Stochastic Inverse Optimization

John R. Birge[†] Xiaocheng Li[‡] Chunlin Sun[◊]

[†] The University of Chicago Booth School of Business, John.Birge@chicagobooth.edu
[‡] Imperial College Business School, Imperial College London, xiaocheng.li@imperial.ac.uk
[◊] Institute for Computational and Mathematical Engineering, Stanford University, chunlin@stanford.edu

Abstract

The inverse linear optimization considers the problem of inferring the objective coefficient vector of a linear program through the observation(s) of the constraints and the optimal solution. We consider a stochastic setting where the objective coefficient vector follows some unknown distribution and the goal becomes to estimate the distribution from observations of constraints and the optimal solution. The existing approaches of inverse linear optimization all rely on the existence of one objective coefficient vector that is consistent with all the observations and thus will fail in a stochastic setting where each observation may correspond to a different objective vector. We consider two settings for stochastic inverse optimization: a Gaussian setting where the objective coefficient follows von Mises-Fisher distribution, and a δ-corruption setting where the objective coefficient distribution concentrates on one vector with high probability and it is arbitrarily corrupted otherwise. We devise algorithms to estimate the underlying distribution under both settings and develop theoretical guarantees accordingly. We illustrate the algorithm performance through numerical experiments.

1 Introduction

The problem of inverse linear programming considers the problem of inferring the objective coefficient vector of a linear program through the observation(s) of one single or multiple pairs of the constraints and the optimal solution. Early works (Zhang and Liu, 1996; Ahuja and Orlin, 2001) study the case of one single observation and aim to find one objective coefficient vector that (i) is consistent with the observation and (ii) minimizes the distance to a pre-specified vector. Different variants of the problem are later studied: the inverse convex programming (Keshavarz et al., 2011), the inverse optimal value problem (Ahmed and Guan, 2005), the multiple-objective setting (Chan et al., 2014), etc. While the forward optimization problem solves for the optimal solution with the knowledge of the objective function, the inverse optimization problem seeks to recover the unknown objective function through the observations of the optimal solution. The inverse optimization problem finds its application under a wider range of contexts: geoscience (Burton and Toint, 1992), finance (Bertsimas et al., 2012), energy (Aswani et al., 2018), health (Chan et al., 2014), market analysis (Birge et al., 2017) etc.

In this paper, we study the problem under a stochastic setting where the unobserved coefficient vector of the LP is sampled from some unknown distribution. The goal is to estimate the distribution through observations of the constraints and the optimal solutions. Such a stochastic setting can be motivated from many application contexts. For example, one stream of literature (Beigman and Vohra, 2006; Aswani et al., 2018) employs the inverse optimization models to study customer behaviour where the objective function encodes the customer latent preference and the observations of the optimal solution correspond to the customer's purchasing behaviour. In such a context, the customers may have different preferences, and thus the stochastic utility (objective coefficient vector) captures the heterogeneity of the customer preference. Accordingly, the goal becomes to uncover the distribution of customer preferences rather

than to identify one shared preference for all the customers. To the best of our knowledge, our paper provides the first result in the literature of inverse optimization that allows the unobserved objective function to be stochastically distributed.

We consider two settings for stochastic inverse optimization: a Gaussian setting and a δ-corruption setting. For the Gaussian setting, we assume the objective coefficient vector follows the von Mises–Fisher distribution, i.e., the restriction of a multivariate Gaussian distribution to the unit sphere. The setting can be viewed as the most straightforward generalization of the deterministic setting. For the δ-corruption setting, we assume the objective coefficient vector concentrates on an unknown vector with probability $1-\delta$ and follows an arbitrarily corrupted distribution with probability δ. The setting can be viewed as a robust version of the traditional inverse optimization where the observations are subject to adversarial corruption or contaminated with outliers. For both settings, a key distinction from the traditional setting is that there does not exist an objective coefficient vector that is consistent with all the observations. The existing approaches more or less rely on this consistency for algorithm design and analysis and thus will fail for the stochastic setting.

1.1 Related Literature

Traditional inverse optimization.

Burton and Toint (1992, 1994) first study the inverse optimization problem for the shortest path problem. The formulation is then generalized and studied in (Zhang and Liu, 1996; Ahuja and Orlin, 2001) as inverse linear optimization with one single observation. Keshavarz et al. (2011) investigate the inference of a more general convex objective function, and a line of works consider the multi-objective setting (Chan et al., 2014; Dong and Zeng, 2020b). The inverse optimal value problem (Ahmed and Guan, 2005) considers a setting where the optimal objective value instead of the optimal solution is observed. Optimization-based approaches are usually adopted for this line of works with the aim to translate the inverse problem as a convex or linear programming problem. Recently, Keshavarz et al. (2011) and Aswani et al. (2018) study a statistical setting where the observations are noisily sampled from some distribution. Specifically, Aswani et al. (2018) consider a setting where the optimal solutions are contaminated with some independent noises (justified as bounded rationality (Tversky and Kahneman, 1985)). We remark that all these existing works consider a deterministic regime where the objective function is parameterized by some unknown but fixed parameters (e.g., objective coefficients for linear programs), and the goal is to estimate the parameters through observations. These works aim to develop efficient procedures to identify a vector of parameters that is consistent with all the observations. In this light, our stochastic formulation complements the line of work and solves the inverse optimization for the case when such consistency no longer exists.

Data-driven inverse optimization.

A line of work (Mohajerin Esfahani et al., 2018; Tan et al., 2020; Besbes et al., 2021) studies the inverse optimization under a data-driven setting. Under the data-driven setting, the goal is still to recover the objective function, but each observation (of the constraints and the optimal solution) is associated with a vector of covariates. The covariate vector is also known as context or side information, and it presumably encodes information about the unobserved objective function. Mohajerin Esfahani et al. (2018) take a distributionally robust approach and derive generalization bounds on the suboptimality gap for the estimated parameter. Tan et al. (2020) take a machine learning perspective and draw a connection with the line of work on contextual linear programs (Elmachtoub and Grigas, 2022). Besbes et al. (2021) generalize the inverse linear optimization problem and apply a non-linear contextual function on the decision variables in the objective function. Several recent works (Bärmann et al., 2018; Dong et al.,

2018a; Dong and Zeng, 2020a; Chen and Kılınç-Karzan, 2020) cast the inverse optimization problem in an online context and develop (online) gradient-based algorithms. In Section 4.2, we also develop online algorithms for the δ-corruption setting which can be viewed as a generalization of the algorithms along this line. As before, this line of literature on data-driven inverse optimization also assumes the existence of a vector of parameters that is consistent with all the observations, so the settings should be viewed as deterministic rather than stochastic.

Revealed preference.

A separate line of works focuses on a special case of inverse optimization where there is only one single constraint. The problem is often referred to as the revealed preference problem and has a long history in economics (Samuelson, 1948; Afriat, 1967). Recent works (Beigman and Vohra, 2006; Zadimoghaddam and Roth, 2012; Balcan et al., 2014; Amin et al., 2015) formulate the problem as a learning problem that aims to find a utility function which explains a set of past observations and predicts the future behavior of a utility-maximization agent. Beigman and Vohra (2006) initiate this line of research and study a statistical setup where the input data is a set of observations and the performance of a learning algorithm is evaluated by sample complexity bounds. Different from the inverse optimization literature, the literature of learning from revealed preference also studies the query-based model where the observations of the constraints and the optimal solution are collected in an active manner. Through querying an oracle for the optimal decision (solution) of customers, the goal is to minimize the number of samples required for a sufficiently accurate estimation of the utility function. Zadimoghaddam and Roth (2012) study the case of a linear or linearly separable concave utility function, and Balcan et al. (2014) generalize the setting and devises learning algorithms for several classes of utility functions. Some subsequent works along this line study the associated revenue management problem (Amin et al., 2015) and a game-theoretic setting (Dong et al., 2018b) where the agents act strategically to hide the true actions. While the existing works all consider a deterministic preference setting, our paper considers a stochastic preference setting where the customer's preference is sampled from an unknown distribution.

Linear programming with side information.

One stream of research studies the linear programming problem with the presence of side information. The problem is studied under different names such as prescriptive analytics (Bertsimas and Kallus, 2020), smart-predict-then-optimize (Elmachtoub and Grigas, 2022), and end-to-end optimization (Ho-Nguyen and Kılınç-Karzan, 2022). The formulation and methodology lie in the intersection of operations research and machine learning. The goal is to predict the unseen objective function/optimal solution of an optimization problem using the available covariates/side information. Different from the inverse optimization problem, the decision maker is equipped with history observations that contain not only the covariates and the constraints, but also the objective function.

2 Model Setup

In this section, we formulate the problem of stochastic inverse (linear) optimization, which takes the following standard-form linear program (LP) as its underlying form

$$\begin{aligned} \mathrm{LP}(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}) \coloneqq \max_{\boldsymbol{x}} \ & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \tag{1}$$

where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, and $\boldsymbol{b} \in \mathbb{R}^m$ are the inputs of the LP. The decision variables $\boldsymbol{x} = (x_1, ..., x_n)^\top \in \mathbb{R}^n$. While the formulation (1) covers all the LPs as special cases, one can interpret the objective as

maximizing a linear utility function subject to some resource constraints. Throughout the paper, we use n to denote the number of decision variables and m to denote the number of constraints.

We make the following distributional assumption on the LP's inputs.

Assumption 1 (Distribution)**.** *We assume:*

- *The objective coefficient vector* $\boldsymbol{c}$ *follows some unknown distribution* $\mathcal{P}_c$*.*
- *The LP's input* $(\boldsymbol{A}, \boldsymbol{b})$ *follows some unknown distribution* $\mathcal{P}_{A,b}$ *independent of* $\mathcal{P}_c$*.*
- *The LP's optimal objective value is finite with probability* 1*.*

The conventional setting of inverse optimization aims to recover the unobserved objective coefficient vector through the observation(s) of the optimal solution and the constraints. For our stochastic setting, the objective then becomes to estimate the underlying distribution of $\mathcal{P}_c$ using the observations. Formally, the aim of stochastic inverse optimization is to estimate the distribution through the dataset

$$\mathcal{D}_T = \{(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\}_{t=1}^T . \tag{2}$$

Here the t-th sample corresponds to an unobserved $\boldsymbol{c}_t$ generated from $\mathcal{P}_c$ and $\boldsymbol{x}_t^*$ is one optimal solution of $\mathrm{LP}(\boldsymbol{c}_t, \boldsymbol{A}_t, \boldsymbol{b}_t)$. We note that there is a *scale invariance* of the objective coefficient, in that $\boldsymbol{c}_t$ and $\alpha \cdot \boldsymbol{c}_t$ give the same optimal solution $\boldsymbol{x}_t^*$ for any $\alpha > 0$. Hence these two are indistinguishable based on the observations of $\boldsymbol{x}_t^*$'s. Therefore, we restrict the distribution $\mathcal{P}_c$ to the unit sphere $\mathcal{S}^{n-1} = \{\boldsymbol{c} : \|\boldsymbol{c}\|_2 = 1\}$.

In this paper, we study two settings: (i) a Gaussian stochastic setting and (ii) a δ-corruption setting. For (i), we assume that $\mathcal{P}_c$ follows the von Mises–Fisher distribution, i.e., the restriction of a multivariate Gaussian distribution to the unit sphere. The results developed under this setting can also be generalized to other spherical distributions, and the Gaussian distribution is chosen for demonstration purpose. For (ii), we assume $\mathcal{P}_c$ concentrates on an unknown vector $\boldsymbol{c}^*$ with probability $1-\delta$ and follows an arbitrarily corrupted distribution with probability δ. The setting can be viewed as a robust version of the traditional inverse optimization where the observations are subject to adversarial corruption or contaminated with outliers.

The challenge of the problem.

We first discuss the challenge of the stochastic settings in general and thus distinguish our results from the existing approaches of inverse optimization. Specifically, each observation $(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$ in the dataset $\mathcal{D}_T$ prescribes a set $\mathcal{C}_t \subset \mathcal{S}^{n-1}$,

$$\mathcal{C}_t \coloneqq \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \boldsymbol{x}_t^* \text{ is an optimal solution of } \mathrm{LP}(\boldsymbol{c}, \boldsymbol{A}_t, \boldsymbol{b}_t)\right\}.$$

The set $\mathcal{C}_t$ captures all the possible values of $\boldsymbol{c}_t$ that is consistent with the t-th observation. The following proposition states a standard result of linear programming (**?**) that the set $\mathcal{C}_t$ can be derived from the optimality condition and expressed by a group of linear constraints.

Proposition 1. *For each set* $\mathcal{C}_t$*, there exists a matrix* $\boldsymbol{V}_t$ *and a vector* $\boldsymbol{u}_t$ *dependent on* $(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$ *such that*

$$\mathcal{C}_t = \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \boldsymbol{V}_t \boldsymbol{c} \le \boldsymbol{u}_t\right\}.$$

In the conventional setting of inverse optimization, there exists at least one $\boldsymbol{c}^*$ that is consistent with all the observations, and thus the problem reduces to finding one feasible $\boldsymbol{c}$ in the set of $\cap_{t=1}^T \mathcal{C}_t$. In contrast, for a stochastic setting, it may happen that the set of $\cap_{t=1}^T \mathcal{C}_t$ is empty. To the best of

our knowledge, all the existing approaches of inverse optimization build upon the assumption that the intersection set is non-empty and thus will fail under this stochastic setting. Figure 1 further illustrates the intuition of "empty intersection", and it provides a conceptual visualization of the problem under a 1-dimensional and 2-dimensional case ($n = 1, 2$). We remark that the figure is just for illustrative purpose as the problem may not be well-defined in the 1-dimensional case.

