F}} = 4d\log(n^3\log(n)d^3 d_{\mathcal{Y}}^2\log(1/\delta)\varsigma/(\delta\Lambda_n)).$$

By combining the above analysis, we can conclude that

$$\mathbb{P}\big(\|\widehat{Z}_n\|^2_{\widehat{V}_n^{-1}} \geq 16d\eta^2\sigma^2\log(n)\mathfrak{C}_{\mathcal{F}}\big) \leq 3\delta.$$

□

6.3 Verification of Assumption 3

Lemma 4 *Consider the statistical loss function ℓ to be the square loss, the negative log-likelihood or the cross entropy. If for every $\zeta \in \Omega$, $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, ℓ satisfies the restricted secant inequality at θ^*, then Assumption 3 holds.*

Proof Let us first verify that

$$\nabla\ell(y, \phi(f_\theta(\zeta, x))) = \nabla f_\theta(\zeta, x)^\top(\phi(f_\theta(\zeta, x)) - y).$$

For the square loss $\ell(y, \phi(f_\theta(\zeta, x))) = \frac{1}{2}\|y - \phi(f_\theta(\zeta, x))\|_2^2$ with ϕ being the identity map, we have $\nabla\ell(y, \phi(f_\theta(\zeta, x))) = \nabla f_\theta(\zeta, x)^\top(\phi(f_\theta(\zeta, x)) - y)$.

For the log-likelihood, when $d_{\mathcal{Y}} = 1$, $\ell(y, \phi(f_\theta(\zeta, x))) = -y\log(\phi(f_\theta(\zeta, x))) - (1-y)\log(1-\phi(f_\theta(\zeta, x)))$, where $\phi(y) = \frac{\exp(y)}{1+\exp(y)}$, we have

$$\nabla\ell(y, \phi(f_\theta(\zeta, x))) = -\frac{y\phi'\nabla f_\theta(\zeta, x)}{\phi(f_\theta(\zeta, x))} + \frac{(1-y)\phi'\nabla f_\theta(\zeta, x)}{1-\phi(f_\theta(\zeta, x))}.$$

Since

$$-\frac{y\phi'}{\phi} + \frac{(1-y)\phi'}{1-\phi} = \frac{-y\phi(1-\phi)}{\phi} + \frac{(1-y)\phi(1-\phi)}{1-\phi} = -y + \phi(f_\theta(\zeta, x)),$$

we calculate that $\nabla\ell(y, f_\theta(\zeta, x)) = \nabla f_\theta(\zeta, x)^\top(\phi(f_\theta(\zeta, x)) - y)$.

On the other hand, when $d_{\mathcal{Y}} \geq 2$, $\ell(y, \phi(f_\theta)) = \sum_{i=1}^{d_{\mathcal{Y}}} -y_i\log(\phi(f_\theta(\zeta, x)_i))$, where $\phi(y)_i = \frac{\exp(y_i)}{1+\sum_{i=1}^{d_{\mathcal{Y}}}\exp(y_i)}$. Let $z_\theta := f_\theta(\zeta, x)$. Thus,

$$\frac{\partial\ell(y, \phi(f_\theta))}{\partial z_{\theta,k}} = \frac{\partial}{\partial z_{\theta,k}}\sum_{i=1}^{d_{\mathcal{Y}}} -y_i\log(\phi(f_\theta(\zeta, x)_i)) = -y_k + \phi(f_\theta(\zeta, x)_k).$$

Then, it follows from the chain rule that

$$\nabla\ell(y, \phi(f_\theta(\zeta, x))) = \nabla f_\theta(\zeta, x)^\top(\phi(f_\theta(\zeta, x)) - y).$$

When ℓ satisfies the restricted secant inequality with parameter $\tilde{\alpha}$, we have

$$\begin{aligned}
&\left(\nabla f_\theta(\zeta,x)^\top\left(\phi(f_\theta(\zeta,x))-\phi(f_{\theta^*}(\zeta,x))\right)\right)^\top(\theta-\theta^*)\\
=&\nabla\ell(y,\phi(f_\theta))^\top(\theta-\theta^*)\mid_{y=\phi(f^*_\theta(\zeta,x))}\\
\geq&\tilde{\alpha}\|\theta-\theta^*\|_2^2\\
\geq&\tilde{\alpha}/\kappa_f^2(\theta-\theta^*)^\top\nabla f_\theta(\zeta,x)^\top\nabla f_\theta(\zeta,x)(\theta-\theta^*).
\end{aligned}$$

