A Sharp Comparison of Prescriptive Analytic Frameworks for The Big Data Newsvendor Problem

Zhen Qiao
Engineering Systems and Design Pillar,
Singapore University of Technology and Design,
Singapore
zhen_qiao@mymail.sutd.edu.sg

Karthyek Murthy
Daniel J. Epstein Department of Industrial & Systems Engineering,
University of Southern California,
Los Angeles, USA
karthyek@usc.edu

Abstract

We study the feature-based newsvendor problem in high dimensions and provide sharp regret characterizations for three widely used prescriptive analytics approaches: the conventional Estimate-Then-Optimize (ETO), and the more recent end-to-end methods of Integrated-Learning-Optimization (ILO) & Direct Policy Optimization (DPO). Under well-specified linear demand model, we derive the regret results using convex Gaussian minmax theorem. Numerical explorations enabled by the results offer robust evidence of the superiority of ETO persisting even in high-dimensional regimes. Further, we highlight substantial performance gains attainable over all 3 methods by utilizing downstream optimization only in the model selection stage, instead of embedding it in training as with ILO & DPO.

1 Introduction

Recent advances in end-to-end learning of prediction models tailored for downstream optimization [1, 2, 3, 4] present modern operations managers with a critical design choice: Should estimation and optimization be treated as separate stages, or integrated into a single, end-to-end pipeline? The central premise of the latter is that being optimization task-aware enables learning that is directly aligned with decision costs and quality. But does this promise hold in out-of-sample deployment? In particular, does the in-sample optimality achieved by integrated methods consistently translate into improved out-of-sample performance? The anticipation of such gains has spurred a surge of decision-focused learning approaches in recent years; see [5] and references therein for a comprehensive overview of integrating learning and optimization in data-driven contextual optimization.

Out-of-sample generalization bounds for integrated-learning-optimization (ILO) approaches (see, eg., [2, 6, 7, 8]), while useful for gaining insights into specific learning algorithms, tend to fall short of revealing the benefits of ILO, if any, over the conventional Estimate-Then-Optimize (ETO) approaches. For instance, in predicting demand for a newsvendor problem, is there a compelling reason to favor ILO over established procedures such as least-squares or Lasso regression? Under model misspecification, recent efforts [9, 10] have shown that ILO can achieve smaller regret in contextual optimization and deliver additional first-order improvement in non-contextual settings.

By contrast, when the learning model is well-specified, theoretical evidence thus far points to stronger performance guarantees for ETO, including faster regret convergence in contextual linear optimization [11] and stochastic dominance for regret in broader data-driven optimization [9]. It should be noted that the latter results are obtained in fixed-dimensional large-sample regime, which typically asserts the optimality of unregularized ETO (eg., Cramer-Rao lower bound, see also [12]). Such optimality observations drawn from fixed dimensional analysis may appear at odds with the

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

widespread need for model selection tools such as regularization or dropout-type techniques that have proven essential in high-dimensional learning contexts.

Does ETO continue to dominate ILO even in high dimensional settings, where the first-step estimation needs to learn a prediction model from a high-dimensional feature representation? This paper is devoted to examining this question in the well-studied feature-based newsvendor model treated in [2]. Using the Convex Gaussian MinMax Theorem (CGMT) [13], we derive sharp regret characterizations in the regime where the ambient dimension p is proportionately large as the sample size n under a well-specified model. Unlike in finite-dimensional regimes, the resulting regret does not vanish asymptotically, underscoring the intrinsic difficulty of this setting. Numerical explorations enabled by the regret characterizations reveal the superiority of ETO over ILO & DPO. The end-to-end coupling in ILO and DPO distort the KKT optimality conditions, inflating both prediction error and prescriptive costs. Furthermore, they highlight substantial performance gains attainable over ETO, ILO & DPO when the downstream optimization task is utilized only while cross-validating for model selection, instead of embedding it directly in training as with ILO & DPO.

2 The Feature-based Newsvendor Model & Estimation approaches

We consider the feature-based newsvendor problem treated in [2]: A decision-maker, responsible for making purchase orders for a product sold by their firm, is faced with uncertain demand D for the product. While determining the number of units to be ordered, the decision-maker has the flexibility to leverage observable contextual information $\boldsymbol{X} \in \mathbb{R}^p$ (such as weather, spatio-temporal attributes, economic indicators, etc.) which may have a role in influencing/explaining the demand D. The decision-maker is also equipped with the dataset $\mathcal{D} = \{(\boldsymbol{X}_1, D_1), \ldots, (\boldsymbol{X}_n, D_n)\}$ comprising n independent historical observations of the demand paired with the respective contextual feature information. Upon ordering w units of the product, the decision-maker pays either a holding cost of $h(w-D)^+$ to store the excess $(w-D)^+$, or, a back-order cost of $b(D-w)^+$ as a penalty for the shortfall $(D-w)^+$, depending on the realized value of the demand D. Here $h, b > 0$.

If the decision-maker is risk-neutral, it would be ideal if they could identify an ordering rule $w : \mathbb{R}^p \to \mathbb{R}$ that minimizes the conditional expected cost $E[L_{nv}(w(\boldsymbol{X}) - D) \mid \boldsymbol{X}]$ almost surely, where

$$L_{nv}(w-d) = h(w-d)^+ + b(d-w)^+ \tag{1}$$

is the combined holding and back-order costs incurred by ordering w units of the product. Identifying such conditionally optimal order quantity is an impossible proposition though, given that the decision-maker has access to the distribution of $D \mid \boldsymbol{X}$ only via the limited dataset $\mathcal{D}$. They may nevertheless utilize the dataset $\mathcal{D}$ to either (a) learn how the distribution of the demand D is determined by the contextual information $\boldsymbol{X}$ and use this subsequently to determine the order quantity, or, (b) rather directly learn an ordering rule which is optimal within a class of pre-specified ordering rules. Assuming a linear demand model

$$D = \theta_0^* + \boldsymbol{\theta}^{*\top} \boldsymbol{X} + \varepsilon, \tag{2}$$

where $\theta_0^* > 0, \boldsymbol{\theta}^* \in \mathbb{R}^p$, and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, this paper is devoted to sharp performance analyses of the following prominent contextual optimization approaches aligning between either possibilities.