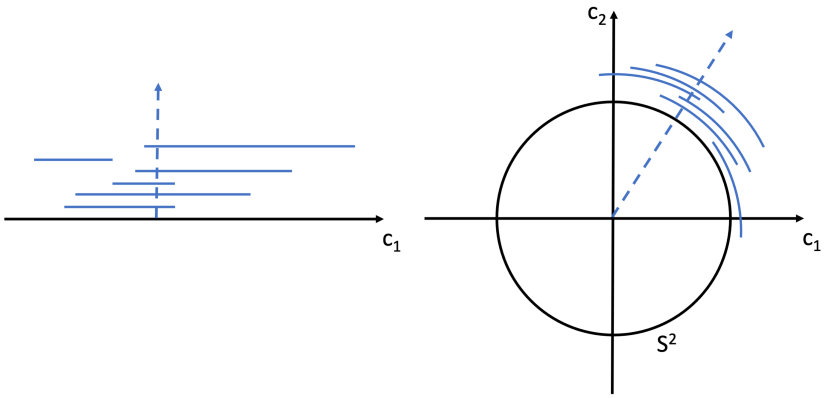

Figure 1: Visualizing the estimation problem in 1-dimensional and 2-dimensional case. For both panels, the blue solid segments denote the sets $\mathcal{C}_t$'s, and the blue dashed lines represent a value of $\boldsymbol{c}$ that appears most frequently in these $\mathcal{C}_t$'s. In the conventional setting of inverse optimization, there must exist one vector $\boldsymbol{c}$ that intersects with all the $\mathcal{C}_t$ (blue solid segments), but this may not be the case for the stochastic setting.

Another perspective is to view the stochastic inverse problem as a statistical estimation problem. The goal is to estimate the distribution of $\mathcal{P}_c$ (for example, von Mises-Fisher distribution) without the observations of the realized samples $\boldsymbol{c}_t$'s, but merely with the knowledge of $\mathcal{C}_t$ to which $\boldsymbol{c}_t$ belongs. At the first place, the estimation problem is more challenging than the canonical setting where the realized samples are fully observed. Also, the sample efficiency of the estimation procedure is naturally contingent on the dispersion of $\mathcal{C}_t$ which is essentially determined by the generation of $(\boldsymbol{A}_t, \boldsymbol{b}_t)$. For example, if all the $\mathcal{C}_t$'s coincide with each other, then one can hardly learn much about the underlying $\mathcal{P}_c$. In this paper, we aim to pinpoint conditions on $\mathcal{P}_{A,b}$ under which the estimation of $\mathcal{P}_c$ is possible. Furthermore, an alternative way to measure the estimation accuracy is to evaluate the predictive performance of the estimated model on new observations generated from $\mathcal{P}_{A,b}$, and such performance bounds generally bear less dependency on the distribution of $\mathcal{P}_{A,b}$. We also provide theoretical guarantees in this sense.

3 Inverse Optimization with Gaussian Distributed Objective

In this section, we study a setting where the distribution $\mathcal{P}_c$ follows the von Mises-Fisher distribution. Specifically, the distribution is parameterized by $\boldsymbol{\theta} = (\boldsymbol{\mu}, \kappa)$ with the density function

$$f(\boldsymbol{c}; \boldsymbol{\theta}) \coloneqq \frac{\exp\left(\kappa \boldsymbol{\mu}^\top \boldsymbol{c}\right)}{\int_{\boldsymbol{c}' \in \mathcal{S}^{n-1}} \exp(\kappa \boldsymbol{\mu}^\top \boldsymbol{c}') d\boldsymbol{c}'} \propto \exp\left(\kappa \boldsymbol{\mu}^\top \boldsymbol{c}\right).$$

Here the vector $\boldsymbol{\mu} \in \mathbb{R}^n$ represents the mean direction and the parameter $\kappa > 0$ controls the concentration of the distribution. The conventional setting of the inverse optimization problem can be viewed as the case when $\kappa = \infty$, and then the objective coefficient vector becomes deterministic and is equal to $\boldsymbol{\mu}$ with probability one. Denote the true parameters of the distribution $\mathcal{P}_c$ by $\boldsymbol{\theta}^* = (\boldsymbol{\mu}^*, \kappa^*)$. The goal is to estimate the true parameters using the dataset $\mathcal{D}_T$.

The likelihood of the dataset $\mathcal{D}_T$ under parameter $\boldsymbol{\theta}$ is given by

$$l(\mathcal{D}_T|\boldsymbol{\theta}) \coloneqq \prod_{t=1}^{T} l\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}\right)$$
$$= \prod_{t=1}^{T} \mathbb{P}(\boldsymbol{c}_t \in \mathcal{C}_t|\boldsymbol{\theta}) = \prod_{t=1}^{T} \int_{\boldsymbol{c}\in\mathcal{C}_t} f(\boldsymbol{c};\boldsymbol{\theta})d\boldsymbol{c}.$$

Recall that the set $\mathcal{C}_t$ comes from the LP's optimality condition in Proposition 1 and it contains all the objective coefficient vectors that are consistent with the t-th observation in the dataset.

We remark that since the likelihood function is not closed-form in $\boldsymbol{\theta}$, it prevents a direct maximum likelihood estimation approach. Specifically, this is because the integration of $f(\boldsymbol{c};\boldsymbol{\theta})$ over the region $\mathcal{C}_t$ is not closed-form, and, in fact, the issue is not only pertaining to the Gaussian parameterization of $\mathcal{P}_c$. The scale-invariant property of the objective coefficient vector restricts the distribution $\mathcal{P}_c$ to a unit sphere (or a simplex), and consequently, the likelihood function inevitably involves a non-closed-form integration. The issue can be partially resolved by using the Monte Carlo method to approximate the integration, but this will prevent the usage of gradient-based algorithms to obtain the maximum likelihood estimation. To make things even worse, we show in the following that the (negative) likelihood or log-likelihood function is non-convex and can have local minimum with arbitrarily large gap against the global minimum.

3.1 Non-convexity of Likelihood Function

Proposition 2 gives a negative result on the approach of maximum likelihood estimation (MLE). The idea of its proof is illustrated by Figure 2 for the case of two samples ($T = 2$). From the plot, it is easy to see that the mean direction $\boldsymbol{\mu} = (1, 0)^\top$ gives a large likelihood and is the global maximizer of the likelihood function. In fact, its opposite direction, as a drastically wrong answer to the estimation problem, is a local maximizer of the likelihood function. Thus this prevents the application of gradient-based approach for MLE as it may get stuck at the local optimum.

Proposition 2. *For the negative likelihood function $-l(\mathcal{D}_T|\boldsymbol{\theta})$ as a function of $\boldsymbol{\theta}$, there exists local minimum whose objective value has an arbitrarily large gap against the global minimum.*

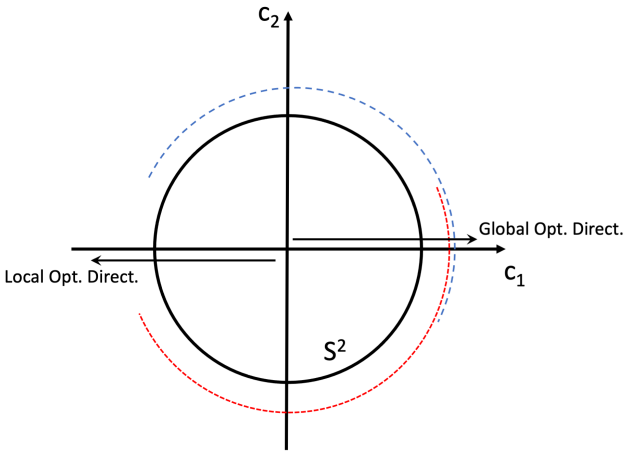

Figure 2: A 2-dimensional illustration of the non-convexity of the likelihood function with two samples ($T = 2$). The blue and red dashed lines denote the sets $\mathcal{C}_1$ and $\mathcal{C}_2$. The mean direction $\boldsymbol{\mu} = (1, 0)^\top$ is the global minimum of the negative likelihood function, while the mean direction $\boldsymbol{\mu} = (-1, 0)^\top$ is a local minimum.

Arguably, this type of non-convexity phenomenon illustrated in Figure 2 is not rare when the sample

number T is small. In the next subsection, we provide a positive result that when T is large, although the likelihood function may still be non-convex, it exhibits a concentration behaviour around the true parameters $\boldsymbol{\theta}^*$. Together with Proposition 2, it characterizes the landscape of the likelihood function for the estimation problem.

3.2 Posterior Sampling and Concentration

Now we propose a posterior sampling approach for the problem. The idea is, instead of maximizing the likelihood function, we directly draw samples from the posterior distribution. In this subsection, we justify this approach by establishing a concentration property of the posterior distribution.

First, we suppose there is a prior distribution $\mathbb{P}_0(\boldsymbol{\theta})$ satisfying the following assumption. The only role of the prior distribution is to ensure that it does not exclude any possible value of the true $\boldsymbol{\theta}^*$.

Assumption 2. *We assume that the concentration parameter $\kappa^* \in (\underline{\kappa}, \bar{\kappa})$ where $\underline{\kappa}, \bar{\kappa}$ are two known positive constants, and we choose the prior distribution $\mathbb{P}_0(\boldsymbol{\theta})$ as a uniform distribution over $\mathcal{S}^{n-1} \times (\underline{\kappa}, \bar{\kappa})$.*

Then the posterior distribution is defined by

$$\begin{aligned}\mathbb{P}_T(\boldsymbol{\theta}) &:= \frac{\mathbb{P}_0(\boldsymbol{\theta}) \cdot \mathbb{P}\left(\mathcal{D}_T|\boldsymbol{\theta}\right)}{\mathbb{P}\left(\mathcal{D}_T\right)} \\ &\propto \mathbb{P}_0(\boldsymbol{\theta}) \cdot \prod_{t=1}^{T} \int_{\boldsymbol{c}\in\mathcal{C}_t} f(\boldsymbol{c};\boldsymbol{\theta}) d\boldsymbol{c}.\end{aligned}$$

With slight abuse of notation, we use $\mathbb{P}_T(\cdot)$ (or $\mathbb{P}_0(\cdot)$) to refer to both the density function and the probability measure of the posterior (or prior) distribution.

The following theorem justifies the approach of posterior sampling. We first remark that the posterior sampling approach is just proposed to estimate the parameters, but all the theoretical results are stated in the frequentist language. The proof of Theorem 1 mimics the idea of the convergence of the posterior distribution (Ghosal et al., 2000; Chae et al., 2021). The main difficulty of the proof is to adapt the existing analysis for observable samples to the case of unobservable samples; notably, for our case, we do not observe $\boldsymbol{c}_t$'s but only know $\boldsymbol{c}_t \in \mathcal{C}_t$. While similar results also hold for other underlying distribution of $\mathcal{P}_c$, the von Mises-Fisher distribution provides more analytical convenience in deriving the result.

Theorem 1. *Let*

$$\Theta_T := \left\{ \boldsymbol{\theta} \in \mathcal{S}^{n-1} \times (\underline{\kappa}, \bar{\kappa}) : D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*\right)\right) \le \max\left(8, 8\bar{\kappa}\right) \frac{\sqrt{n} \cdot \log T}{\sqrt{T}} \right\}$$

where $D_{TV}(\cdot,\cdot)$ is the total variation distance between two distributions supported on $\mathbb{R}^n \times \mathbb{R}_+^n \times \mathbb{R}_+$ equipped with Euclidean metric. Then, under Assumptions 1-2,

$$1 - \mathbb{P}_T(\Theta_T) \to 0 \text{ in probability as } T \to \infty.$$

Specifically, the following inequality holds

$$\mathbb{E}\left[\mathbb{P}_T\left(\Theta_T\right)\right] \ge 1 - \frac{3}{T}.$$

where the expectation is taken with respect to the random distribution $\mathbb{P}_T(\cdot)$ (essentially, with respect to the dataset $\mathcal{D}_T$).

To interpret Theorem 1, we first note that each $\boldsymbol{\theta}$, together with the distribution of $\mathcal{P}_{A,b}$, defines a distribution over the space of $(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$. As we use observations $(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$'s to identify the true $\boldsymbol{\theta}^*$, the set Θ_T defines a set of indistinguishable $\boldsymbol{\theta}$'s based on the total variation distance between distributions of $(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$. The set Θ_T shrinks as $T \to \infty$. The posterior sampling approach samples from the distribution $\mathbb{P}_T(\cdot)$, and Theorem 1 states that the samples will be concentrated in set Θ_T with high probability. The posterior distribution $\mathbb{P}_T(\cdot)$ is dependent on the dataset $\mathcal{D}_T$, so it is a random distribution itself, and the results in Theorem 1 are stated in either convergence in probability or expectation. As a side note, the total variation distance in the theorem is not critical and it can be replaced with other distances such as the Hellinger distance.

Intuitively, Theorem 1 says that for some $\boldsymbol{\theta}$ such that the likelihood distribution $\mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}\right)$ differs from $\mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*\right)$ to a certain extent, the posterior $\mathbb{P}_T(\cdot)$ is unlikely to generate such $\boldsymbol{\theta}$. Intuitively, this is the best one can do in that the only observed data are $(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$'s. If some $\boldsymbol{\theta}$ gives a similar likelihood as the true $\boldsymbol{\theta}^*$, then we cannot distinguish $\boldsymbol{\theta}$ from the true $\boldsymbol{\theta}^*$ based on the dataset. In other words, the posterior distribution identifies the true $\boldsymbol{\theta}^*$ up to some "equivalence" in the likelihood distribution space. The following corollary further exploits this intuition that if there is an equivalence between the likelihood distribution space and the underlying parameter space, then the posterior distribution is capable of identifying the true parameter.

Corollary 1. *Suppose*

$$D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*\right)\right) > 0$$

for all $\boldsymbol{\theta} \neq \boldsymbol{\theta}^* \in \mathcal{S}^{n-1} \times [\underline{\kappa}, \bar{\kappa}]$. Then the posterior distribution $\mathbb{P}_T(\cdot)$ will converge to the point-mass distribution supported on $\boldsymbol{\theta}^*$ almost surely as $T \to \infty$. Moreover, suppose there exists a constant $M > 0$ satisfying

$$D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*\right)\right) \geq M \cdot \|\kappa\boldsymbol{\mu} - \kappa^*\boldsymbol{\mu}^*\|_2, \tag{3}$$

for all $\boldsymbol{\theta} = (\boldsymbol{\mu}, \kappa) \neq \boldsymbol{\theta}^* = (\boldsymbol{\mu}^*, \kappa^*) \in \mathcal{S}^{n-1} \times [\underline{\kappa}, \bar{\kappa}]$.

Under Assumptions 1-2, the following inequality holds with probability no less than $1 - \frac{1+4M}{\sqrt{T}}$,

$$\mathbb{E}_T\left[\|\kappa_T\boldsymbol{\mu}_T - \kappa^*\boldsymbol{\mu}^*\|_2\right] \leq \max\left(9, 9\bar{\kappa}\right) \cdot \frac{\sqrt{n}\log T}{M\sqrt{T}}$$

where $\boldsymbol{\theta}_T = (\boldsymbol{\mu}_T, \kappa_T)$ is sampled from the posterior distribution $\mathbb{P}_T(\cdot)$.