The proof is finalized by setting $\alpha=\tilde{\alpha}/\kappa_f^2$. □

6.4 Proof of Theorem 1

Define $\Delta_n=\widehat{\theta}_n-\theta^*$. It suffices to show that $\|\widehat{Z}_n\|^2_{\widehat{V}_n^{-1}}\geq C\|\Delta_n\|^2_{\widehat{V}_n}$ for some properly chosen constant C.

Using the mean value theorem, for $i=1,\cdots,n$, and $k=1,\cdots,d_{\mathcal{Y}}$, there exists $\bar{\theta}_{i,k}$ which is a convex combination of $\widehat{\theta}_n$ and θ^* such that

$$\phi(f_{k,\widehat{\theta}_n}(\zeta_i,X_i))-\phi(f_{k,\theta^*}(\zeta_i,X_i))=g_{k,\bar{\theta}_{i,k}}(\zeta_i,X_i)(\widehat{\theta}_n-\theta^*),$$

where g is defined in (7). For each $i=1,\cdots,d_{\mathcal{Y}}$, a compact form is given by

$$\phi(f_{\widehat{\theta}_n}(\zeta_i,X_i))-\phi(f_{\theta^*}(\zeta_i,X_i))=g_{\bar{\theta}_i}(\zeta_i,X_i)(\widehat{\theta}_n-\theta^*)$$

where $g_{\bar{\theta}_i}(\zeta_i,X_i):=\begin{pmatrix}g_{1,\bar{\theta}_{i,1}}(\zeta_i,X_i)\\ \vdots\\ g_{d_{\mathcal{Y}},\bar{\theta}_{i,d_{\mathcal{Y}}}}(\zeta_i,X_i)\end{pmatrix}$.

Using the definition of $\widehat{Z}_n$ and the (6) that $\widehat{\theta}_n$ satisfies, it follows that

$$\begin{aligned}
\widehat{Z}_n&=\sum_{i=1}^n\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top\epsilon_i\\
&=\sum_{i=1}^n\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top(Y_i-\phi(f_{\theta^*}(\zeta_i,X_i)))\\
&=\sum_{i=1}^n\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top(\phi(f_{\widehat{\theta}_n}(\zeta_i,X_i))-\phi(f_{\theta^*}(\zeta_i,X_i)))\\
&=\sum_{i=1}^n\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top g_{\bar{\theta}_i}(\zeta_i,X_i)(\widehat{\theta}_n-\theta^*).
\end{aligned}$$

Define a matrix B as

$$B:=\sum_{i=1}^n g_{\bar{\theta}_i}(\zeta_i,X_i)^\top\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)-\sum_{i=1}^n\underline{\kappa}_\phi\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i).$$

It follows that

$$\begin{aligned}
\|\widehat{Z}_n\|^2_{\widehat{V}_n^{-1}}&=\Delta_n^\top\left(\underline{\kappa}_\phi\sum_{i=1}^n\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)+B\right)\widehat{V}_n^{-1}\cdot\\
&\quad\left(\underline{\kappa}_\phi\sum_{i=1}^n\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)^\top\nabla f_{\widehat{\theta}_n}(\zeta_i,X_i)+B^\top\right)\Delta_n\\
&=\underline{\kappa}_\phi^2\|\Delta_n\|^2_{\widehat{V}_n}+2\underline{\kappa}_\phi\Delta_n^\top B\Delta_n+\Delta_n^\top B\widehat{V}_n^{-1}B^\top\Delta_n.
\end{aligned}\tag{11}$$

If $\Delta_n^\top B\Delta_n \geq 0$, since $\Delta_n^\top B\widehat{V}_n^{-1}B^\top\Delta_n \geq 0$, we conclude that

$$\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}}^2 \geq \underline{\kappa}_\phi^2\|\Delta_n\|_{\widehat{V}_n}^2,$$

thus we reach the conclusion.