Estimate-Then-Optimize (ETO): As the name suggests, this approach focuses on first estimating the unknown model parameters $\theta_0^*, \boldsymbol{\theta}^*$, and σ via a conventional regression procedure such as

$$(\hat{\theta}_0, \hat{\boldsymbol{\theta}}) \in \arg\min_{\theta_0, \boldsymbol{\theta}} \frac{1}{n} \sum_{i=1}^n \left(D_i - \theta_0 - \boldsymbol{\theta}^\top \boldsymbol{X}_i\right)^2 \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^n \left(D_i - \hat{\theta}_0 - \hat{\boldsymbol{\theta}}^\top \boldsymbol{X}_i\right)^2,$$

that focuses primarily on prediction error. One may alternatively consider a regularized variant as in

$$(\hat{\theta}_0^{(\eta)}, \hat{\boldsymbol{\theta}}^{(\eta)}) \in \arg\min_{\theta_0, \boldsymbol{\theta}} \frac{1}{n} \sum_{i=1}^n \left(D_i - \theta_0 - \boldsymbol{\theta}^\top \boldsymbol{X}_i\right)^2 + \eta \|\boldsymbol{\theta}\|_2^2, \quad \text{where } \eta > 0. \tag{3}$$

The distinguishing feature of this conventional ETO approach is that model estimation is undertaken entirely *agnostic* to the newsvendor's downstream cost (1). Supposing the estimated model parameters $(\hat{\theta}_0, \hat{\boldsymbol{\theta}}, \hat{\sigma})$ to be well-describing the ground truth, ETO then prescribes the decision rule,

$$w(\boldsymbol{X}) = \hat{F}^{-1}_{D|\boldsymbol{X}}\left(\frac{b}{b+h}\right) = \hat{\theta}_0 + \hat{\boldsymbol{\theta}}^\top \boldsymbol{X} + \hat{\sigma} \Phi^{-1}\left(\frac{b}{b+h}\right), \tag{4}$$

as it minimizes the resulting estimate of $E[L_{nv}(w(\boldsymbol{X}) - D) \mid \boldsymbol{X}]$. Here, $\hat{F}_{D|\boldsymbol{X}}(\cdot)$ denotes the CDF of the estimated conditional distribution $D \mid \boldsymbol{X} \sim \mathcal{N}(\hat{\theta}_0 + \hat{\boldsymbol{\theta}}^\top \boldsymbol{X}, \hat{\sigma}^2)$ and $\Phi^{-1}(p) = \inf\{z \in \mathbb{R} : P(Z \leq z) \geq p\}$ is the inverse CDF of the standard normal distribution.

Integrated-Learning-Optimization (ILO): Observe that the ordering rule $w(\boldsymbol{X}) = \theta_0 + \boldsymbol{\theta}^\top \boldsymbol{X} + s\Phi^{-1}(b/(b+h))$ is optimal for the demand model $D \mid \boldsymbol{X} \sim \mathcal{N}(\theta_0 + \boldsymbol{\theta}^\top \boldsymbol{X}, s^2)$. Hence, the task of finding model parameters which minimize the sample average newsvendor cost associated with the respective optimal ordering rule can be accomplished by,

$$\left(\hat{\theta}_{0,\mathtt{ILO}}, \hat{\boldsymbol{\theta}}_{\mathtt{ILO}}, \hat{\sigma}_{\mathtt{ILO}}\right) \in \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^p, \theta_0, s \geq 0} \frac{1}{n} \sum_{i=1}^{n} L_{nv}\left(\theta_0 + \boldsymbol{\theta}^\top \boldsymbol{X}_i + s\Phi^{-1}(b/(b+h)) - D_i\right). \quad (5)$$

Here, unlike ETO, the model estimation is directly informed by the optimization objective (1). By definition of this integrated approach, the in-sample newsvendor cost associated with the estimated model parameters $(\hat{\theta}_{0,\mathtt{ILO}}, \hat{\boldsymbol{\theta}}_{\mathtt{ILO}}, \hat{\sigma}_{\mathtt{ILO}})$ is smaller than that incurred by deploying the ETO optimal ordering rule (4). Refer [3, 14, 9, 10] and references therein for a comprehensive account of ILO.

Decision-Rule Optimization (or) Direct Policy Optimization (DPO): Considering linear decision rules of the form $w(\boldsymbol{X}) = a + \boldsymbol{b}^\top \boldsymbol{X}$, this approach bypasses the model estimation step. It seeks to directly find optimal decision rule coefficients minimizing the empirical newsvendor loss as in,

$$(\hat{a}_{\mathtt{DPO}}, \hat{\boldsymbol{b}}_{\mathtt{DPO}}) \in \arg\min_{a, \boldsymbol{b}} \frac{1}{n} \sum_{i=1}^{n} L_{nv}\big(a + \boldsymbol{b}^\top \boldsymbol{X}_i - D_i\big).$$

Refer [2] for an account of DPO in the newsvendor model and [4, 1] for their broader use.

3 Sharp Characterizations for Out-of-Sample Costs in High-Dimensions

Does being optimization task-aware, as in the end-to-end learning adopted in ILO & DPO, translate into gains in out-of-sample performance? In particular, does the in-sample optimality of ILO manifest as an improvement in out-of-sample performance as well? In order to answer these questions in an asymptotic regime representative of the challenges posed by modern datasets, we consider the high-dimensional setting where the the dimension of the feature vector $\boldsymbol{X}$, denoted by p, is comparable with the number of samples n. In particular, we consider $n \to \infty, n/p \to \delta \in (1, \infty)$, a setting canonical in high-dimensional statistics, but novel for analysis in contextual optimization literature.

Assumption 1. The p-dimensional feature vector $\boldsymbol{X}$ is standardized to satisfy $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, \frac{1}{p}\mathbf{I}_p)$.

Extensions to the below regret characterizations, which are based on convex Gaussian Min-max Theorem (CGMT, [13, Thm. 3]), to suit $\boldsymbol{X}$ featuring elliptical distributions and more general random feature model are conceivable along the lines pursued in [15, 16].