The corollary states that when there is some equivalence between the likelihood distribution space and the parameter space as (3), the true parameter is identifiable. The first part of the corollary states a consistency result that as long as all the $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$ are distinguishable from $\boldsymbol{\theta}^*$ through the likelihood function, then the posterior sampling will eventually identify the true $\boldsymbol{\theta}^*$. The second part relates to the convergence rate with an equivalence parameter M.

In Assumption 1, we assume the constraint input $(\boldsymbol{A}_t, \boldsymbol{b}_t)$ is generated from some distribution $\mathcal{P}_{A,b}$. We note that Theorem 1 and Corollary 1 hold without any additional assumption on $\mathcal{P}_{A,b}$, but the space topology of the likelihood distribution is highly dependent on $\mathcal{P}_{A,b}$. Specifically, the distribution of $(\boldsymbol{A}_t, \boldsymbol{b}_t)$ determines the separateness of the parameter space through affecting the value of M in (3) or even its existence. Practically, the value M for a specific distribution of $\mathcal{P}_{A,b}$ can be estimated through simulation. So if the decision maker has some flexibility in choosing the distribution of $\mathcal{P}_{A,b}$, the better choice would be the one that corresponds to a larger value of M. If the constraint input $(\boldsymbol{A}_t, \boldsymbol{b}_t)$ is not randomly generated but can be actively chosen, the results in Theorem 1 and Corollary 1 still hold by conditioning on all the $(\boldsymbol{A}_t, \boldsymbol{b}_t)$'s.

Corollary 2. *Let $(\boldsymbol{A}, \boldsymbol{b})$ be a new sample from $\mathcal{P}_{A,b}$, i.e., independent from the dataset $\mathcal{D}_T$, and let*

$\boldsymbol{\theta}_T = (\boldsymbol{\mu}_T, \kappa_T)$ be a sample from the posterior distribution $\mathbb{P}_T(\cdot)$. Denote $\tilde{\boldsymbol{x}}^*$ and $\boldsymbol{x}^*$ as the optimal solutions of LP$(\boldsymbol{\mu}_T, \boldsymbol{A}, \boldsymbol{b})$ and LP$(\boldsymbol{\mu}^*, \boldsymbol{A}, \boldsymbol{b})$, respectively. Then, under Assumptions 1-2, the following inequality holds with probability no less than $1 - \frac{5}{\sqrt{T}}$,

$$\mathbb{E}\left[\|\tilde{\boldsymbol{x}}^* - \boldsymbol{x}^*\|_2\right] \leq \max\left(16, 16\bar{\kappa}\right) \frac{\sqrt{n} \cdot \log T}{\sqrt{T}},$$

where the expectation is taken with respect to both the posterior distribution $\mathbb{P}_T(\cdot)$ and $(\boldsymbol{A}, \boldsymbol{b})$.

Corollary 2 provides an upper bound on the predictive performance of the posterior distribution. Specifically, we want to predict the optimal solution of a linear program specified by $\boldsymbol{\mu}^*$ (proportionally to $\mathbb{E}[\boldsymbol{c}]$) and a new sample of the constraint $(\boldsymbol{A}, \boldsymbol{b})$, and the prediction $\tilde{\boldsymbol{x}}^*$ is based on a posterior sample. We know from Theorem 1 that the posterior distribution concentrates on those $\boldsymbol{\theta}$'s that are indistinguishable from the true $\boldsymbol{\theta}^*$ in terms of the likelihood. Speaking of the predictive performance, we are only concerned with the distribution of the optimal solution (equivalently, the likelihood), but do not require the identification of the exact true $\boldsymbol{\theta}^*$, so Corollary 2 does not require the additional condition (3) to hold. In other words, even if the true $\boldsymbol{\theta}^*$ is not identifiable (for certain distribution of $(\boldsymbol{A}, \boldsymbol{b})$), the posterior sampling approach can still provide some performance guarantee in terms of predicting the optimal solution.

4 Inverse Optimization with δ-Corrupted Objective

In this section, we consider a second setting where $\boldsymbol{c}_t$, the t-th unobserved objective coefficient vector in the dataset $\mathcal{D}_T$ is specified by

$$\boldsymbol{c}_t = \begin{cases} \boldsymbol{c}^*, & \text{w.p. } 1-\delta, \\ \mathcal{P}_c', & \text{w.p. } \delta, \end{cases} \tag{4}$$

where $\boldsymbol{c}^* \in \mathcal{S}^{n-1}$ is a fixed vector, and $\mathcal{P}_c'$ is an arbitrary and unknown distribution. The parameter $\delta \in [0, 1]$ controls the intensity of corruption which makes the δ-corruption setting an interpolation between the conventional setting of inverse optimization and the Gaussian setting in the previous section. Specifically, the conventional setting corresponds to the case of $\delta = 0$, and the Gaussian setting in the previous section corresponds to the case of $\delta = 1$ and $\mathcal{P}_c'$ being the von-Mises Fisher distribution. While we allow the corruption distribution $\mathcal{P}_c'$ to be arbitrarily adversarially chosen, we only aim to recover the parameter $\boldsymbol{c}^*$ using the dataset $\mathcal{D}_T$. To this end, the results developed in this section can also be seen as a robust version of the existing inverse optimization approaches against adversarial corruption or statistical outliers.

4.1 Recovery of true objective

A natural approach to estimate $\boldsymbol{c}^*$ is via solving the following optimization problem:

$$\mathrm{OPT}_\delta := \max_{\boldsymbol{c} \in \mathcal{S}^{n-1}} \sum_{t=1}^{T} I_{\mathcal{C}_t}(\boldsymbol{c})$$

where the indicator function $I_{\mathcal{E}}(e) = 1$ if $e \in \mathcal{E}$ and $I_{\mathcal{E}}(e) = 0$ otherwise. The rationale for the optimization problem is that for the t-th observation, a vector $\boldsymbol{c}$ is consistent with the observation, i.e., $\boldsymbol{x}_t^*$ is the optimal solution of LP$(\boldsymbol{c}, \boldsymbol{A}_t, \boldsymbol{b}_t)$, if and only if $I_{\mathcal{C}_t}(\boldsymbol{c}) = 1$. Thus the optimization problem essentially finds a vector $\boldsymbol{c}$ that is consistent with the maximal number of observations in $\mathcal{D}_T$. For a small δ, say $\delta < 0.5$, the true $\boldsymbol{c}^*$ will give a large objective value and should ideally stay close to the maximizer. From

an optimization perspective, the objective function is discontinuous in its argument $\boldsymbol{c}$, so we propose the simulated annealing algorithm – Algorithm 4 in the next section to solve for its optimal solution. In this subsection, we focus on the theoretical properties of the optimization problem.

We first build some connection between the optimization problem OPT_δ and that of the conventional setting with $\delta = 0$. Let $\bar{\boldsymbol{x}}_t^*$ be the optimal solution of $\mathrm{LP}(\boldsymbol{c}^*, \boldsymbol{A}_t, \boldsymbol{b}_t)$ and define

$$\bar{\mathcal{C}}_t \coloneqq \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \bar{\boldsymbol{x}}_t^* \text{ is an optimal solution of } \mathrm{LP}(\boldsymbol{c}^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\right\}.$$

Then the conventional setting of inverse optimization solves

$$\bar{\mathrm{OPT}}_\delta \coloneqq \max_{\boldsymbol{c} \in \mathcal{S}^{n-1}} \sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t}(\boldsymbol{c}).$$

By the setup of the problem, $\boldsymbol{c}^*$ is an optimizer of $\bar{\mathrm{OPT}}_\delta$ and the optimal objective value is T. The following proposition establishes that the optimization problem OPT_δ can be viewed as a contaminated version of $\bar{\mathrm{OPT}}_\delta$. And the effect that the contamination has on the objective function can be bounded using δ.

Proposition 3. *Under Assumption 1, the following inequality holds*

$$\mathbb{P}\left(\max_{\boldsymbol{c} \in \mathcal{S}^{n-1}} \left| \frac{1}{T}\sum_{t=1}^{T} I_{\mathcal{C}_t}(\boldsymbol{c}) - \frac{1}{T}\sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t}(\boldsymbol{c}) \right| \le \delta + \frac{\log T}{\sqrt{T}}\right) \le \frac{1}{T}.$$

As in the Gaussian setting, it may happen that the generation of $(\boldsymbol{A}_t, \boldsymbol{b}_t)$'s may prevent an exact recovery of $\boldsymbol{c}^*$. In other words, there might exist some vector $\boldsymbol{c}'$ that is indistinguishable from the true $\boldsymbol{c}^*$ based on the observations $\mathcal{D}_T$, i.e., $\boldsymbol{c}' \in \mathcal{C}_t$ if and only if $\boldsymbol{c}^* \in \mathcal{C}_t$ for all t. So, we aim to derive a bound on the predicted performance of an estimator $\hat{\boldsymbol{c}}$. Specifically, we define

$$\mathrm{Acc}(\hat{\boldsymbol{c}}) \coloneqq \mathbb{E}\left[I_{\mathcal{C}}(\hat{\boldsymbol{c}})\right] \text{ with } \mathcal{C} \coloneqq \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \boldsymbol{x}^* \text{ is an optimal solution of } \mathrm{LP}(\boldsymbol{c}_{\text{new}}, \boldsymbol{A}_{\text{new}}, \boldsymbol{b}_{\text{new}})\right\}$$

where $\hat{\boldsymbol{c}}$ is our estimator of $\boldsymbol{c}^*$, $(\boldsymbol{A}_{\text{new}}, \boldsymbol{b}_{\text{new}})$ is a new sample from the distribution $\mathcal{P}_{A,b}$, $\boldsymbol{c}_{\text{new}}$ is a new sample following the law of (4), and $\boldsymbol{x}^*$ is the optimal solution of the corresponding LP. That is, the quantity captures the probability that $\hat{\boldsymbol{c}}$ is consistent with a new (unseen) observation. As a benchmark, we know that for the true parameter satisfies $\mathrm{Acc}(\boldsymbol{c}^*) \ge 1 - \delta$.

The challenge for deriving a bound on $\mathrm{Acc}(\hat{\boldsymbol{c}})$ arises from the discontinuity of the objective function OPT_δ. The existing methods for deriving a statistical generalization bound largely rely on the continuity and the Lipschitzness of the loss function. Moreover, from Proposition 1, we know that $\mathcal{C}_t$ is specified by $(\boldsymbol{V}_t, \boldsymbol{u}_t)$ and the $\boldsymbol{V}_t$'s are of different dimensions for different t's. To overcome these challenges, we devise the following γ-margin objective function. Specifically, we first define a parameterized version of $\mathcal{C}_t$ by

$$\mathcal{C}_t(\gamma) \coloneqq \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \boldsymbol{V}_t\boldsymbol{c} \le \boldsymbol{u}_t - \gamma\boldsymbol{e}\right\}$$

where $\gamma \ge 0$ is a constant and $\boldsymbol{e}$ is an all-one vector. It is obvious that $\mathcal{C}_t(\gamma) \subset \mathcal{C}_t$. Accordingly, we define the γ-margin optimization problem by

$$\mathrm{OPT}_\delta(\gamma) \coloneqq \max_{\boldsymbol{c} \in \mathcal{S}^{n-1}} \sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}).$$

We introduce two groups of assumptions and then derive theoretical results on the optimizer of $\mathrm{OPT}_\delta(\gamma)$.

Assumption 3 (Nondegeneracy)**.** *We assume the following:*

- *(Uniqueness). The optimal solution* $\boldsymbol{x}^* = (x_1^*, ..., x_n^*)$ *of LP*$(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$ *is unique almost surely.*
- *(Primal nondegeneracy). The number of the nonzero entries of* $\boldsymbol{x}^*$ *equals the number of constraints almost surely, i.e.*

$$|B \coloneqq \{i : x_i^* > 0, i = 1, ..., n\}| = m.$$

  These nonzero entries are also called as basic variables.

- *(Non-singularity). Denote* $\boldsymbol{A}_B$ *as the submatrix of* $\boldsymbol{A}$ *corresponding to the index set* B*. The matrix* $\boldsymbol{A}_B$ *satisfies*

$$\|\boldsymbol{A}_B^{-1}\|_2 \leq \sigma. \tag{5}$$

The first two parts of the assumption require that the underlying linear program LP$(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$ is nondegenerate. This is a mild requirement in that any LP can satisfy the assumption with an arbitrarily small perturbation (Megiddo and Chandrasekaran, 1989). The third part of the assumption concerns the singularity of the submatrix of $\boldsymbol{A}$ that corresponds to the optimal basis. The intuition for this lower bound is that σ characterizes the instability of the optimal solution $\boldsymbol{x}^*$. Specifically, a larger σ indicates a stronger collinearity between the columns of $\boldsymbol{A}_B$, and consequently, the optimal solution $\boldsymbol{x}^*$ is less stable under small perturbation of the LP's input (including $\boldsymbol{c}$). Moreover, in the assumption, the inequality (5) is required to hold in an almost surely sense. We remark that this inequality can be relaxed to hold in a high probability sense, and accordingly, the remaining results in this subsection will only be affected by an additional small probability term.

Assumption 4 (L-stability)**.** *The following inequality holds for some* L*,*

$$\mathbb{P}\left(I_{\bar{\mathcal{C}}(\gamma)}(\boldsymbol{c}^*) \neq I_{\bar{\mathcal{C}}}(\boldsymbol{c}^*)\right) \leq L\gamma$$

where $\bar{\mathcal{C}} = \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \bar{\boldsymbol{x}}^* \text{ is an optimal solution of LP}(\boldsymbol{c}^*, \boldsymbol{A}, \boldsymbol{b})\right\}$ *and the probability space refers to the generation of* $(\boldsymbol{A}, \boldsymbol{b}) \sim \mathcal{P}_{A,b}$*.*

Assumption 4 requires the γ-margin function's value evaluate at $\boldsymbol{c}^*$ does not change much when γ is small. Specifically, we note that $I_{\bar{\mathcal{C}}}(\boldsymbol{c}^*) = 1$ almost surely, and the assumption states that with a probability of at least $1 - L\gamma$, the optimality condition in Proposition 1 is satisfied at least by a margin of γ. In Section XXX, we provide an intuitive interpretation of the assumption for the case of single constraint, which corresponds to the problem of revealed preference.