Otherwise if $\Delta_n^\top B\Delta_n < 0$, using Assumption 3 we have

$$\begin{aligned}&\left(\nabla f_{\widehat{\theta}_n}(\zeta_i, X_i)^\top(\phi(f_{\widehat{\theta}_n}(\zeta_i, X_i)) - \phi(f_{\theta^*}(\zeta_i, X_i)))\right)^\top \Delta_n\\ \geq&\alpha\Delta_n^\top\nabla f_\theta(w,x)^\top\nabla f_\theta(w,x)\Delta_n.\end{aligned}$$

Summing s from 1 to n yields

$$\begin{aligned}&\sum_{i=1}^n\left(\nabla f_{\widehat{\theta}_n}(\zeta_i, X_i)^\top(\phi(f_{\widehat{\theta}_n}(\zeta_i, X_i)) - \phi(f_{\theta^*}(\zeta_i, X_i)))\right)^\top \Delta_n\\ &\geq \alpha\sum_{i=1}^n\Delta_n^\top\nabla f_\theta(\zeta_i, X_i)^\top\nabla f_\theta(\zeta_i, X_i)\Delta_n.\end{aligned}$$

It follows that

$$\begin{aligned}\Delta_n^\top B\Delta_n &= \sum_{i=1}^n(\phi(f_{\widehat{\theta}_n}(\zeta_i, X_i)) - \phi(f_{\theta^*}(\zeta_i, X_i)))^\top\nabla f_{\widehat{\theta}_n}\Delta_n - \underline{\kappa}_\phi\|\Delta_n\|_{\widehat{V}_n}^2\\ &\geq (\alpha - \underline{\kappa}_\phi)\|\Delta_n\|_{\widehat{V}_n}^2.\end{aligned}$$

By Cauchy-Schwarz inequality, we have that

$$(\Delta_n^\top B\widehat{V}_n^{-1}B^\top\Delta_n)^{\frac{1}{2}} = \|B\Delta_n\|_{\widehat{V}_n^{-1}} \geq \frac{|\Delta_n^\top B\Delta_n|}{\|\Delta_n\|_{\widehat{V}_n}} \geq (\alpha - \underline{\kappa}_\phi)\|\Delta_n\|_{\widehat{V}_n}.$$

Combined with (11), we conclude that

$$\begin{aligned}\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}}^2 &\geq \underline{\kappa}_\phi^2\|\Delta_n\|_{\widehat{V}_n}^2 + 2\underline{\kappa}_\phi\Delta_n^\top B\Delta_n + \frac{(\Delta_n^\top B\Delta_n)^2}{\|\Delta_n\|_{\widehat{V}_n}^2}\\ &= \left(\underline{\kappa}_\phi\|\Delta_n\|_{\widehat{V}_n} + \frac{\Delta_n^\top B\Delta_n}{\|\Delta_n\|_{\widehat{V}_n}}\right)^2 \geq \alpha^2\|\Delta_n\|_{\widehat{V}_n}^2.\end{aligned}$$

Thus by Lemma 1, it holds with probability at least $1-3\delta$ that

$$\Lambda_n\|\Delta_n\|_2^2 \leq \|\Delta_n\|_{\widehat{V}_n}^2 \leq \frac{1}{\min(\underline{\kappa}_\phi^2, \alpha^2)}\|\widehat{Z}_n\|_{\widehat{V}_n^{-1}}^2 \leq \frac{16d\eta^2\sigma^2\log(n)\mathfrak{C}_\mathcal{F}}{\min(\underline{\kappa}_\phi^2, \alpha^2)}.$$

Hence the proof is completed. □

7 Conclusion

In this paper, we establish non-asymptotic performance guarantees of PTO and ETO approaches for contextual stochastic optimization with endogenous uncertainty, providing bounds on approximation error and regret. Several potential extensions to our work are worth investigating. While we develop statistical inference for a general parametric function class, exploring how our framework could extend to nonparametric function classes presents an interesting direction for future research. Moreover, evaluating the performance of other frameworks such as ILO and decision-rule optimization under endogenous uncertainty are left for future work.