We define the out-of-sample *regret* of a decision rule $w(\cdot)$ as $E[L_{nv}(w(\boldsymbol{X}) - D)] - E[L_{nv}(w^*(\boldsymbol{X}) - D)]$, capturing the out-of-sample cost incurred by deploying $w(\cdot)$ when compared to the oracle optimal decision rule $w^*(\boldsymbol{X}) = \theta_0^* + \boldsymbol{\theta}^{*\top} \boldsymbol{X} + \sigma\Phi^{-1}(b/(b+h))$. To state our results, define

$$\ell(\mu, s) := E\left[L_{nv}\left(\mu + sZ\right)\right] = (b+h)\left[s\,\phi\left(\frac{\mu}{s}\right) + \mu\Phi\left(\frac{\mu}{s}\right)\right] - \mu b,$$

where $Z \sim \mathcal{N}(0,1)$ and $\kappa := \Phi^{-1}(b/(b+h))$. Additionally, let

$$\psi(z, \lambda) := \max_{u \in [-b, h]}\left\{zu - \frac{\lambda u^2}{2}\right\} - \frac{s^2 - \sigma^2}{2\lambda\delta}.$$

Theorem 1. *For the demand model (2) satisfying Assump. 1, the out-of-sample regret $\mathcal{R}_n$ satisfies*

$$\boxed{\mathcal{R}_n \to \ell\big(\mu^*, s^*\big) - (b+h)\sigma\phi(\kappa)}, \qquad \textit{as } n \to \infty, \textit{ where}$$

$$\begin{cases} \textit{under the ETO-optimal ordering rule (4):} \quad \mu^* = \sigma\kappa,\ s^* = \sigma(1 - 1/\delta)^{-1/2}, \textit{ and} \\ \textit{under ILO \& DPO-optimal ordering rules: } \mu^*,\ s^* \textit{ solve } \min_{\mu \in \mathbb{R}, s \geq \sigma} \max_{\lambda > 0} E\left[\psi\left(\mu + sZ; \lambda\right)\right]. \end{cases}$$

Observe from Theorem 1 that the regret does not vanish, primarily due to the difficulty of learning when the number of samples are roughly only as much as the ambient dimensions. Contrast this with the finite dimensional setting where $O_p(1/n)$ regret convergence is obtainable: see, eg., [9]. As ILO & DPO's regret limits coincide, we shall only compare ETO vs ILO from here onwards.

Lemma 2. *Any solution to the lower dimensional problem* $\min_{\mu \in \mathbb{R}, s \geq \sigma} \max_{\lambda > 0} E\left[\psi\left(\mu + sZ; \lambda\right)\right]$ *satisfies the deterministic equivalent (DE) optimality conditions* (6a) - (6c) *in Appendix* A.

4 Discriminating ETO vs ILO numerically via DE Optimality Conditions

Considering various choices for δ and the cost parameters b, h, we numerically demonstrate in Figure 1 the remarkable accuracy of the limiting regret predictions obtained in Theorem 1 by comparing with their finite sample counterparts.

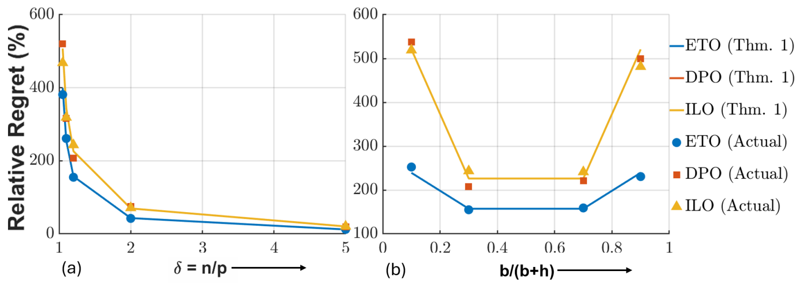

Figure 1: Finite-Sample performances (vs) regret characterizations: $p = 1000, \sigma = 1$, with $b/(b + h) = 0.3$ fixed in the left panel and $\delta = 1.2$ in the right panel

Fig 2 compares the limiting regret of ETO & ILO in Theorem 1. Interestingly, the heatmap in Fig 2(a) asserts the uniform dominance of ETO over ILO across all parameter settings $n/p \in (1, 5)$ and $b/(b + h) \in (0, 1)$. The disadvantage for ILO is more pronounced when either the dataset is smaller or costs b, h are skewed. Fig 2(b) demonstrates (i) the disadvantage with ILO even when the noise variance σ is taken to be known in ILO, and (b) the applicability of regularized ETO when $n < p$. Fig 3 asserts the benefits of being optimization-aware in a subtler manner than ILO: While performing model selection via cross-validation, we choose the ridge regularization strength η that minimizes the respective optimal cost $E[L_{nv}(\hat{\theta}_0^{(\eta)} + \boldsymbol{X}^\top \hat{\boldsymbol{\theta}}^{(\eta)} + \hat{\sigma}\kappa - D)]$ evaluated over a test dataset. All the panels in Figure 3 illustrate the robust performance gains one may obtain across various signal-to-noise ratios and dataset sizes. The benefits are more pronounced in smaller datasets.

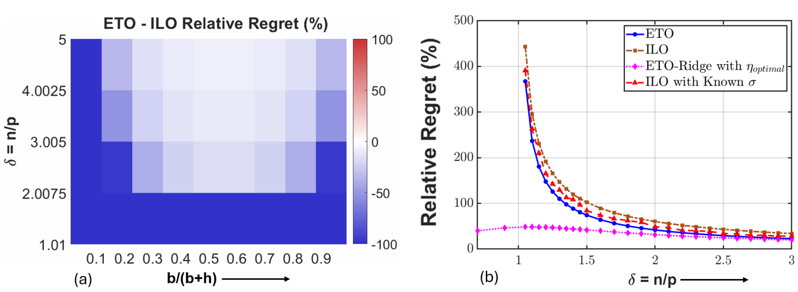

Figure 2: Comparing ETO vs ILO regret predictions from Theorem 1

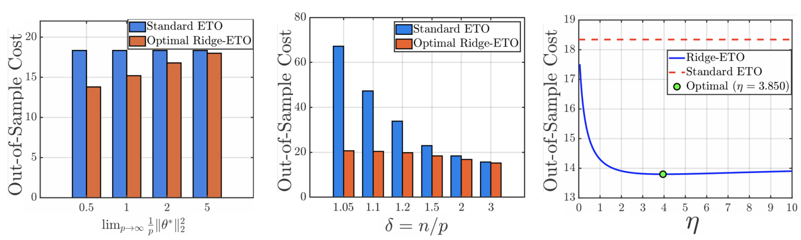

Figure 3: Illustration of lower out-of-sample costs with optimization-aware crossvalidation

References

[1] Priya Donti, Brandon Amos, and J Zico Kolter. Task-based end-to-end model learning in stochastic optimization. *Advances in neural information processing systems*, 30, 2017.