Proposition 4. *Under Assumption 1 and Assumption 3, the following inequality holds with probability no less than* $1 - \epsilon$*,*

$$\max_{\boldsymbol{c} \in \mathcal{S}^{n-1}} \frac{1}{T} \sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}) - \mathrm{Acc}(\boldsymbol{c}) \leq 4\sqrt{\frac{(1 + m\sigma^2)\log T}{\gamma^2 T}} + 4\sqrt{\frac{\log(T/\epsilon)}{T}}$$

for $\epsilon \in (0, 1)$*.*

Proposition 4 relates the generalization accuracy of any arbitrary $\boldsymbol{c}$ with the corresponding objective value of the γ-margin optimization problem. As γ increases, the objective function will decrease, so the right-hand-side becomes tighter. Importantly, the accuracy on the left-hand-side is defined by the original indicator function (or equivalently, $\mathcal{C}_t$), while the objective value (the first summation on the left-hand-side) is defined by the γ-margin indicator function (or equivalently, $\mathcal{C}_t(\gamma)$). Thus it justifies the

usage of the γ-margin optimization objective. The implication is that when we optimize the γ-margin objective, we can still obtain a bound on the original accuracy $\mathrm{Acc}(\boldsymbol{c})$ for a not-so-small γ.

Theorem 2. *Under Assumptions 1, 3 and 4, the following inequality holds*

$$\mathbb{P}\left(\mathrm{Acc}(\hat{\boldsymbol{c}}) \geq 1-\delta-\frac{(9+4L^{1/2}(1+m\sigma^2)^{1/4})\log T}{T^{1/4}}\right) \geq 1-\frac{5}{T},$$

where $\hat{\boldsymbol{c}}$ is one optimal solution of $OPT_\delta(\gamma)$ with $\gamma = \frac{2(1+m\sigma^2)^{1/4}}{L^{1/2}T^{1/4}}$.

Theorem 2 states a generalization bound on the accuracy of our estimator $\hat{\boldsymbol{c}}$. From Proposition 4, a larger γ leads to a smaller gap between the accuracy and the γ-margin objective function. Meanwhile, a smaller γ leads to a smaller gap between the optimal objective value $\mathrm{OPT}_\delta(\gamma)$ and $1-\delta$. In the extreme case of $\gamma = 0$, $\mathbb{E}[\mathrm{OPT}_\delta(0)] = \mathrm{Acc}(\boldsymbol{c}^*) \geq 1-\delta$. Theorem 2 optimizes the value of γ to trade off these two aspects. Assumption 3 covers the first aspect through Proposition 4, while Assumption 4 is the key to bound the second aspect. At first sight, the choice of γ may require the knowledge of parameters such as σ and L. We remark that knowledge of these parameters is not critical: one can simply choose $\gamma = (m/T)^{1/4}$ and this will only change the bound slightly in terms of its dependency on σ and L. Finally, we remark that the design of the γ-margin loss function is inspired from the max-margin classifier, but the analysis is entirely different. For the max-margin classifier, the introduction of the margin aims to make the underlying loss function 1-Lipschitz so that a generalization bound using Rademacher complexity can be derived. But for here, our γ-margin objective function is still a discontinuous one. Compared to the best achievable accuracy of $1-\delta$, the accuracy bound in Theorem 2 has a gap on the order of $\frac{1}{T^{1/4}}$. While many statistical estimation problems allow a better rate of $\frac{1}{T^{1/2}}$, we interpret the slower rate here as a price paid for the discontinuity of the objective function.

4.2 Online Setting and Suboptimality Loss

The methods proposed in the Gaussian setting and in the previous subsection all require some Markov chain Monte Carlo (MCMC)-type implementation. Computationally, these methods can be very inefficient when the dimension n is large (see next section for discussions); in that case, the MCMC algorithm essentially conducts a sampling procedure over a high-dimensional space. In this section, we consider an online setting with a different objective for the δ-corrupted setting, and the setting facilitates the development of more efficient algorithms to solve the problem. The results here also draw a connection with the existing works on the conventional setting of inverse optimization ($\delta = 0$) under the online setting (Bärmann et al., 2018; Chen and Kılınç-Karzan, 2020). Moreover, the results here do not rely on Assumption 3 and Assumption 4.

Specifically, we consider the following performance measure instead of $\mathrm{Acc}(\hat{\boldsymbol{c}})$,

$$l_{\mathrm{sub}}(\hat{\boldsymbol{c}}) \coloneqq \mathbb{E}\left[\hat{\boldsymbol{c}}^\top \hat{\boldsymbol{x}}^* - \hat{\boldsymbol{c}}^\top \boldsymbol{x}^*\right]$$

where $\hat{\boldsymbol{c}}$ is our estimator of $\boldsymbol{c}^*$, $(\boldsymbol{A}_{\mathrm{new}}, \boldsymbol{b}_{\mathrm{new}})$ is a new sample from the distribution $\mathcal{P}_{A,b}$, and $\boldsymbol{c}_{\mathrm{new}}$ is a new sample following the law of (4). Here $\hat{\boldsymbol{x}}^*$ and $\boldsymbol{x}^*$ are the optimal solutions of $\mathrm{LP}(\hat{\boldsymbol{c}}, \boldsymbol{A}_{\mathrm{new}}, \boldsymbol{b}_{\mathrm{new}})$ and $\mathrm{LP}(\boldsymbol{c}_{\mathrm{new}}, \boldsymbol{A}_{\mathrm{new}}, \boldsymbol{b}_{\mathrm{new}})$, respectively. The expectation is taken with respect to $(\boldsymbol{A}_{\mathrm{new}}, \boldsymbol{b}_{\mathrm{new}})$ and $\boldsymbol{c}_{\mathrm{new}}$. The performance measure l_{sub} is known as the suboptimality loss in the literature of inverse optimization (Mohajerin Esfahani et al., 2018; Bärmann et al., 2018; Chen and Kılınç-Karzan, 2020). The rationale is that when our estimator $\hat{\boldsymbol{c}}$ stays close to the realized (unobserved) $\boldsymbol{c}_{\mathrm{new}}$, $\hat{\boldsymbol{x}}^*$ and $\boldsymbol{x}^*$ should be the same to each other, which leads to a zero sub-optimality loss. Intuitively, one can also interpret the loss as a measure of explainability – how well the estimator explains the observed optimal solution $\boldsymbol{x}^*$. Also, the

measure can be viewed as a soft version of the binary-valued $Acc(\hat{\boldsymbol{c}})$ in that

$$I_{\mathcal{C}}(\hat{\boldsymbol{c}}) = 1 \;\Rightarrow\; \hat{\boldsymbol{c}}^\top \hat{\boldsymbol{x}}^* - \hat{\boldsymbol{c}}^\top \boldsymbol{x}^* = 0.$$

For the online setting, the observations in the dataset $\mathcal{D}_T$ arrive sequentially. The decision maker follows a policy/algorithm π. At each time t, the policy π makes a guess $\hat{\boldsymbol{c}}_t$ without using the observations of future samples (including the observation at time t). The performance of the policy is evaluated by the regret defined by

$$\mathrm{Reg}_T(\pi) = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\boldsymbol{c}}_t^\top \hat{\boldsymbol{x}}_t^* - \hat{\boldsymbol{c}}_t^\top \boldsymbol{x}_t^*\right) - \min_{\boldsymbol{c}\in\mathbb{R}^n} l_{\mathrm{sub}}(\boldsymbol{c})$$

where $\boldsymbol{x}_t^*$ is the observed optimal solution at time t and $\hat{\boldsymbol{x}}_t^*$ is the optimal solution of $\mathrm{LP}(\hat{\boldsymbol{c}}_t, \boldsymbol{A}_t, \boldsymbol{b}_t)$. We emphasize that this online setting may not fit all the application contexts of inverse optimization, but it provides us a generic framework for algorithm design and analysis.

Note that the regret depends on the magnitude of the observation $\boldsymbol{x}_t^*$ by its definition. Even if we can find a objective coefficient vector close to the true one, the regret may be amplified due to a large size of the feasible region. Thus, we introduce the following assumption where the probability space is with respect to the distribution $\mathcal{P}_{A,b}$. The bound is natural and the value of D can be pre-determined in many application scenarios.

Assumption 5. *There exists a constant D such that the following bound holds almost surely*

$$\|\boldsymbol{x}\|_2 \le D$$

for any $\boldsymbol{x}$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}$ where $(\boldsymbol{A}, \boldsymbol{b}) \sim \mathcal{P}_{A,b}$.

We first remark that the function $l_{\mathrm{sub}}(\cdot)$ is a convex function with respect to its argument. To see this, the function can be rewritten as

$$l_{\mathrm{sub}}(\boldsymbol{c}) = \mathbb{E}\left[\max_{\boldsymbol{x}\ge 0:\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}} \boldsymbol{c}^\top \boldsymbol{x} - \boldsymbol{c}^\top \boldsymbol{x}^*\right] \tag{6}$$

where the first term takes maximum over a set of linear functions and the second term is a linear function. This provides a foundation for the algorithm design in (Bärmann et al., 2018; Chen and Kılınç-Karzan, 2020). The results here can be viewed as an extension of theirs from a corruption-free setting of $\delta = 0$ to a δ-corrupted setting with $\delta > 0$.

As in the previous section, we first provide a benchmark as follows. It concerns the sub-optimality loss using the true $\boldsymbol{c}^*$. Its proof is straightforward and it is based on the generation mechanism of the objective coefficient vector (4). The right-hand-side of the bound reflects the inconsistency caused by the corruption, and it captures the proportion of the generated data samples that cannot be explained using the true $\boldsymbol{c}^*$.

Proposition 5. *Under Assumption 1 and Assumption 5, suppose that the feasible region of the LP (1) is bounded by D. The following inequality holds.*

$$l_{sub}(\boldsymbol{c}^*) \le 2D\delta.$$

With the benchmark in mind, we proceed to analyze two algorithms for the online setting. Both algorithms are inspired from standard algorithms for the problem of online convex optimization.

Our first algorithm – Algorithm 1 is described as follows. At each time t, the algorithm utilizes the past observations $\{(\boldsymbol{x}_s^*, \boldsymbol{A}_s, \boldsymbol{b}_s)\}_{s=1}^{t-1}$ to construct an empirical version of the suboptimality loss function

$l_{\text{sub}}(\cdot)$. It solves the optimization problem (7) to obtain the estimator $\hat{c}_t$ for each time step. The optimizer of (7) can be interpreted as a "leader" that has the best performance on the past samples (though some of which are corrupted).

Algorithm 1 Follow the Corrupted Leader

1: Input: dataset $\mathcal{D}_T = \{(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\}_{t=1}^T$
2: Let $\hat{\boldsymbol{c}}_1 = \mathbf{1}/\sqrt{n}$
3: **for** $t = 2, ..., T$ **do**
4: Compute

$$\hat{\boldsymbol{c}}_t = \arg\min_{\boldsymbol{c}\in\mathcal{S}^{n-1}} \sum_{s=1}^{t-1} \left(\max_{\boldsymbol{x}_s \geq \mathbf{0}: \boldsymbol{A}_s \boldsymbol{x}_s = \boldsymbol{b}_s} \boldsymbol{c}^\top \boldsymbol{x}_s \right) - \boldsymbol{c}^\top \boldsymbol{x}_s^* \tag{7}$$

5: **end for**
6: Output: $\hat{\boldsymbol{c}}_1, ..., \hat{\boldsymbol{c}}_T$

Theorem 3. *Under Assumption 1 and Assumption 5, the following bound holds for any $T \geq 3$,*

$$\text{Reg}_T(\pi_1) \leq 2D\delta + \frac{12D \log T}{\sqrt{T}} + \frac{3D}{T}$$

where π_1 represents the policy given by Algorithm 1.

Theorem 3 gives the regret bound of Algorithm 1. Its proof is based on a standard application of Rademacher complexity. We first note that the objective function (7) is convex, so the optimization problem can be solved much more efficiently than the problem of $\text{OPT}_\delta(\gamma)$. Second, compared to the benchmark in Proposition 5, the regret bound of Algorithm 1 has additional terms on the order of $O(1/\sqrt{T})$. The result shows the effectiveness of the "follow-the-leader" scheme and the unnecessity of introducing a margin loss as the case of $\text{OPT}_\delta(\gamma)$.

Algorithm 2 Online Corrupted Gradient Descent

1: Input: dataset $\mathcal{D}_T = \{(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\}_{t=1}^T$, a convex set $\mathcal{K}$
2: Let $\hat{\boldsymbol{c}}_1$ be an arbitrary point in $\mathcal{K}$
3: **for** $t = 1, ..., T$ **do**
4: Compute

$$\hat{\boldsymbol{x}}_t^* = \arg\max_{\boldsymbol{x}_t \geq \mathbf{0}: \boldsymbol{A}_t \boldsymbol{x}_t \leq \boldsymbol{b}_t} (\hat{\boldsymbol{c}}_t)^\top \boldsymbol{x}_t$$

5: Update

$$\hat{\boldsymbol{c}}_{t+1} = \text{Proj}_{\mathcal{K}} \left(\hat{\boldsymbol{c}}_t - \frac{1}{D\sqrt{T}} (\hat{\boldsymbol{x}}_t^* - \boldsymbol{x}_t^*) \right)$$

6: **end for**
7: Output: $\hat{\boldsymbol{c}}_1, ..., \hat{\boldsymbol{c}}_T$

Our second algorithm – Algorithm 2 implements the standard online gradient descent procedure to the problem. As we noted earlier, the loss function $l_{\text{sub}}(\cdot)$ is convex. At each time t, the algorithm computes the gradient $\nabla l_{\text{sub}}(\hat{\boldsymbol{c}}_t)$ and uses the gradient for an update in the space of objective coefficient vector. This algorithm has a shortcoming in that it requires the candidate set for the objective coefficient vector $\mathcal{K} = \{\boldsymbol{c}^*\} \cup \text{supp}\{\mathcal{P}_c'\}$ to be convex. This is the case (after some scaling) when the support of the corruption distribution $\text{supp}\{\mathcal{P}_c'\}$ stays within a half sphere that contains $\boldsymbol{c}^*$. In the algorithm, a projection step ensures $\boldsymbol{c}_t \in \mathcal{K}$ after the gradient update.