References

1. Abbasi-Yadkori Y, Pál D, Szepesvári C (2011) Online least squares estimation with self-normalized processes: An application to bandit problems. arXiv preprint arXiv:11022670
2. Ban GY, Rudin C (2019) The big data newsvendor: Practical insights from machine learning. Operations Research 67(1):90–108
3. Bazier-Matte T, Delage E (2020) Generalization bounds for regularized portfolio selection with market side information. INFOR: Information Systems and Operational Research 58(2):374–401
4. Bertsimas D, Kallus N (2020) From predictive to prescriptive analytics. Management Science 66(3):1025–1044
5. Bertsimas D, Koduri N (2022) Data-driven optimization: A reproducing kernel hilbert space approach. Operations Research 70(1):454–471
6. Bertsimas D, Van Parys B (2022) Bootstrap robust prescriptive analytics. Mathematical Programming 195(1):39–78
7. Boucheron S, Lugosi G, Massart P (2013) Concentration Inequalities: A Nonasymptotic Theory of Independence. Oxford University Press
8. Chen N, Lagzi S, Milner J (2022) Using neural networks to guide data-driven operational decisions. Available at SSRN 4217092
9. Chen R, Paschalidis I (2019) Selecting optimal decisions via distributionally robust nearest-neighbor regression. Advances in Neural Information Processing Systems 32
10. Chu W, Li L, Reyzin L, Schapire R (2011) Contextual bandits with linear payoff functions. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp 208–214
11. Dani V, Hayes TP, Kakade SM (2008) Stochastic linear optimization under bandit feedback. Conference on Learning Theory
12. El Balghiti O, Elmachtoub AN, Grigas P, Tewari A (2019) Generalization bounds in the predict-then-optimize framework. Advances in neural information processing systems 32
13. El Balghiti O, Elmachtoub AN, Grigas P, Tewari A (2023) Generalization bounds in the predict-then-optimize framework. Mathematics of Operations Research 48(4):2043–2065
14. Elmachtoub AN, Grigas P (2022) Smart "predict, then optimize". Management Science 68(1):9–26
15. Elmachtoub AN, Lam H, Zhang H, Zhao Y (2023) Estimate-then-optimize versus integrated-estimation-optimization: A stochastic dominance perspective. arXiv preprint arXiv:230406833
16. Feng Q, Shanthikumar JG, Wu J (2023) Contextual data-integrated newsvendor solution with operational data analytics (ODA). Available at SSRN 4668658
17. Filippi S, Cappe O, Garivier A, Szepesvári C (2010) Parametric bandits: The generalized linear case. In: Advances in Neural Information Processing Systems, pp 586–594
18. Gupta V, Rusmevichientong P (2021) Small-data, large-scale linear optimization with uncertain objectives. Management Science 67(1):220–241
19. Gupta V, Huang M, Rusmevichientong P (2024) Debiasing in-sample policy performance for small-data, large-scale optimization. Operations Research 72(2):848–870
20. Hellemo L, Barton PI, Tomasgard A (2018) Decision-dependent probabilities in stochastic programs with recourse. Computational Management Science 15(3):369–395
21. Ho-Nguyen N, Kılınç-Karzan F (2022) Risk guarantees for end-to-end prediction and optimization processes. Management Science 68(12):8680–8698
22. Hu Y, Kallus N, Mao X (2022) Fast rates for contextual linear optimization. Management Science 68(6):4236–4245
23. Kannan R, Bayraksan G, Luedtke JR (2020) Data-driven sample average approximation with covariate information. Optimization Online
24. Kannan R, Bayraksan G, Luedtke J (2021) Heteroscedasticity-aware residuals-based contextual stochastic optimization. arXiv preprint arXiv:210103139
25. Kannan R, Bayraksan G, Luedtke JR (2023) Residuals-based distributionally robust optimization with covariate information. Mathematical Programming pp 1–57
26. Karimi H, Nutini J, Schmidt M (2016) Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, pp 795–811
27. Kveton B, Zaheer M, Szepesvari C, Li L, Ghavamzadeh M, Boutilier C (2020) Randomized exploration in generalized linear bandits. In: International Conference on Artificial Intelligence and Statistics, pp 2066–2076