[2] Gah-Yi Ban and Cynthia Rudin. The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1):90–108, 2019.

[3] Adam N Elmachtoub and Paul Grigas. Smart predict, then optimize. *Management Science*, 68(1):9–26, 2022.

[4] Dimitris Bertsimas and Nathan Kallus. From predictive to prescriptive analytics. *Management Science*, 66(3):1025–1044, 2020.

[5] Utsav Sadana, Abhilash Chenreddy, Erick Delage, Alexandre Forel, Emma Frejinger, and Thibaut Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 2024.

[6] Othman El Balghiti, Adam N Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. *Mathematics of Operations Research*, 48(4):2043–2065, 2023.

[7] Meng Qi and Zuo-Jun Shen. Integrating prediction/estimation and optimization with applications in operations management. In *Tutorials in Operations Research: Emerging and Impactful Topics in Operations*, pages 36–58. INFORMS, 2022.

[8] Michael Huang and Vishal Gupta. Decision-focused learning with directional gradients. *arXiv preprint arXiv:2402.03256*, 2024.

[9] Adam N Elmachtoub, Henry Lam, Haofeng Zhang, and Yunfan Zhao. Estimate-then-optimize versus integrated-estimation-optimization: A stochastic dominance perspective. *arXiv preprint arXiv:2304.06833*, 2023.

[10] Adam N. Elmachtoub, Henry Lam, Haixiang Lan, and Haofeng Zhang. Dissecting the impact of model misspecification in data-driven optimization. In *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pages 1594–1602. PMLR, 03–05 May 2025.

[11] Yichun Hu, Nathan Kallus, and Xiaojie Mao. Fast rates for contextual linear optimization. *Management Science*, 2022.

[12] Henry Lam. On the impossibility of statistically improving empirical optimization: A second-order stochastic dominance perspective. *arXiv preprint arXiv:2105.13419*, 2021.

[13] Christos Thrampoulidis, Ehsan Abbasi, and Babak Hassibi. Precise error analysis of regularized m-estimators in high dimensions. *IEEE Transactions on Information Theory*, 64(8):5592–5628, 2018.

[14] Paul Grigas, Meng Qi, and Zuo-Jun Shen. Integrated conditional estimation-optimization. *arXiv preprint arXiv:2110.12351*, 2021.

[15] Noureddine El Karoui. On the impact of predictor geometry on the performance on high-dimensional ridge-regularized generalized robust regression estimators. *Probability Theory and Related Fields*, 170(12):95175, January 2017.

[16] David Bosch, Ashkan Panahi, Ayca Ozcelikkale, and Devdatt Dubhashi. Random features model with general convex regularization: A fine grained analysis with precise asymptotic learning curves. In Francisco Ruiz, Jennifer Dy, and Jan-Willem van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 11371–11414. PMLR, 25–27 Apr 2023.

A Deterministic Equivalent (DE) Optimality Conditions

The deterministic equivalent (DE) optimality conditions, indicated in Lemma 2 for the low-dimension optimization $\min_{\mu\in\mathbb{R},s\geq\sigma}\max_{\lambda>0}E\left[\psi\left(\mu+sZ;\lambda\right)\right]$, are given by the following non-linear equations in the variables $\mu\in\mathbb{R}$, $s\geq\sigma$, and $\lambda>0$:

$$[\lambda b+\mu]\,\Phi\left(-\frac{\lambda b+\mu}{s}\right)+[\lambda h-\mu]\,\Phi\left(\frac{\lambda h-\mu}{s}\right)+s\left[\phi\left(\frac{\lambda h-\mu}{s}\right)-\phi\left(\frac{\lambda b+\mu}{s}\right)\right]=\lambda h, \quad (6a)$$

$$\begin{aligned}&\left[\lambda^2b^2-\mu^2-s^2\right]\Phi\left(-\frac{\lambda b+\mu}{s}\right)+\left[\mu^2+s^2-\lambda^2h^2\right]\Phi\left(\frac{\lambda h-\mu}{s}\right)\\&\quad-\lambda^2 s\left[\mu+\lambda h\right]\phi\left(\frac{\lambda h-\mu}{s}\right)+\lambda^2 s\left[\mu-\lambda b\right]\phi\left(\frac{\lambda b+\mu}{s}\right)=\frac{s^2-\sigma^2}{\delta}-\lambda^2h^2, \quad \text{and} \quad (6b)\end{aligned}$$

$$s\left[\Phi\left(\frac{\lambda h-\mu}{s}\right)-\Phi\left(-\frac{\lambda b+\mu}{s}\right)\right]+(1-\lambda)\left[h\phi\left(\frac{\lambda h-\mu}{s}\right)+b\phi\left(\frac{\lambda b+\mu}{s}\right)\right]=\frac{s}{\lambda\delta}. \quad (6c)$$

B The Convex Gaussian Min-max Theorem (CGMT)

This section provides a brief overview of the specific tool from the Convex Gaussian Min-max Theorem (CGMT) framework [13] that is instrumental in our derivation of the ILO deterministic equivalent in Appendix D. Our goal is to analyze a high-dimensional optimization problem by relating it to a simpler, lower-dimensional "auxiliary" problem.

B.1 The Primary and Auxiliary Optimization Problems

The CGMT provides a powerful way to analyze saddle-point (min-max) problems that contain a specific type of bilinear term involving a large Gaussian matrix.