Theorem 4. *Suppose the set $\mathcal{K}$ is convex and bounded by 1 under the Euclidean norm. Under Assumption 1 and Assumption 5, the following bound holds for any $T \geq 3$,*

$$\mathrm{Reg}_T(\pi_2) \leq 2D\delta + \frac{10D\log T}{\sqrt{T}} + \frac{2D}{T^2}$$

where π_2 represents the policy given by Algorithm 2.

Theorem 4 gives the regret bound of Algorithm 2. Its proof adapts the standard analysis of online gradient descent to our case of corrupted observations. The result can also be viewed as a generalization of the analysis in (Bärmann et al., 2018) from the case of $\delta = 0$ to the case of $\delta > 0$. We remark that the regret bound here matches that of Theorem 3, while Algorithm 2 is computationally more efficient than Algorithm 1.

5 Numerical Experiments, Extensions, and Discussions

5.1 Algorithms for Posterior Sampling

Now we discuss the computational aspects of the proposed algorithms to complement the theoretical development in the previous sections. As mentioned earlier, the posterior sampling avoids the complication of optimizing over $\boldsymbol{\theta}$ in the maximum likelihood estimation, but still inevitably needs to deal with the sampling and numeric approximation of the likelihood function. Algorithm 3 describes a standard Metropolis–Hastings algorithm to sample from the posterior distribution $\mathbb{P}_T(\cdot)$. In the numerical experiments, we choose the proposal distribution $\mathbb{Q}$ to be a Gaussian random perturbation, i.e., $\boldsymbol{\theta}' = \mathrm{Proj}(\boldsymbol{\theta}^{(k-1)} + \boldsymbol{\epsilon})$ where $\boldsymbol{\epsilon}$ follows a Gaussian distribution and the projection ensures that $\boldsymbol{\theta}'$ stays on the sphere $\mathcal{S}^{n-1} \times [\underline{\kappa}, \bar{\kappa}]$. For the acceptance ratio, as the posterior distribution is not in closed form, a Monte Carlo subroutine is needed to estimate the ratio.

Algorithm 3 Posterior Sampling for the Gaussian Setting

1: Input: dataset $\mathcal{D}_T = \{(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\}_{t=1}^T$, number of iterations K
2: Initialize $\boldsymbol{\theta}^{(0)}$ by randomly sampling from the prior distribution $\mathbb{P}_0(\boldsymbol{\theta})$
3: **for** $k = 1, ..., K$ **do**
4: Draw a random $\boldsymbol{\theta}'$ from a pre-determined proposal distribution $\mathbb{Q}(\boldsymbol{\theta}'|\boldsymbol{\theta}^{(k-1)})$
5: Compute the acceptance rate:

$$r = \min\left\{\frac{\mathbb{P}_T(\boldsymbol{\theta}')}{\mathbb{P}_T(\boldsymbol{\theta}^{(k-1)})}, 1\right\}$$

6: Set

$$\boldsymbol{\theta}^{(k)} = \begin{cases} \boldsymbol{\theta}', & \text{w.p. } r \\ \boldsymbol{\theta}^{(k-1)}, & \text{w.p. } 1-r \end{cases}$$

7:
8: **end for**
9: Output: $\boldsymbol{\theta}^{(K)}$

Algorithm 4 presents a simulated annealing algorithm to solve the optimization problem $\mathrm{OPT}_\delta(\gamma)$ in Section 4.1. It takes a similar MCMC routine as in Algorithm 3 and we use the same Gaussian random perturbation for the proposal distribution $\mathbb{Q}$. As the temperature parameter η decreases, the sampling distribution in Algorithm 4 will gradually be more concentrated on the optimal solution set of $\mathrm{OPT}_\delta(\gamma)$. Algorithm 4 can be implemented more efficiently than Algorithm 3 in that the likelihood ratio calculation in (8) is analytical.

Algorithm 4 Simulated annealing algorithm for δ-corruption

Input: dataset $\mathcal{D}_T = \{(\boldsymbol{x}_t^*), \boldsymbol{A}_t, \boldsymbol{b}_t\}_{t=1}^T$, margin γ, number of iterations K, interval length τ
Initialize an initial (temperature) $\eta > 0$ and the reduction rate $\alpha \in (0, 1)$
Randomly generate the first estimate $\boldsymbol{c}^{(0)}$
for $k = 1, ..., K$ **do**
    if k mod $\tau = 0$ **then**
        Update $\eta \leftarrow \alpha \cdot \eta$
    end if
    Draw a proposal $\boldsymbol{c}'$ from a predetermined proposal distribution $\mathbb{Q}(\boldsymbol{c}'|\boldsymbol{c}^{(k-1)})$
    Compute the acceptance rate:

$$r = \min\left\{\exp\left\{\frac{1}{\eta}\cdot\left(\sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}') - \sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}^{(k-1)})\right)\right\}, 1\right\} \tag{8}$$

    Set

$$\boldsymbol{c}^{(k)} = \begin{cases} \boldsymbol{c}', & \text{w.p. } r \\ \boldsymbol{c}^{(k-1)}, & \text{w.p. } 1-r \end{cases}$$

end for
Output: $\boldsymbol{c}^{(K)}$

Numerical Experiments.

We first illustrate Algorithm 3 and Algorithm 4 for a single-constraint case due to the high computational complexity to compute the acceptance rate. Specifically, we consider the underlying LP that has the following form

$$\begin{aligned} &\max_{\boldsymbol{x}} \ \boldsymbol{c}^\top \boldsymbol{x} \\ &\text{s.t. } \boldsymbol{a}^\top \boldsymbol{x} \le b, \quad \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{1}. \end{aligned}$$

The reason why we consider the LP with inequality constraints is to guarantee feasibility. It can be easily transferred to the standard form by adding slack variables. Table 1 reports some numerical results for the two algorithms. For both the Gaussian and δ-corruption settings, we consider three distributions of $\mathcal{P}_{\boldsymbol{a},b}$: (i) a uniform distribution where $\boldsymbol{a} \sim$Unif$([1,2]^n)$ and $b \sim$Unif$([1,n])$; (ii) a discrete distribution where $\boldsymbol{a} \sim$ Unif$(\{1,2\}^n)$ and $b \sim$Unif$(1,...,n)$; (iii) a fixed-$\boldsymbol{a}$ distribution where $\boldsymbol{a} = (1,...,1)^\top$ and $b \sim$Unif$(1,...,n)$. For the Gaussian case, the true parameters $(\boldsymbol{\mu}^*, \kappa^*)$ are uniformly generated from $\mathcal{S}^{n-1} \times [1, 10]$, and the accuracy is calculated by $(\boldsymbol{\mu}^*\boldsymbol{x}^* - \boldsymbol{\mu}^*\tilde{\boldsymbol{x}}^*)/\boldsymbol{\mu}^*\boldsymbol{x}^*$ where $\boldsymbol{x}^*$ and $\tilde{\boldsymbol{x}}^*$ are defined in Corollary 2. For the δ-corruption case, $\boldsymbol{c}^*$ is uniformly generated from $\mathcal{S}^{n-1}$, δ is set to be 0.1, and the accuracy is calculated by $\text{Acc}(\hat{\boldsymbol{c}})/\text{Acc}(\boldsymbol{c}^*)$ where $\text{Acc}(\boldsymbol{c})$ is defined in Section 4. The numbers in Table 1 are reported based on an average of 20 simulation trials, and we run both Algorithm 3 and Algorithm 4 for $K = 1000$ iterations with $T = 1000$ samples.

We make the following observations from the numerical experiments. First, we remark that the theoretical results in the previous sections provide strong guarantees on the convergence property of the posterior distribution. So the deterioration of the algorithm performance for the case when $n = 25$ is solely caused by the inaccuracy of the approximate sampling in either Algorithm 3 or Algorithm 4. Such inaccuracy can definitely be mitigated to some extent by a more efficient algorithm implementation such as parallel computing. However, we argue that the performance deterioration as n grows may point to a curse of dimensionality that is intrinsic to this estimation problem. Essentially, we aim to estimate a high-dimensional distribution only through partial information, i.e., the sets $\mathcal{C}_t$'s. On the positive end, the algorithms work well for $n \le 10$, so if the decision maker has the power of choosing $(\boldsymbol{a}_t, b_t)$, s/he can

break up the high-dimensional estimation problem into a number of low-dimensional estimation problems by focusing on a handful of dimensions each time. Moreover, we provide a visualization of the condition (3) in Figure 3 for $n = 5$ calculated based on simulation. The visualization supports the existence of M and thus the identifiability of the true parameters when the posterior sampling can be accurately fulfilled.

For the multiple-constraint case, we only illustrate the performance of Algorithm 2 under the δ-corruption case. Here, we also consider the LP with inequality constraints as following

$$\max_{\boldsymbol{x}} \boldsymbol{c}^\top \boldsymbol{x}$$
$$\text{s.t. } \boldsymbol{A}^\top \boldsymbol{x} \leq \boldsymbol{b}, \quad \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{1}.$$

Moreover, we also consider three same distributions of $\mathcal{P}_{\boldsymbol{A},\boldsymbol{b}}$:(i) a uniform distribution where $\boldsymbol{A}$ $\sim$Unif$([1,2]^{m\times n})$ and $\boldsymbol{b}$ $\sim$Unif$([1,n]^m)$; (ii) a discrete distribution where $\boldsymbol{A} \sim$ Unif$(\{1,2\}^{m\times n})$ and $\boldsymbol{b}$ $\sim$Unif$(\{1,...,n\}^m)$; (iii) a fixed-$\boldsymbol{A}$ distribution where $\boldsymbol{A}$ is an all-one matrix and $\boldsymbol{b}$ $\sim$Unif$(\{1,...,n\}^m)$. The underlying true objective vector $\boldsymbol{c}^*$ is uniformly generated from the standard simplex $\{\boldsymbol{c} \geq \boldsymbol{0} : \boldsymbol{1}^\top \boldsymbol{c} = 1\}$, and δ is set to be 0.1. In Table 2, we report the regret defined in Section 4.2 and the accuracy calculated by Acc$(\hat{\boldsymbol{c}})$/Acc$(\boldsymbol{c}^*)$. The numbers in Table 2 are reported based on an average of 100 simulation trials with $m = 5$ constraints and $T = 1000$ samples. We observe that the regret is consistent with our theoretical results that the regret is in the order of $O(1/\sqrt{T})$. Moreover, the results also show that this gradient based method though with a good predictive performance, can hardly recover the decisions.

		$n=3$	$n=5$	$n=10$	$n=25$
Gaussian	(i)	99.9%	99.9%	98.8%	59.9%
	(ii)	99.9%	99.3%	96.9%	56.9%
	(iii)	99.9%	94.8%	92.6%	68.5%
δ-corru.	(i)	99.9%	96.7%	96.0%	55.7%
	(ii)	99.6%	97.9%	97.6%	63.1%
	(iii)	99.9%	98.7%	87.1%	58.7%

Table 1: Predictive Performance of Algorithm 3 and Algorithm 4

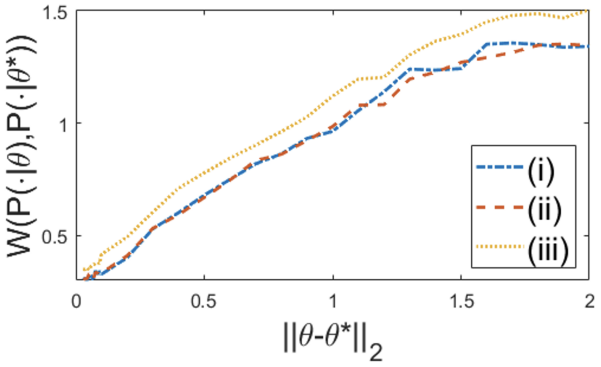

Figure 3: Visualization of (3).

5.2 Extensions to Other Settings

Noisy Observations

		$n=3$	$n=5$	$n=10$	$n=25$
Regret	(i)	2×10^{-4}	4×10^{-3}	1×10^{-2}	1×10^{-2}
	(ii)	2×10^{-3}	2×10^{-3}	1×10^{-2}	3×10^{-2}
	(iii)	2×10^{-4}	5×10^{-4}	2×10^{-3}	1×10^{-3}
Accuracy	(i)	71.1%	54.9%	26.0%	4.9%
	(ii)	78.6%	64.6%	30.1%	4.1%
	(iii)	82.5%	73.8%	55.0%	19.9%

Table 2: Predictive Regret and Accuracy by Algorithm 2

So far we have been focused on the setting where the optimal solutions are accurately observed. However, these observations can be noisy in a practical context which corresponds to the noisy setting of inverse optimization (Aswani et al., 2018). Specifically, we observe a contaminated or suboptimal version of the optimal solution:

$$\tilde{\boldsymbol{x}}_t = \boldsymbol{x}_t^* + \boldsymbol{\eta}_t$$

where $\boldsymbol{\eta}_t$ denotes the random noise. The setting is also known as the bounded rationality (Tversky and Kahneman, 1985). For our stochastic inverse optimization problem, the goal becomes to estimate the distribution of $\mathcal{P}_c$ from

$$\tilde{\mathcal{D}}_T \coloneqq \{(\tilde{\boldsymbol{x}}_t, \boldsymbol{A}_t, \boldsymbol{b}_t)\}_{t=1}^T.$$

For the Gaussian case, our analysis can be naturally extended to this noisy setting as the following theorem. Compared to Theorem 1, the new definition of the set $\tilde{\Theta}_T$ is based on the likelihood/distribution of the noisy observation. All the remaining statements and proofs of Theorem 5 are identical to Theorem 1. Specifically, the theorem says that the posterior distribution will also concentrate around the true underlying distribution given by the true parameter $\boldsymbol{\theta}^*$ even in this noisy case.

Theorem 5. *Let*

$$\tilde{\Theta}_T \coloneqq \left\{ \boldsymbol{\theta} \in \mathcal{S}^{n-1} \times (\underline{\kappa}, \bar{\kappa}) : D_{TV}\left(\mathbb{P}\left((\tilde{\boldsymbol{x}}_t, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}\right), \mathbb{P}\left((\tilde{\boldsymbol{x}}_t, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*\right)\right) \leq \max\left(8, 8\bar{\kappa}\right) \frac{\sqrt{n}\cdot\log T}{\sqrt{T}} \right\}$$

where $D_{TV}(\cdot,\cdot)$ is the total variation distance between two distributions supported on $\mathbb{R}^n \times \mathbb{R}_+^n \times \mathbb{R}_+$ equipped with Euclidean metric. Then, under Assumptions 1-2,

$$1 - \mathbb{P}_T(\tilde{\Theta}_T) \to 0 \text{ in probability as } T \to \infty.$$

Specifically, the following inequality holds

$$\mathbb{E}\left[\mathbb{P}_T\left(\tilde{\Theta}_T\right)\right] \geq 1 - \frac{3}{T}.$$

where the expectation is taken with respect to the random distribution $\mathbb{P}_T(\cdot)$ (essentially, with respect to the dataset $\tilde{\mathcal{D}}_T$).