28. Lattimore T, Hao B (2021) Bandit phase retrieval. arXiv preprint arXiv:210601660
29. Lattimore T, Szepesvári C (2020) Bandit algorithms. Cambridge University Press
30. Li L, Chu W, Langford J, Schapire RE (2010) A contextual-bandit approach to personalized news article recommendation. In: Proceedings of the 19th international conference on World wide web, pp 661–670
31. Li L, Lu Y, Zhou D (2017) Provably optimal algorithms for generalized linear contextual bandits. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR.org, pp 2071–2080
32. Liu H, Grigas P (2021) Risk bounds and calibration for a smart predict-then-optimize method. Advances in Neural Information Processing Systems 34:22083–22094
33. Loke GG, Tang Q, Xiao Y (2022) Decision-driven regularization: A blended model for predict-then-optimize. Available at SSRN 3623006
34. Luo F, Mehrotra S (2020) Distributionally robust optimization with decision dependent ambiguity sets. Optimization Letters 14(8):2565–2594
35. Nohadani O, Sharma K (2018) Optimization under decision-dependent uncertainty. SIAM Journal on Optimization 28(2):1773–1795
36. Noyan N, Rudolf G, Lejeune M (2022) Distributionally robust optimization under a decision-dependent ambiguity set with applications to machine scheduling and humanitarian logistics. INFORMS Journal on Computing 34(2):729–751
37. de la Pena VH, Klass MJ, Lai TL (2004) Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. Annals of probability pp 1902–1933
38. Peña VH, Lai TL, Shao QM (2008) Self-normalized processes: Limit theory and Statistical Applications. Springer Science & Business Media
39. Petersen KB, Pedersen MS, et al. (2008) The matrix cookbook. Technical University of Denmark 7(15):510
40. Qi M, Shen ZJ (2022) Integrating prediction/estimation and optimization with applications in operations management. In: Tutorials in Operations Research: Emerging and Impactful Topics in Operations, INFORMS, pp 36–58
41. Qi M, Grigas P, Shen ZJM (2021) Integrated conditional estimation-optimization. arXiv preprint arXiv:211012351
42. Rusmevichientong P, Tsitsiklis JN (2010) Linearly parameterized bandits. Mathematics of Operations Research 35(2):395–411
43. Sadana U, Chenreddy A, Delage E, Forel A, Frejinger E, Vidal T (2024) A survey of contextual optimization methods for decision-making under uncertainty. European Journal of Operational Research
44. Shapiro A, Dentcheva D, Ruszczynski A (2021) Lectures on stochastic programming: modeling and theory. SIAM
45. Srivastava PR, Wang Y, Hanasusanto GA, Ho CP (2021) On data-driven prescriptive analytics with side information: A regularized nadaraya-watson approach. arXiv preprint arXiv:211004855
46. Sutter T, Van Parys BP, Kuhn D (2024) A pareto dominance principle for data-driven optimization. Operations Research
47. Valko M, Korda N, Munos R, Flaounas I, Cristianini N (2013) Finite-time analysis of kernelised contextual bandits. arXiv preprint arXiv:13096869
48. Van Parys BP, Esfahani PM, Kuhn D (2021) From data to decisions: Distributionally robust optimization is optimal. Management Science 67(6):3387–3402
49. Zahavy T, Mannor S (2019) Deep neural linear bandits: Overcoming catastrophic forgetting through likelihood matching. arXiv preprint arXiv:190108612
50. Zhang H, Yin W (2013) Gradient methods for convex minimization: better rates under weaker conditions. arXiv preprint arXiv:13034645
51. Zhang L, Yang J, Gao R (2024) Optimal robust policy for feature-based newsvendor. Management Science 70(4):2315–2329
52. Zhang Y, Liu J, Zhao X (2023) Data-driven piecewise affine decision rules for stochastic programming with covariate information. arXiv preprint arXiv:230413646
53. Zhou D, Li L, Gu Q (2020) Neural contextual bandits with ucb-based exploration. In: International Conference on Machine Learning, PMLR, pp 11492–11502
54. Zhu T, Xie J, Sim M (2022) Joint estimation and robustness optimization. Management Science 68(3):1659–1677