Let $\boldsymbol{A}\in\mathbb{R}^{n\times p}$ be a matrix with i.i.d. $\mathcal{N}(0,1)$ entries. Consider the following optimization problem, referred to as the **Primary Optimization (PO)** problem:

$$\Phi(\boldsymbol{A})=\min_{\boldsymbol{w}\in S_w}\max_{\boldsymbol{u}\in S_u}\left\{\boldsymbol{u}^\top\boldsymbol{A}\boldsymbol{w}+\psi(\boldsymbol{w},\boldsymbol{u})\right\},$$

where $S_w\subset\mathbb{R}^p$ and $S_u\subset\mathbb{R}^n$ are compact convex sets, and ψ is a continuous function that is convex in $\boldsymbol{w}$ and concave in $\boldsymbol{u}$.

The CGMT states that, under suitable conditions, the optimal value of the PO problem is asymptotically equivalent to the optimal value of a much simpler **Auxiliary Optimization (AO)** problem. The AO problem is defined by replacing the Gaussian bilinear term $\boldsymbol{u}^\top\boldsymbol{A}\boldsymbol{w}$ with terms involving the norms of the variables and two standard Gaussian vectors, $\boldsymbol{g}\in\mathbb{R}^n$ and $\boldsymbol{h}\in\mathbb{R}^p$:

$$\phi(\boldsymbol{g},\boldsymbol{h})=\min_{\boldsymbol{w}\in S_w}\max_{\boldsymbol{u}\in S_u}\left\{\|\boldsymbol{w}\|_2\,\boldsymbol{g}^\top\boldsymbol{u}+\|\boldsymbol{u}\|_2\,\boldsymbol{h}^\top\boldsymbol{w}+\psi(\boldsymbol{w},\boldsymbol{u})\right\}.$$

The key insight of the theorem is that the complex, high-dimensional interactions within the matrix $\boldsymbol{A}$ can be captured by the much simpler interactions in the AO problem.

B.2 Application to the ILO Problem in Appendix D

Our derivation of the ILO regret leverages this PO-AO correspondence directly. As shown in Equation (11) of Appendix D, the ILO empirical loss minimization can be formulated as a PO problem:

$$\mathcal{P}_n=\min_{\boldsymbol{d}_\theta\in\mathbb{R}^p,\ \mu\in\mathbb{R}}\ \max_{\boldsymbol{u}\in[-b,h]^n}\ \frac{1}{n}\Big(\boldsymbol{u}^\top\boldsymbol{X}\boldsymbol{d}_\theta+\mu\,\boldsymbol{u}^\top\boldsymbol{1}_n-\boldsymbol{u}^\top\boldsymbol{\varepsilon}\Big).$$

To match the CGMT framework, we perform the following identifications:

- The Gaussian matrix $\boldsymbol{A}$ corresponds to our rescaled feature matrix, $\sqrt{p}\,\boldsymbol{X}\in\mathbb{R}^{n\times p}$.
- The primal optimization variable $\boldsymbol{w}$ corresponds to the rescaled parameter error vector, $\boldsymbol{v}=\boldsymbol{d}_\theta/\sqrt{p}$.

- The dual optimization variable $\boldsymbol{u}$ is the vector of dual variables from the conjugate of the newsvendor loss, with $\boldsymbol{u} \in [-b, h]^n$.
- The function $\psi(\boldsymbol{w}, \boldsymbol{u})$ collects all remaining terms that do not involve the Gaussian matrix:

$$\psi(\boldsymbol{v}, \boldsymbol{u}; \mu) = \mu\, \boldsymbol{u}^\top \mathbf{1}_n - \boldsymbol{u}^\top \boldsymbol{\varepsilon}.$$

The application of the CGMT, as detailed in Appendix D, allows us to replace the challenging term $\boldsymbol{u}^\top(\sqrt{p}\,\boldsymbol{X})\boldsymbol{v}$ with its auxiliary counterpart. This step is crucial as it decouples the optimization variables, enabling the subsequent simplifications that lead to the final scalar deterministic equivalent.

C Proof for Theorem 1 (ETO Part)

C.1 Setup and Goal

We derive the asymptotic out-of-sample regret for the Estimate-Then-Optimize (ETO) policy under the setup of the main text. The demand model is $D = \theta_0^\star + \boldsymbol{\theta}^{\star\top}\boldsymbol{X} + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, and the feature vectors have i.i.d. components $X_{ij} \sim \mathcal{N}(0, 1/p)$. We operate in the high-dimensional asymptotic regime where $n, p \to \infty$ such that $n/p \to \delta \in (1, \infty)$.

The ETO policy, based on the Ordinary Least Squares (OLS) estimates $(\hat{\theta}_0, \hat{\boldsymbol{\theta}}, \hat{\sigma})$, is given by:

$$w_{\text{ETO}}(\boldsymbol{X}) = \hat{\theta}_0 + \hat{\boldsymbol{\theta}}^\top \boldsymbol{X} + \hat{\sigma}\,\kappa, \quad \text{where} \quad \kappa := \Phi^{-1}\left(\frac{b}{b+h}\right).$$

The oracle policy is $w^\star(\boldsymbol{X}) = \theta_0^\star + \boldsymbol{\theta}^{\star\top}\boldsymbol{X} + \sigma\,\kappa$. The out-of-sample regret is defined as $\mathcal{R}_n = E[L_{nv}(w_{\text{ETO}}(\boldsymbol{X}) - D)] - E[L_{nv}(w^\star(\boldsymbol{X}) - D)]$. Our goal is to derive the limit of $\mathcal{R}_n$.

C.2 Characterizing the Out-of-Sample Decision Error

The core of the regret calculation lies in characterizing the distribution of the out-of-sample decision error, $w_{\text{ETO}}(\boldsymbol{X}) - D$. Let $(\boldsymbol{X}, D)$ be a new test sample, independent of the training data used to obtain $(\hat{\theta}_0, \hat{\boldsymbol{\theta}}, \hat{\sigma})$. The error term can be decomposed as:

$$\begin{aligned}
w_{\text{ETO}}(\boldsymbol{X}) - D &= \left(\hat{\theta}_0 + \hat{\boldsymbol{\theta}}^\top \boldsymbol{X} + \hat{\sigma}\,\kappa\right) - \left(\theta_0^\star + \boldsymbol{\theta}^{\star\top}\boldsymbol{X} + \varepsilon\right) \\
&= (\hat{\theta}_0 - \theta_0^\star) + (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)^\top \boldsymbol{X} + (\hat{\sigma} - \sigma)\kappa + (\sigma\kappa - \varepsilon).
\end{aligned}$$

Conditional on the training data (i.e., for fixed $\hat{\boldsymbol{\theta}}, \hat{\sigma}, \hat{\theta}_0$), the term $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)^\top \boldsymbol{X} - \varepsilon$ is a sum of Gaussian random variables and is thus Gaussian. Asymptotically, the entire expression converges in distribution to a Gaussian random variable. We now find its limiting mean and variance.