For the δ-corruption case, we first note that as long as $\tilde{\boldsymbol{x}}_t$ and $\tilde{\boldsymbol{x}}_t$ share the same optimal basis, they correspond to the same set $\mathcal{C}_t$. For the case that $\tilde{\boldsymbol{x}}_t$ and $\tilde{\boldsymbol{x}}_t$ correspond to different optimal bases, we can view it as the part of the δ-corruption. In this way, the results in the δ-corruption case can also be applied to here.

Contextual Setting

Another interesting setting is that the decision maker observes not only the optimal action $\boldsymbol{x}_t^*$ and

the coefficient pair $(\boldsymbol{A}_t, \boldsymbol{b}_t)$, but also some context vector $\boldsymbol{z}_t$ for each time period. We remark that our posterior sampling result in Section 3 is also applicable to this contextual setting. Specifically, we can reconstruct the observation set as

$$\tilde{D}_T \coloneqq \{(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t, \boldsymbol{z}_t)\}_{t=1}^T.$$

Then, as in Section 3, we compute the posterior distribution $\mathbb{P}(\tilde{D}_T|\boldsymbol{\theta})$. Finally, we can either sample one parameter $\boldsymbol{\theta}_T$ from this posterior distribution if this distribution is concentrated, or find the parameter maximize posterior distribution if the distribution is computationally tractable.

5.3 Discussions

We conclude our discussion with the following remarks.

Query-based model with user-chosen $(\boldsymbol{A}_t, \boldsymbol{b}_t)$: In this paper, we have focused on the case where the constraints $(\boldsymbol{A}_t, \boldsymbol{b}_t)$'s are stochastically generated. When the concentration parameter κ is known for the Gaussian case, there is an efficient way of learning $\boldsymbol{\mu}$ through choosing $(\boldsymbol{A}_t, \boldsymbol{b}_t)$'s (See the Appendix D.2). In addition, the numerical experiments above also inspire a potential approach that dismantles the high-dimensional estimation problem into a number of low-dimensional problems. Another interesting and important question is whether there exist designs of $(\boldsymbol{A}_t, \boldsymbol{b}_t)$'s such that the posterior sampling can be more efficiently carried out.

Better algorithm design: We formulate the problem of stochastic inverse optimization and propose several algorithms to solve the problems in various settings. However, these algorithms all suffer from certain limitations. For example, the posterior sampling-based algorithms (Algorithm 3 and Algorithm 4) suffer from the curse of high-dimensionality, while the online gradient descent algorithm (Algorithm 2) requires the unknown objective coefficient vector to stay within a convex region. There might be some other algorithm design such as the expectation-maximization algorithm that deserves more future investigation.

Choice modeling: The stochastic utility model in our paper also draws an interesting connection with the literature on choice modeling, which is a pillar for the pricing and assortment problems in revenue management (Talluri et al., 2004; Gallego et al., 2019). For most of the existing choice models, the learning problem can be viewed as a special case of our study by letting $\boldsymbol{A} = (1, ..., 1)^\top$ and $b = 1$. The results in our paper complement this line of literature in developing a model where customers can make multiple purchases.

References

Afriat, Sydney N. 1967. The construction of utility functions from expenditure data. *International economic review* **8**(1) 67–77.

Ahmed, Shabbir, Yongpei Guan. 2005. The inverse optimal value problem. *Mathematical programming* **102**(1) 91–110.

Ahuja, Ravindra K, James B Orlin. 2001. Inverse optimization. *Operations Research* **49**(5) 771–783.

Amin, Kareem, Rachel Cummings, Lili Dworkin, Michael Kearns, Aaron Roth. 2015. Online learning and profit maximization from revealed preferences. *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 29.

Aswani, Anil, Zuo-Jun Shen, Auyon Siddiq. 2018. Inverse optimization with noisy data. *Operations Research* **66**(3) 870–892.

Balcan, Maria-Florina, Amit Daniely, Ruta Mehta, Ruth Urner, Vijay V Vazirani. 2014. Learning economic parameters from revealed preferences. *International Conference on Web and Internet Economics*. Springer, 338–353.

Baricz, Árpád. 2010. Bounds for modified bessel functions of the first and second kinds. *Proceedings of the Edinburgh Mathematical Society* **53**(3) 575–599.

Bärmann, Andreas, Alexander Martin, Sebastian Pokutta, Oskar Schneider. 2018. An online-learning approach to inverse optimization. *arXiv preprint arXiv:1810.12997* .

Beigman, Eyal, Rakesh Vohra. 2006. Learning from revealed preference. *Proceedings of the 7th ACM Conference on Electronic Commerce*. 36–42.

Bertsimas, Dimitris, Vishal Gupta, Ioannis Ch Paschalidis. 2012. Inverse optimization: A new perspective on the black-litterman model. *Operations research* **60**(6) 1389–1403.

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

Besbes, Omar, Yuri Fonseca, Ilan Lobel. 2021. Contextual inverse optimization: Offline and online learning. *Available at SSRN 3863366* .

Birge, John R, Ali Hortaçsu, J Michael Pavlin. 2017. Inverse optimization for the recovery of market structure from market outcomes: An application to the miso electricity market. *Operations Research* **65**(4) 837–855.

Boucheron, Stéphane, Gábor Lugosi, Pascal Massart. 2013. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press.

Burton, Didier, Ph L Toint. 1992. On an instance of the inverse shortest paths problem. *Mathematical programming* **53**(1) 45–61.

Burton, Didier, Ph L Toint. 1994. On the use of an inverse shortest paths algorithm for recovering linearly correlated costs. *Mathematical Programming* **63**(1) 1–22.

Chae, Minwoo, Pierpaolo De Blasi, Stephen G Walker. 2021. Posterior asymptotics in wasserstein metrics on the real line. *Electronic Journal of Statistics* **15**(2) 3635–3677.

Chan, Timothy CY, Tim Craig, Taewoo Lee, Michael B Sharpe. 2014. Generalized inverse multiobjective optimization with application to cancer therapy. *Operations Research* **62**(3) 680–695.

Chen, Violet Xinying, Fatma Kılınç-Karzan. 2020. Online convex optimization perspective for learning from dynamically revealed preferences. *arXiv preprint arXiv:2008.10460* .

Dong, Chaosheng, Yiran Chen, Bo Zeng. 2018a. Generalized inverse optimization through online learning. *Advances in Neural Information Processing Systems* **31**.

Dong, Chaosheng, Bo Zeng. 2020a. Expert learning through generalized inverse multiobjective optimization: Models, insights, and algorithms. *International Conference on Machine Learning*. PMLR, 2648–2657.

Dong, Chaosheng, Bo Zeng. 2020b. Inverse multiobjective optimization through online learning. *arXiv preprint arXiv:2010.06140* .

Dong, Jinshuo, Aaron Roth, Zachary Schutzman, Bo Waggoner, Zhiwei Steven Wu. 2018b. Strategic classification from revealed preferences. *Proceedings of the 2018 ACM Conference on Economics and Computation*. 55–70.

Elmachtoub, Adam N, Paul Grigas. 2022. Smart "predict, then optimize". *Management Science* **68**(1) 9–26.

Gallego, Guillermo, Huseyin Topaloglu, et al. 2019. *Revenue management and pricing analytics*, vol. 209. Springer.

Ghosal, Subhashis, Jayanta K Ghosh, Aad W Van Der Vaart. 2000. Convergence rates of posterior distributions. *Annals of Statistics* 500–531.

Gibbs, Alison L, Francis Edward Su. 2002. On choosing and bounding probability metrics. *International statistical review* **70**(3) 419–435.

Gray, Robert M. 2011. *Entropy and information theory*. Springer Science & Business Media.

Hazan, Elad. 2016. Introduction to online convex optimization. *Foundations and Trends in Optimization* **2**(3-4) 157–325.

Ho-Nguyen, Nam, Fatma Kılınç-Karzan. 2022. Risk guarantees for end-to-end prediction and optimization processes. *Management Science* .

Keshavarz, Arezou, Yang Wang, Stephen Boyd. 2011. Imputing a convex objective function. *2011 IEEE international symposium on intelligent control*. IEEE, 613–619.

Liese, Friedrich, Igor Vajda. 2006. On divergences and informations in statistics and information theory. *IEEE Transactions on Information Theory* **52**(10) 4394–4412.

Luenberger, David G, Yinyu Ye. 2021. *Linear and Nonlinear Programming*, vol. 228. Springer Nature.

McAllester, David A. 2003. Pac-bayesian stochastic model selection. *Machine Learning* **51**(1) 5–21.

Megiddo, Nimrod, Ramaswamy Chandrasekaran. 1989. On the ε-perturbation method for avoiding degeneracy. *Operations Research Letters* **8**(6) 305–308.

Mohajerin Esfahani, Peyman, Soroosh Shafieezadeh-Abadeh, Grani A Hanasusanto, Daniel Kuhn. 2018. Data-driven inverse optimization with imperfect information. *Mathematical Programming* **167**(1) 191–234.

Romanazzi, Mario. 2014. Discriminant analysis with high dimensional von mises-fisher distributions. *8th Annual International Conference on Statistics*. Athens Institute for Education and Research, 1–16.

Samuelson, Paul A. 1948. Consumption theory in terms of revealed preference. *Economica* **15**(60) 243–253.

Talluri, Kalyan T, Garrett Van Ryzin, Garrett Van Ryzin. 2004. *The theory and practice of revenue management*, vol. 1. Springer.

Tan, Yingcong, Daria Terekhov, Andrew Delong. 2020. Learning linear programs from optimal decisions. *Advances in Neural Information Processing Systems* **33** 19738–19749.

Tversky, Amos, Daniel Kahneman. 1985. The framing of decisions and the psychology of choice. *Behavioral decision making*. Springer, 25–41.

Van der Vaart, Aad W. 2000. *Asymptotic statistics*, vol. 3. Cambridge university press.

Wainwright, Martin J. 2019. *High-dimensional statistics: A non-asymptotic viewpoint*, vol. 48. Cambridge University Press.

Zadimoghaddam, Morteza, Aaron Roth. 2012. Efficiently learning from revealed preference. *International Workshop on Internet and Network Economics*. Springer, 114–127.

Zhang, Jianzhong, Zhenhong Liu. 1996. Calculating some inverse linear programming problems. *Journal of Computational and Applied Mathematics* **72**(2) 261–273.

A Auxiliary Lemmas

We present some preliminary lemmas in this section, most of which are basic inequalities from information theory.

Lemma 1 (Pinsker's inequality)**.** *For any two distributions $\mathcal{P}_1$ and $\mathcal{P}_2$,*

$$D_{TV}(\mathcal{P}_1, \mathcal{P}_2) \leq \sqrt{\frac{1}{2} D_{KL}(\mathcal{P}_1, \mathcal{P}_2)},$$

where $D_{TV}(\cdot, \cdot)$ and $D_{KL}(\cdot, \cdot)$ denote the total variation distance and the KL-divergence between two distributions, respectively.

Proof. We refer to Lemma 6.2 of Gray (2011). □

Lemma 2 (Data processing inequality)**.** *Let $\{K_\lambda\}_{\lambda \in \mathcal{X}}$ be a set of random variables indexed by parameter λ in some space $\mathcal{X}$. Consider two random variables Λ_1, Λ_2 taking values in $\mathcal{X}$. The following inequality holds*

$$D_{KL}(K_{\Lambda_1}, K_{\Lambda_2}) \leq D_{KL}(\Lambda_1, \Lambda_2),$$

where $D_{KL}(\cdot, \cdot)$ denotes the KL-divergence between two distributions. Here $\{K_\lambda\}_{\lambda \in \mathcal{X}}$ is usually called as a Markov kernel, a transition probability distribution, or a statistical kernel.

Proof. We refer to Theorem 14 of Liese and Vajda (2006). □

Lemma 3 (Packing number)**.** *Let $\mathcal{B}_r$ denote the ball in $\mathbb{R}^n$ centered at original point (or any point) with radius r. Then, the ϵ-packing number of $\mathcal{B}_r$ is upper bounded by*

$$\left(1 + \frac{2r}{\epsilon}\right)^n.$$

In other words, there exist at most $\left(1 + \frac{2r}{\epsilon}\right)^n$ disjoint balls with radius of $\frac{\epsilon}{2}$ in $\mathcal{B}_r$.

Proof. Assume that there are M disjoint $\epsilon/2$-balls. Then, the total volume of those M balls cannot be larger than the volume of a $r + \epsilon/2$ ball, i.e.,

$$M \cdot \left(\frac{\epsilon}{2}\right)^n \leq (r + \frac{\epsilon}{2})^n,$$

which implies

$$M \leq \left(1 + \frac{2r}{\epsilon}\right)^n.$$

□

Lemma 4 (Hoeffding's inequality). *Let $X_1, ..., X_T$ be independent random variables such that X_t takes its values in $[u_t, v_t]$ almost surely for all $t \leq T$. Then the following inequality holds*

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{n} X_t - \mathbb{E}X_t\right| \geq s\right) \leq 2\exp\left(-\frac{2T^2 s^2}{\sum_{i=1}^{n}(u_t - v_t)^2}\right)$$

for any $s > 0$.

Proof. We refer to Chapter 2 of Boucheron et al. (2013). □

Lemma 5 (Doob's consistency theorem). *Suppose that $\mathbb{P}(\cdot|\boldsymbol{\theta}) \neq \mathbb{P}(\cdot|\boldsymbol{\theta}')$ whenever $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$. Then, for every prior probability measure on the parameter space, the sequence of posterior measures converges to the point mass distribution of the true parameter in distribution for almost every $\boldsymbol{\theta}$.*

Proof. We refer to Chapter 10.10 of Van der Vaart (2000). □

Lemma 6. *Let $\mathcal{F}$ be a set of functions whose domain is the support of the distribution of $(\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})$. For any probability distribution $\tilde{\mathbb{P}}$ on $\mathcal{F}$, the following inequality holds for all $f \in \mathcal{F}$ and all distributions $\tilde{\mathbb{Q}}$ on $\mathcal{F}$ simultaneously*

$$\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\mathbb{E}[f(\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})]\right] \leq \mathbb{E}_{\tilde{\mathbb{Q}}}\left[\sum_{t=1}^{T} f(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\right] + \sqrt{\frac{D_{KL}(\tilde{\mathbb{P}}, \tilde{\mathbb{Q}}) + \log\frac{T}{\epsilon} + 2}{2T - 1}}.$$

with probability no less than $1 - \epsilon$. Here the inner expectation on the left-hand-side is taken with respect to $(\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})$.

Proof. We refer to Theorem 1 in McAllester (2003). □

The following lemma provides a useful bound for the modified Bessel function of the first kind. This function is closely related to the density of the von Mises–Fisher distribution. Typically, the modified Bessel function of the first kind is denoted by $I_\nu(x)$ with the parameter ν. In this paper, to distinguish between the indicator function and this modified Bessel function, we denote the modified Bessel function of the first kind by $\tilde{I}_\nu(x)$, which is defined as

$$\tilde{I}_\nu(x) := \sum_{t=0}^{\infty} \frac{1}{t\,!\,\Gamma(t + \nu + 1)}\left(\frac{x}{2}\right)^{2t+\nu},$$

where $\Gamma(\cdot)$ denotes the gamma function.

Lemma 7. *For all $0 < x < y$ and $\nu > 0$,*

$$e^{x-y}\left(\frac{x}{y}\right)^{\nu} \leq \frac{\tilde{I}_\nu(x)}{\tilde{I}_\nu(y)} \leq e^{y-x}\left(\frac{x}{y}\right)^{\nu},$$

where $\tilde{I}_\nu(\cdot)$ denotes the modified Bessel function of the first kind.

Proof. We refer to Chapter 2 of Baricz (2010). □

Lemma 8 (Rademacher Complexity). *Let $\mathcal{F}$ be a class of functions $f : \mathcal{X} \to [a, b]$, and $\{X_t\}_{t=1}^T$ be i.i.d. random variables taking values in $\mathcal{X}$. Then the following inequality holds for any $s > 0$*

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}}\left|\frac{1}{T}\sum_{t=1}^{T} f(X_t) - \mathbb{E}[f(X_1)]\right| \leq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\left|\frac{1}{T}\sum_{t=1}^{T}\sigma_t f(X_t)\right|\right] + s\right) \leq \exp\left(\frac{2Ts^2}{(b-a)^2}\right),$$

where $\{\sigma_t\}_{t=1}^T$ denotes a set of i.i.d. random signs satisfying $\mathbb{P}(\sigma_t = 1) = \mathbb{P}(\sigma_t = -1) = \frac{1}{2}$.

Proof. We refer to Theorem 4.10 in Wainwright (2019). □

Lemma 9 (Theorem 3.1 in Hazan (2016)). *Let $\{f_t(\boldsymbol{c})\}_{t=1}^T$ be a sequence of convex functions defined on a convex set $\mathcal{K}$. Suppose that $\mathcal{K}$ is contained in the unit ball with 2-norm, and $\|\nabla f_t(\boldsymbol{c})\|_2 \leq G$ for all $\boldsymbol{c} \in \mathcal{K}$ and all $t = 1, ..., T$. Let $\hat{\boldsymbol{c}}_{t+1} = \hat{\boldsymbol{c}}_t - \frac{2}{G\sqrt{t}}\nabla f_t(\hat{\boldsymbol{c}}_t)$. Then, the following inequality holds*

$$\sum_{t=1}^{T} f_t(\hat{\boldsymbol{c}}_t) - \min_{\boldsymbol{c}\in\mathcal{K}} \sum_{t=1}^{T} f_t(\boldsymbol{c}) \leq 3G\sqrt{T}.$$

Proof. We refer to Theorem 3.1 in Hazan (2016). □

A.1 Proof of Proposition 1

Proof. For an observation of $(\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})$, let

$$\mathcal{C} \coloneqq \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \boldsymbol{x}^* \text{ is an optimal solution of LP}(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})\right\}.$$

Denote $N = \{i : x_i^* = 0\}$ be the index set of zero entries of $\boldsymbol{x}^*$. The corresponding variables are called non-basic variables.

From the LP's optimality conditions, we know that $\boldsymbol{c} \in \mathcal{C}$ if and only if the following inequality holds:

$$\boldsymbol{r}_N \coloneqq \boldsymbol{c}_N - \boldsymbol{A}_N(\boldsymbol{A}_B)^{-1}\boldsymbol{c}_B \geq 0, \tag{9}$$

which can be expressed by linear constraints. Here, $\boldsymbol{c}_B$ denotes the vector consisting of entries of $\boldsymbol{c}$ corresponding to the index set B, $\boldsymbol{c}_B$ denotes the vector consisting of entries of $\boldsymbol{c}$ corresponding to the index set N, and $\boldsymbol{A}_N$ denotes the matrix consisting of columns of $\boldsymbol{A}$ corresponding to the index set N. For more detials about LP's optimality conditions, we refer to Section 4 in Luenberger and Ye (2021).

□

B Proof of Section 3

In this section, we provide the proofs of Section 3 and analyze the convergence of the posterior distribution. We remark that the total variation distance in Theorem 1 is not essential, and other metrics such as the Wasserstein distance, the Hellinger distance, and the Prokhorov metric are also valid (See Gibbs and Su (2002) for more about the relationships between different probability metrics). We first show several lemmas and then prove the main theorem.

B.1 Proof of Proposition 2

Proof. Consider the case that $n = 2$ and κ is fixed. In this case, we only have two observations $\mathcal{D}_2 = \{(\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)\}_{t=1}^2$. Let $\mathcal{C}_1$, $\mathcal{C}_2$ be the sets of all possible values of the objective vectors that is consistent with the two observations, respectively. Suppose

$$\mathcal{C}_1 = \left\{(\cos\alpha, \sin\alpha) : \alpha \in \left[-\frac{\pi}{4}, \frac{3\pi}{4}\right]\right\},$$
$$\mathcal{C}_2 = \left\{(\cos\alpha, \sin\alpha) : \alpha \in \left[-\frac{3\pi}{4}, \frac{\pi}{4}\right]\right\}$$

as shown in Figure 2.

Now, we show that for a fixed and large enough κ, $\boldsymbol{\mu}_1 = (1, 0)$ and $\boldsymbol{\mu}_2 = (-1, 0)$ are local minima of the negative likelihood function. To see this, we first compute the probability of the dataset $\mathcal{D}_2$ for any given $\boldsymbol{\mu} = (\cos\phi, \sin\phi)$ with fixed known κ,

$$\mathbb{P}\left(\mathcal{D}_2|(\boldsymbol{\mu}, \kappa)\right) = \frac{1}{4\pi^2 I_0^2(\kappa)} \int_{-\pi/4}^{3\pi/4} \exp(\kappa\cos(\psi + \phi))\mathrm{d}\psi \int_{-\pi/4}^{3\pi/4} \exp(\kappa\cos(\psi - \phi))\mathrm{d}\psi, \tag{10}$$

where $\phi \in [0, 2\pi)$ is the radian corresponding to the vector $\boldsymbol{\mu}$. The first order derivative of (10) with respect to ϕ is

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}\left(\mathcal{D}_2|(\boldsymbol{\mu}, \kappa)\right)}{\mathrm{d}\phi} =& \frac{1}{4\pi^2 I_0^2(\kappa)} \left[\left(\exp\left(\kappa\cos(-\pi/4 - \phi)\right) - \exp\left(\kappa\cos(3\pi/4 - \phi)\right)\right) \int_{-\pi/4}^{3\pi/4} \exp(\kappa\cos(\psi + \phi))\mathrm{d}\psi \right. \\ & \left. - \left(\exp\left(\kappa\cos(-\pi/4 + \phi)\right) - \exp\left(\kappa\cos(3\pi/4 + \phi)\right)\right) \int_{-\pi/4}^{3\pi/4} \exp(\kappa\cos(\psi - \phi))\mathrm{d}\psi \right],\end{aligned}$$

which implies $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are critical points. That is,

$$\left.\frac{\mathrm{d}\mathbb{P}\left(\mathcal{D}_2|(\boldsymbol{\mu}, \kappa)\right)}{\mathrm{d}\phi}\right|_{\phi=0} = \left.\frac{\mathrm{d}\mathbb{P}\left(\mathcal{D}_2|(\boldsymbol{\mu}, \kappa)\right)}{\mathrm{d}\phi}\right|_{\phi=\pi} = 0.$$

Similarly, we have that the second order derivatives at $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are

$$\begin{aligned}\left.\frac{\mathrm{d}^2\mathbb{P}\left(\mathcal{D}_2|(\boldsymbol{\mu}, \kappa)\right)}{\mathrm{d}\phi^2}\right|_{\phi=0} =& -\frac{1}{4\pi^2 I_0^2(\kappa)} \left[\sqrt{2}\left(\exp\left(\frac{\sqrt{2}\kappa}{2}\right) + \exp\left(-\frac{\sqrt{2}\kappa}{2}\right)\right) \int_{-\pi/4}^{3\pi/4} \exp(\kappa\cos(\psi))\mathrm{d}\psi \right. \\ & \left. + 2\left(\exp\left(\frac{\sqrt{2}\kappa}{2}\right) - \exp\left(-\frac{\sqrt{2}\kappa}{2}\right)\right)^2 \right],\end{aligned} \tag{11}$$

$$\begin{aligned}\left.\frac{\mathrm{d}^2\mathbb{P}\left(\mathcal{D}_2|(\boldsymbol{\mu}, \kappa)\right)}{\mathrm{d}\phi^2}\right|_{\phi=\pi} =& \frac{1}{4\pi^2 I_0^2(\kappa)} \left[\sqrt{2}\left(\exp\left(\frac{\sqrt{2}\kappa}{2}\right) + \exp\left(-\frac{\sqrt{2}\kappa}{2}\right)\right) \int_{3\pi/4}^{7\pi/4} \exp(\kappa\cos(\psi))\mathrm{d}\psi \right. \\ & \left. - 2\left(\exp\left(\frac{\sqrt{2}\kappa}{2}\right) - \exp\left(-\frac{\sqrt{2}\kappa}{2}\right)\right)^2 \right].\end{aligned} \tag{12}$$

Then, from (11), it is easy to see that $\left.\frac{\mathrm{d}^2\mathbb{P}(\mathcal{D}_2|(\boldsymbol{\mu},\kappa))}{\mathrm{d}\phi^2}\right|_{\phi=0} < 0$, which implies $\boldsymbol{\mu}_1 = (1, 0)$ is a minimizer of the negative likelihood function. Moreover, since the numerator of (12) satisfies

$$\lim_{\kappa\to\infty} \frac{\sqrt{2}\left(\exp\left(\frac{\sqrt{2}\kappa}{2}\right) + \exp\left(-\frac{\sqrt{2}\kappa}{2}\right)\right) \int_{3\pi/4}^{7\pi/4} \exp(\kappa\cos(\psi))\mathrm{d}\psi - 2\left(\exp\left(\frac{\sqrt{2}\kappa}{2}\right) - \exp\left(-\frac{\sqrt{2}\kappa}{2}\right)\right)^2}{\exp\left(\sqrt{2}\kappa\right)} = -2,$$

we have $\left.\frac{\mathrm{d}^2\mathbb{P}(\mathcal{D}_2|(\boldsymbol{\mu},\kappa))}{\mathrm{d}\phi^2}\right|_{\phi=\pi} < 0$ when κ is large enough. Thus, for large κ, we have $\boldsymbol{\mu}_2 = (-1, 0)$ is also a minimizer of the negative likelihood function.

Next, we show that the gap between the likelihood function value at $\boldsymbol{\mu}_1$ and that at $\boldsymbol{\mu}_2$ can be arbitrarily close to 1. It is sufficient to directly show that the gap between $\mathbb{P}(\mathcal{D}_2|(\boldsymbol{\mu}_1, \kappa))$ and $\mathbb{P}(\mathcal{D}_2|(\boldsymbol{\mu}_2, \kappa))$

goes to 1 as κ goes to infinity. To see this, we have

$$\begin{aligned}\mathbb{P}(\mathcal{D}_2|(\boldsymbol{\mu}_1,\kappa)) &= \left(\frac{1}{2\pi I_0(\kappa)}\int_{-\pi/4}^{3\pi/4}\exp\left(\kappa\cos\psi\right)\mathrm{d}\psi\right)^2 \\ &= \left(\frac{\int_{-\pi/4}^{3\pi/4}\exp\left(\kappa\cos\psi\right)\mathrm{d}\psi}{\int_{-\pi/4}^{3\pi/4}\exp\left(\kappa\cos\psi\right)\mathrm{d}\psi+\int_{3\pi/4}^{7\pi/4}\exp\left(\kappa\cos\psi\right)\mathrm{d}\psi}\right)^2 \\ &\to 1,\ \text{,as } \kappa\to\infty,\end{aligned}$$

where the first line comes from (10), the second line comes from the definition of $I_0(\kappa)$, and the last line comes from $\lim_{\kappa\to\infty}\frac{\int_{3\pi/4}^{7\pi/4}\exp(\kappa\cos\psi)\mathrm{d}\psi}{\int_{-\pi/4}^{3\pi/4}\exp(\kappa\cos\psi)\mathrm{d}\psi}=0$. Similarly, we have

$$\mathbb{P}(\mathcal{D}_2|(\boldsymbol{\mu}_1,\kappa))\to 0,\ \text{,as } \kappa\to\infty.$$

Thus we complete the proof.

□

B.2 Lemmas for the Proof of Theorem 1

Recall $\mathcal{D}_T=\{(\boldsymbol{x}_t^*,\boldsymbol{A}_t,\boldsymbol{b}_t)\}_{t=1}^T$ is the dataset, and $\mathbb{P}((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta})$ is the likelihood distribution under the parameter $\boldsymbol{\theta}$. In the following, we denote the parameter space $\mathcal{S}^{n-1}\times(\underline{\kappa},\bar{\kappa})$ by Θ, denote the total variation distance between $\mathbb{P}((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta}_1)$ and $\mathbb{P}((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta}_2)$ by $D_{TV}(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)$, where $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2\in\Theta$ are two parameters, and denote the ϵ-packing number of a parameter subset $\tilde{\Theta}\subset\Theta$ with a metric D_{metric} by $\mathcal{N}\left(\epsilon,\tilde{\Theta},D_{metric}\right)$. For example, if the underlying metric is the total variation metric, the corresponding ϵ-packing number is denoted by $\mathcal{N}\left(\epsilon,\tilde{\Theta},D_{TV}\right)$.