The intercept estimator $\hat{\theta}_0$ is consistent, i.e., $\hat{\theta}_0 \xrightarrow{P} \theta_0^\star$. The mean of the decision error therefore converges to $E[\sigma\kappa - \varepsilon] = \sigma\kappa$.

The variance of the decision error is:

$$\begin{aligned}
\text{Var}\left(w_{\text{ETO}}(\boldsymbol{X}) - D \mid \text{training data}\right) &= \text{Var}\left((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)^\top \boldsymbol{X} - \varepsilon\right) \\
&= \text{Var}\left((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)^\top \boldsymbol{X}\right) + \text{Var}(\varepsilon) \\
&= (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)^\top \text{Cov}(\boldsymbol{X})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star) + \sigma^2 \\
&= \frac{1}{p}\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star\|_2^2 + \sigma^2.
\end{aligned}$$

Here, we used the fact that $\text{Cov}(\boldsymbol{X}) = E[\boldsymbol{X}\boldsymbol{X}^\top] = (1/p)\mathbf{I}_p$.

C.3 Asymptotic Behavior of OLS Estimators

To find the limit of the variance, we need the asymptotic limits for the parameter error $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star\|_2^2$ and the noise estimate $\hat{\sigma}^2$. These are standard results in high-dimensional statistics under our Gaussian design.

1. Parameter Error. The normalized squared error of the OLS slope coefficients converges in probability to a deterministic constant. Specifically, as shown in [13, Eq. (21)] for squared loss,

$$\frac{1}{p}\|\hat{\boldsymbol{\theta}}_{\text{OLS}} - \boldsymbol{\theta}^\star\|_2^2 \xrightarrow{P} \frac{\sigma^2}{\delta - 1}. \tag{7}$$

2. Noise Variance Estimator. The standard unbiased estimator for the noise variance is $\hat{\sigma}^2 = \frac{1}{n-p}\|D - X\hat{\boldsymbol{\theta}} - \hat{\theta}_0\mathbf{1}\|_2^2$. Since $\frac{(n-p)\hat{\sigma}^2}{\sigma^2}$ follows a χ^2_{n-p} distribution, the weak law of large numbers implies that for $n - p \to \infty$, the estimator is consistent:

$$\hat{\sigma}^2 \xrightarrow{P} \sigma^2. \tag{8}$$

C.4 Deriving the Limiting Regret

Substituting the asymptotic results (7) and (8) into our expressions for the mean and variance of the decision error, we find that $w_{\text{ETO}}(\boldsymbol{X}) - D$ converges in distribution to a Gaussian variable $\mathcal{N}(\mu^\star, (s^\star)^2)$, where:

- The limiting mean is $\mu^\star = \sigma\kappa$.
- The limiting variance is $(s^\star)^2 = \lim_{n,p\to\infty}\left(\frac{1}{p}\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star\|_2^2 + \sigma^2\right) = \frac{\sigma^2}{\delta-1} + \sigma^2 = \sigma^2\left(1 + \frac{1}{\delta-1}\right) = \frac{\sigma^2\delta}{\delta-1} = \frac{\sigma^2}{1-1/\delta}$.

This implies the limiting standard deviation is $s^\star = \sigma(1 - 1/\delta)^{-1/2}$.

The expected out-of-sample cost of the ETO policy thus converges to $E[L_{nv}(\mu^\star + s^\star Z)]$, where $Z \sim \mathcal{N}(0,1)$. Using the notation from the main text, this is precisely $\ell(\mu^\star, s^\star)$.

The expected cost of the oracle policy is a known formula, $E[L_{nv}(w^\star(\boldsymbol{X}) - D)] = (b+h)\sigma\phi(\kappa)$. Therefore, the asymptotic regret converges to:

$$\mathcal{R}_n \to \ell\big(\sigma\kappa, \sigma(1 - 1/\delta)^{-1/2}\big) - (b+h)\sigma\phi(\kappa).$$

This completes the proof for the ETO case in Theorem 1.

D Proof for Theorem 1 (ILO Part)

This section provides the full derivation for the ILO case of Theorem 1. We demonstrate how the high-dimensional empirical risk minimization problem is converted into the scalar saddle-point problem whose optimality conditions are explicitly stated in Appendix A.

D.1 Setup and Primary Optimization (PO) Problem

We adopt the notation and high-dimensional regime ($n, p \to \infty$ with $n/p \to \delta > 1$) from the main text. The ILO empirical program, as defined in (5), minimizes the average newsvendor loss:

$$\min_{\theta_0\in\mathbb{R},\ \boldsymbol{\theta}\in\mathbb{R}^p,\ s\geq\sigma} \frac{1}{n}\sum_{i=1}^n L_{nv}\Big(\theta_0 + \boldsymbol{X}_i^\top\boldsymbol{\theta} + s\,\kappa - D_i\Big), \tag{9}$$

where $L_{nv}(y) = h(y)^+ + b(-y)^+$, and $\kappa = \Phi^{-1}\big(b/(b+h)\big)$.

To analyze this problem, we substitute the true demand model $D_i = \theta_0^\star + \boldsymbol{\theta}^{\star\top}\boldsymbol{X}_i + \varepsilon_i$ into the loss function. The argument of L_{nv} becomes:

$$\theta_0 + \boldsymbol{X}_i^\top\boldsymbol{\theta} + s\,\kappa - D_i = (\theta_0 - \theta_0^\star) + \boldsymbol{X}_i^\top(\boldsymbol{\theta} - \boldsymbol{\theta}^\star) + s\,\kappa - \varepsilon_i.$$

Let the parameter errors be $d_0 = \theta_0 - \theta_0^\star$ and $\boldsymbol{d}_\theta = \boldsymbol{\theta} - \boldsymbol{\theta}^\star$. Since minimizing over $(\theta_0, \boldsymbol{\theta})$ is equivalent to minimizing over the errors $(d_0, \boldsymbol{d}_\theta)$, the optimization problem can be rewritten in terms of these errors:

$$\min_{d_0\in\mathbb{R},\ \boldsymbol{d}_\theta\in\mathbb{R}^p,\ s\geq\sigma} \frac{1}{n}\sum_{i=1}^n L_{nv}\Big(d_0 + \boldsymbol{X}_i^\top\boldsymbol{d}_\theta + s\,\kappa - \varepsilon_i\Big). \tag{10}$$

For analytical convenience, we define a new variable $\mu := d_0 + s\kappa$. The argument of the loss function simplifies to $\mu + \boldsymbol{X}_i^\top \boldsymbol{d}_\theta - \varepsilon_i$. This form is amenable to analysis via the CGMT.

Using the Fenchel conjugate $L_{nv}(y) = \max_{u\in[-b,h]} uy$, the problem can be written in the saddle-point Primary Optimization (PO) form:

$$\mathcal{P}_n = \min_{\boldsymbol{d}_\theta\in\mathbb{R}^p,\ \mu\in\mathbb{R}} \ \max_{\boldsymbol{u}\in[-b,h]^n} \ \frac{1}{n}\Big(\boldsymbol{u}^\top \boldsymbol{X}\boldsymbol{d}_\theta + \mu\,\boldsymbol{u}^\top \mathbf{1}_n - \boldsymbol{u}^\top \boldsymbol{\varepsilon}\Big). \tag{11}$$

D.2 Auxiliary Optimization (AO) via CGMT

To simplify the coupled term $\boldsymbol{u}^\top \boldsymbol{X}\boldsymbol{d}_\theta$, we apply the Convex Gaussian Min-max Theorem [13]. First, we rescale the variables. Let $\boldsymbol{G} := \sqrt{p}\,\boldsymbol{X}$ be a matrix with i.i.d. $\mathcal{N}(0,1)$ entries, and let $\boldsymbol{v} := \boldsymbol{d}_\theta/\sqrt{p}$. The objective of (11), scaled by n, is:

$$n\mathcal{P}_n = \min_{\boldsymbol{v}\in\mathbb{R}^p,\ \mu\in\mathbb{R}} \ \max_{\boldsymbol{u}\in[-b,h]^n} \ \big(\boldsymbol{u}^\top \boldsymbol{G}\boldsymbol{v} + \mu\,\boldsymbol{u}^\top \mathbf{1}_n - \boldsymbol{u}^\top \boldsymbol{\varepsilon}\big).$$

By the CGMT, we can replace the term $\boldsymbol{u}^\top \boldsymbol{G}\boldsymbol{v}$ with $\|\boldsymbol{u}\|_2\,\boldsymbol{\xi}^\top \boldsymbol{v} + \|\boldsymbol{v}\|_2\,\boldsymbol{g}^\top \boldsymbol{u}$, where $\boldsymbol{g} \in \mathbb{R}^n$ and $\boldsymbol{\xi} \in \mathbb{R}^p$ are independent standard Gaussian vectors. This yields the Auxiliary Optimization (AO) problem:

$$\mathcal{A}_n = \min_{\boldsymbol{v}\in\mathbb{R}^p,\ \mu\in\mathbb{R}} \ \max_{\boldsymbol{u}\in[-b,h]^n} \ \Big(\|\boldsymbol{u}\|_2\,\boldsymbol{\xi}^\top \boldsymbol{v} + \|\boldsymbol{v}\|_2\,\boldsymbol{g}^\top \boldsymbol{u} + \mu\,\boldsymbol{u}^\top \mathbf{1}_n - \boldsymbol{u}^\top \boldsymbol{\varepsilon}\Big). \tag{12}$$

We solve the inner minimization over $\boldsymbol{v}$ by decomposing it into magnitude and direction. Let $\alpha_v := \|\boldsymbol{v}\|_2 \geq 0$. The minimization over the direction of $\boldsymbol{v}$ yields $\min_{\|\hat{\boldsymbol{v}}\|=1} \boldsymbol{\xi}^\top \hat{\boldsymbol{v}} = -\|\boldsymbol{\xi}\|_2$. The AO simplifies to:

$$\mathcal{A}_n = \min_{\alpha_v\geq 0,\ \mu\in\mathbb{R}} \ \max_{\boldsymbol{u}\in[-b,h]^n} \ \Big(-\alpha_v\|\boldsymbol{u}\|_2\|\boldsymbol{\xi}\|_2 + \alpha_v\,\boldsymbol{g}^\top \boldsymbol{u} + \mu\,\boldsymbol{u}^\top \mathbf{1}_n - \boldsymbol{u}^\top \boldsymbol{\varepsilon}\Big).$$

The term $-\alpha_v\|\boldsymbol{\xi}\|_2\|\boldsymbol{u}\|_2$ couples all components of $\boldsymbol{u}$ through the norm. We decouple it using the variational identity $-C\|\boldsymbol{u}\|_2 = \min_{\lambda>0}\Big\{-\frac{\lambda}{2}\|\boldsymbol{u}\|_2^2 - \frac{C^2}{2\lambda}\Big\}$, where $C = \alpha_v\|\boldsymbol{\xi}\|_2$. Substituting this into the min-max problem gives:

$$\begin{aligned}
\mathcal{A}_n &= \min_{\alpha_v\geq 0,\ \mu\in\mathbb{R}} \ \min_{\lambda>0} \ \max_{\boldsymbol{u}\in[-b,h]^n} \left\{\alpha_v\,\boldsymbol{g}^\top \boldsymbol{u} + \mu\,\boldsymbol{u}^\top \mathbf{1}_n - \boldsymbol{u}^\top \boldsymbol{\varepsilon} - \frac{\lambda}{2}\|\boldsymbol{u}\|_2^2\right\} - \frac{\alpha_v^2\|\boldsymbol{\xi}\|_2^2}{2\lambda} \\
&= \min_{\alpha_v\geq 0,\ \mu\in\mathbb{R}} \ \min_{\lambda>0} \ \sum_{i=1}^n \max_{u_i\in[-b,h]} \left\{u_i(\alpha_v g_i + \mu - \varepsilon_i) - \frac{\lambda}{2}u_i^2\right\} - \frac{\alpha_v^2\|\boldsymbol{\xi}\|_2^2}{2\lambda}.
\end{aligned} \tag{13}$$

The maximization over the vector $\boldsymbol{u}$ has now decoupled into n independent scalar maximizations. We define the function $\psi(a;\lambda) := \max_{u\in[-b,h]}\{au - \frac{\lambda}{2}u^2\}$. The AO problem is thus:

$$\mathcal{A}_n = \min_{\alpha_v\geq 0,\ \mu\in\mathbb{R}} \ \min_{\lambda>0} \ \sum_{i=1}^n \psi\big(\alpha_v g_i + \mu - \varepsilon_i;\lambda\big) - \frac{\alpha_v^2\|\boldsymbol{\xi}\|_2^2}{2\lambda}. \tag{14}$$

D.3 Deterministic Equivalent (DE) and KKT Characterization

We now divide (14) by n and pass to the high-dimensional limit. By the law of large numbers, $\|\boldsymbol{\xi}\|_2^2/p \to 1$ almost surely, so $\|\boldsymbol{\xi}\|_2^2/n \to 1/\delta$. The sum becomes an expectation over the i.i.d. random variables $g \sim \mathcal{N}(0,1)$ and $\varepsilon \sim \mathcal{N}(0,\sigma^2)$. The random variable in the expectation, $Y := \alpha_v g + \mu - \varepsilon$, follows a Gaussian distribution $\mathcal{N}(\mu, s^2)$ where $s^2 := \alpha_v^2 + \sigma^2$. Substituting $\alpha_v^2 = s^2 - \sigma^2$, the AO objective converges to the Deterministic Equivalent (DE) problem from Theorem 1:

$$\min_{\mu\in\mathbb{R},\ s\geq\sigma} \ \max_{\lambda>0} \ \mathcal{J}(\mu,s,\lambda) := \mathbb{E}\left[\psi(Y;\lambda)\right] - \frac{s^2-\sigma^2}{2\lambda\delta}. \tag{15}$$

The saddle point $(\mu_\star, s_\star, \lambda_\star)$ of this objective is characterized by the Karush-Kuhn-Tucker (KKT) conditions, $\nabla\mathcal{J}(\mu_\star, s_\star, \lambda_\star) = 0$. Taking the partial derivatives with respect to μ, s, and λ yields the following abstract conditions:

$$\partial_\mu \mathcal{J} = \mathbb{E}[\partial_\mu \psi(Y; \lambda_\star)] = \mathbb{E}[u^*(Y; \lambda_\star)] = 0, \tag{K1}$$

$$\begin{aligned}\partial_s \mathcal{J} &= \mathbb{E}\left[\partial_s \psi(Y; \lambda_\star)\right] - \frac{s_\star}{\lambda_\star \delta} = \mathbb{E}\left[u^*(Y; \lambda_\star) \cdot \frac{\partial Y}{\partial s}\right] - \frac{s_\star}{\lambda_\star \delta} \\ &= \frac{1}{s_\star}\,\mathbb{E}\left[(Y - \mu_\star) u^*(Y; \lambda_\star)\right] - \frac{s_\star}{\lambda_\star \delta} = 0,\end{aligned} \tag{K2}$$

$$\partial_\lambda \mathcal{J} = \mathbb{E}[\partial_\lambda \psi(Y; \lambda_\star)] + \frac{s_\star^2 - \sigma^2}{2\lambda_\star^2 \delta} = -\frac{1}{2}\,\mathbb{E}[(u^*(Y; \lambda_\star))^2] + \frac{s_\star^2 - \sigma^2}{2\lambda_\star^2 \delta} = 0. \tag{K3}$$

Here, $u^*(Y; \lambda) = \Pi_{[-b,h]}(Y/\lambda)$ is the optimal dual variable, and $Y \sim \mathcal{N}(\mu_\star, s_\star^2)$.

Connection to Appendix A. The final step is to explicitly compute the Gaussian expectations in (K1)-(K3). This involves standard but lengthy piecewise integration of the function u^* and its moments against the Gaussian PDF. Performing these integrations leads directly to the system of nonlinear equations presented in Appendix A, namely equations (6a), (6b), and (6c). This establishes the final form of the DE optimality conditions and completes the proof for the ILO case of Theorem 1.

D.4 Deriving the Limiting Regret

The core result of the preceding analysis is that the out-of-sample decision error of the ILO policy, $w_{\text{ILO}}(\boldsymbol{X}) - D$, converges in distribution to a Gaussian random variable whose parameters are determined by the solution to the DE problem. Specifically,

$$w_{\text{ILO}}(\boldsymbol{X}) - D \xrightarrow{d} \mathcal{N}(\mu_\star, s_\star^2),$$

where $(\mu_\star, s_\star)$ is part of the solution to the KKT system in Appendix A.

The asymptotic expected out-of-sample cost of the ILO policy is therefore the expected newsvendor loss for this limiting Gaussian error:

$$\mathcal{C}_{\text{ILO}} = \mathbb{E}[L_{nv}(\mu_\star + s_\star Z)], \quad \text{where } Z \sim \mathcal{N}(0, 1).$$

Using the notation from the main text, this cost is precisely $\ell(\mu_\star, s_\star)$.

The regret is the difference between this asymptotic cost and the cost of the oracle policy, $w^\star(\boldsymbol{X})$. The oracle's decision error is $w^\star(\boldsymbol{X}) - D = \sigma\kappa - \varepsilon \sim \mathcal{N}(\sigma\kappa, \sigma^2)$, and its expected cost is a known constant, $(b + h)\sigma\phi(\kappa)$.

Therefore, the asymptotic regret for the ILO policy is:

$$\mathcal{R}_n \to \mathcal{C}_{\text{ILO}} - \mathcal{C}_{\text{oracle}} = \ell(\mu_\star, s_\star) - (b + h)\sigma\phi(\kappa).$$

This completes the derivation for the ILO case in Theorem 1.