Lemma 10 (Theorem 7.1 in (Ghosal et al., 2000))**.** *Suppose the $\epsilon/2$-packing number of the parameter space Θ with the total variation distance is bounded by some constant C, i.e.,*

$$\mathcal{N}\left(\frac{\epsilon}{2},\Theta,D_{TV}\right)\le C.$$

Then, for $t,j\in\mathbb{N}$, there exist a function ϕ_t, which maps a dataset $\mathcal{D}_t$ with t samples to $[0,1]$, such that

$$\mathbb{E}_{\boldsymbol{\theta}^*}[\phi_t]\le\frac{C\exp(-2t\epsilon^2)}{1-\exp(-2t\epsilon^2)}, \tag{13}$$

$$\sup_{D_{TV}(\boldsymbol{\theta},\boldsymbol{\theta}^*)>j\epsilon}\mathbb{E}_{\boldsymbol{\theta}}[1-\phi_t]\le\exp(-2tj^2\epsilon^2), \tag{14}$$

where the expectation is taken with respect to the underlying dataset $\mathcal{D}_t$ under a distribution specified by the corresponding parameter in the subscript.

Proof. We refer to Theorem 7.1 of Ghosal et al. (2000). □

Intuitively, the lemma states that if the packing number is bounded, we can find a function ϕ_t such that its expectation is close to 0 when taken under a distribution with the true parameter, and it is close to 1 otherwise. In statistics, the function ϕ_t naturally serves as a test, and it distinguish the wrong parameters from the true parameters. In the proof of Theorem 1, we will see that this test function ϕ_t plays an important role in bounding the numerator of the posterior distribution.

To apply this lemma for the proof of Theorem 1, we first calculate a bound for the packing number of the parameter space to meet the lemma's condition. Here $D_{KL}(\cdot,\cdot)$ denotes the KL divergence between two distributions.

Lemma 11. *For any two parameters* $\boldsymbol{\theta}_1 = (\boldsymbol{\mu}_1, \kappa_1)$ *and* $\boldsymbol{\theta}_2 = (\boldsymbol{\mu}_2, \kappa_2)$ *in* $\Theta \coloneqq \mathcal{S}^{n-1} \times (\underline{\kappa}, \bar{\kappa})$,

$$D_{TV}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \le \sqrt{\frac{1}{2} D_{KL}(\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}_1), \mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}_2))} \le \sqrt{\|\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2\|_2}. \tag{15}$$

Furthermore, we have

$$\mathcal{N}(\epsilon, \Theta, D_{TV}) \le \left(1 + \frac{\max(2, 2\bar{\kappa})}{\epsilon^2}\right)^n.$$

Proof. The first inequality comes directly from Pinsker's inequality (Lemma 1). For the second inequality, from Lemma 2, we have

$$\begin{aligned}
& D_{KL}(\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}_1), \mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}_2)) \\
\le\ & D_{KL}(\mathcal{P}_c(\cdot|\boldsymbol{\theta}_1), \mathcal{P}_c(\cdot|\boldsymbol{\theta}_2)) \\
=\ & \int_{\boldsymbol{c} \in \mathcal{S}^{n-1}} C_n(\kappa_1) \exp(\kappa_1 \boldsymbol{\mu}_1^\top \boldsymbol{c}) \cdot \log\left(\frac{C_n(\kappa_1) \exp(\kappa_1 \boldsymbol{\mu}_1^\top \boldsymbol{c})}{C_n(\kappa_2) \exp(\kappa_2 \boldsymbol{\mu}_2^\top \boldsymbol{c})}\right) \\
=\ & \int_{\boldsymbol{c} \in \mathcal{S}^{n-1}} C_n(\kappa_1) \exp(\kappa_1 \boldsymbol{\mu}_1^\top \boldsymbol{c}) \cdot \left((\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2)^\top \boldsymbol{c} + \log\left(\frac{\kappa_1^{n/2-1} \tilde{I}_{n/2-1}(\kappa_2)}{\kappa_2^{n/2-1} \tilde{I}_{n/2-1}(\kappa_1)}\right)\right) \\
\le\ & \int_{\boldsymbol{c} \in \mathcal{S}^{n-1}} C_n(\kappa_1) \exp(\kappa_1 \boldsymbol{\mu}_1^\top \boldsymbol{c}) \cdot \left((\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2)^\top \boldsymbol{c} + |\kappa_1 - \kappa_2|\right) \\
\le\ & 2\|\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2\|_2,
\end{aligned}$$

where

$$C_n(\kappa) = \left(\int_{\boldsymbol{c} \in \mathcal{S}^{n-1}} \exp(\kappa \boldsymbol{\mu}^\top \boldsymbol{c})\right)^{-1} = \frac{\kappa^{n/2-1}}{(2\pi)^{n/2} \tilde{I}_{n/2-1}(\kappa)},$$

which is independent of the choice of $\boldsymbol{\mu} \in \mathcal{S}^{n-1}$, and $\tilde{I}_{n/2-1}(\cdot)$ denotes the modified Bessel functions of the first kind. Here, the first line comes from Lemma 2. The second line comes from the definition of the KL-divergence and the density function of the von Mises–Fisher distribution, and the third line comes from the definition of $C_n(\kappa)$. The fourth line comes from Lemma 7 that

$$\left|\log \frac{\kappa_1^{n/2-1} \tilde{I}_{n/2-1}(\kappa_2)}{\kappa_2^{n/2-1} \tilde{I}_{n/2-1}(\kappa_1)}\right| \le \left|\max(\log \mathrm{e}^{\kappa_1 - \kappa_2}, \log \mathrm{e}^{\kappa_2 - \kappa_1})\right| = |\kappa_1 - \kappa_2|,$$

and the last line comes from Cauchy inequality that

$$(\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2)^\top \boldsymbol{c} \le \|\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2\|_2 \|\boldsymbol{c}\|_2 = \|\kappa_1 \boldsymbol{\mu}_1 - \kappa_2 \boldsymbol{\mu}_2\|_2.$$

Therefore, the ϵ^2-ball centered at some $\tilde{\boldsymbol{\theta}}$ with the L_2-norm is a subset of the ϵ-ball centered at $\tilde{\boldsymbol{\theta}}$ with the total variation metric, i.e.,

$$\{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_2 \le \epsilon^2\} \subset \{\boldsymbol{\theta} : D_{TV}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) \le \epsilon\}.$$

Thus, the ϵ-packing number of Θ with the total variance metric is bounded by the ϵ^2-packing number of Θ with the L_2-norm, that is,

$$\mathcal{N}(\epsilon, \Theta, D_{TV}) \le \mathcal{N}(\epsilon^2, \Theta, \|\cdot\|_2).$$

Finally, from Lemma 3, the following inequality holds

$$\mathcal{N}(\epsilon, \Theta, D_{TV}) \leq \left(1 + \frac{\max(2, 2\bar{\kappa})}{\epsilon^2}\right)^n$$

for all $\epsilon < 1$. □

The above lemma establishes the condition of Lemma 10 under our context. The following lemma concerns a bound on the likelihood ratio. Specifically, we will use the test function ϕ_t in (13) and (14) to establish the concentration in Theorem 1. However, the underlying measure for (14) is with respect to the parameter $\boldsymbol{\theta}$ but not the true parameter $\boldsymbol{\theta}^*$. To utilize the function ϕ_t, we need to change the measure using the likelihood ratio, and the following lemma gives a bound for the likelihood ratio.

Lemma 12. *For any $\epsilon > 0$, the following inequality holds with probability no less than $1-\exp(-T\epsilon^2/4\bar{\kappa}^2)$,*

$$\int_\Theta \frac{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\mathbb{P}_0(\boldsymbol{\theta}) \geq \left(\frac{2\epsilon^2}{\bar{\kappa}}\right)^n \cdot \exp(-T\epsilon^2). \tag{16}$$

Proof. Let

$$\Theta_{KL}(\epsilon) \coloneqq \left\{\boldsymbol{\theta} \in \Theta : D_{KL}(\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*), \mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta})) \leq \epsilon^2\right\}.$$

By inequality (15) in Lemma 11, we have that $\Theta_{KL}(\epsilon)$ contains the $2\epsilon^2$-ball centered at $\boldsymbol{\theta}^*$ under the Euclidean norm and, therefore,

$$\mathbb{P}_0(\Theta_{KL}(\epsilon)) \geq \left(\frac{2\epsilon^2}{\bar{\kappa}}\right)^n. \tag{17}$$

Next, we show that with probability no less than $1 - \exp(-T\epsilon^2/4\bar{\kappa}^2)$,

$$\int_{\Theta_{KL}(\epsilon)} \frac{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\mathbb{P}_0(\boldsymbol{\theta}) \geq \exp\left(-T\epsilon^2\right) \cdot \mathbb{P}_0(\Theta_{KL}(\epsilon)). \tag{18}$$

Since the range of the density function of $\mathcal{P}_c(\cdot|\boldsymbol{\theta})$ is between $[e^{-\bar{\kappa}}, e^{\bar{\kappa}}]$ for all $\boldsymbol{\theta} \in \Theta$, we have that

$$\exp(-2\bar{\kappa}) \leq \frac{\mathrm{d}\mathcal{P}_c(\cdot|\boldsymbol{\theta})}{\mathrm{d}\mathcal{P}_c(\cdot|\boldsymbol{\theta}*)} \leq \exp(2\bar{\kappa}), \tag{19}$$

where $\frac{\mathrm{d}\mathcal{P}_c(\cdot|\boldsymbol{\theta})}{\mathrm{d}\mathcal{P}_c(\cdot|\boldsymbol{\theta}^*)}$ denotes the Radon–Nikodym derivative between $\mathcal{P}_c(\cdot|\boldsymbol{\theta})$ and $\mathcal{P}_c(\cdot|\boldsymbol{\theta}^*)$. For any set $\mathcal{D}$ in $\mathcal{X} \times \mathbb{R}^{m\times n} \times \mathbb{R}^m$, we have

$$\begin{aligned}
\int_{\mathcal{D}} \mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}) &= \int_{\boldsymbol{A},\boldsymbol{b}} \int_{\boldsymbol{c}\in\mathcal{S}^{n-1}} I_{\mathcal{D}}(\boldsymbol{x}^*(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}), \boldsymbol{A}, \boldsymbol{b})\mathrm{d}\mathcal{P}_c(\boldsymbol{c}|\boldsymbol{\theta})\mathrm{d}\mathcal{P}_{A,b}(\boldsymbol{A}, \boldsymbol{b}) \\
&= \int_{\boldsymbol{A},\boldsymbol{b}} \int_{\boldsymbol{c}\in\mathcal{S}^{n-1}} I_{\mathcal{D}}(\boldsymbol{x}^*(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}), \boldsymbol{A}, \boldsymbol{b})\mathrm{d}\mathcal{P}_c(\boldsymbol{c}|\boldsymbol{\theta})\mathrm{d}\mathcal{P}_{A,b}(\boldsymbol{A}, \boldsymbol{b}) \\
&= \int_{\boldsymbol{A},\boldsymbol{b}} \int_{\boldsymbol{c}\in\mathcal{S}^{n-1}} \frac{\mathrm{d}\mathcal{P}_c(\cdot|\boldsymbol{\theta})}{\mathrm{d}\mathcal{P}_c(\cdot|\boldsymbol{\theta}^*)} I_{\mathcal{D}}(\boldsymbol{x}^*(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}), \boldsymbol{A}, \boldsymbol{b})\mathrm{d}\mathcal{P}_c(\boldsymbol{c}|\boldsymbol{\theta}^*)\mathrm{d}\mathcal{P}_{A,b}(\boldsymbol{A}, \boldsymbol{b}) \\
&\leq \exp(2\bar{\kappa}) \int_{\boldsymbol{A},\boldsymbol{b}} \int_{\boldsymbol{c}\in\mathcal{S}^{n-1}} I_{\mathcal{D}}(\boldsymbol{x}^*(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}), \boldsymbol{A}, \boldsymbol{b})\mathrm{d}\mathcal{P}_c(\boldsymbol{c}|\boldsymbol{\theta}^*)\mathrm{d}\mathcal{P}_{A,b}(\boldsymbol{A}, \boldsymbol{b}) \\
&= \exp(2\bar{\kappa}) \int_{\mathcal{D}} \mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*),
\end{aligned}$$

where the fourth line comes from (19), and all others come from direct calculations. The inequality above implies

$$\log \frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*)} \leq 2\bar{\kappa}$$

holds almost surely for $\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*)$, where $\frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta}^*)}$ denotes the Radon–Nikodym derivative between those two probability distributions. Hence,

$$-2\bar{\kappa} \leq \log \frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*)} \leq 2\bar{\kappa}.$$

Here, we emphasize that this bound of the Radon–Nikodym derivative still holds even if the optimal solution $\boldsymbol{x}^*$ is not deterministically determined by the objective vector $\boldsymbol{c}$ and the coefficient pair $(\boldsymbol{A}, \boldsymbol{b})$, but by a random distribution conditional on the objective vector and the coefficient pair.

Then, by Hoeffding's inequality (Lemma 4), the following inequality holds with probability no less than $1 - \exp(-T\epsilon^2/4\bar{\kappa}^2)$,

$$\begin{aligned}
&\sum_{t=1}^{T} \int_{\Theta_{KL}(\epsilon)} \log \frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*)} \frac{\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta})}{\mathbb{P}_0(\Theta_{KL}(\epsilon))} \\
\geq & T\mathbb{E}_{\boldsymbol{\theta}^*} \left(\int_{\Theta_{KL}(\epsilon)} \log \frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*)} \frac{\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta})}{\mathbb{P}_0(\Theta_{KL}(\epsilon))} \right) - T\epsilon^2 \\
= & T \int_{\Theta_{KL}(\epsilon)} D_{KL}(\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*), \mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta})) \frac{\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta})}{\mathbb{P}_0(\Theta_{KL}(\epsilon))} - T\epsilon^2 \\
\geq & -T\epsilon^2
\end{aligned} \tag{20}$$

where the second line comes directly from Hoeffding's inequality, the third line comes from Fubini's theorem and the definition of the KL divergence, and the last line comes from the non-negativity of the KL divergence. Then, by Jensen's inequality,

$$\begin{aligned}
\log \left(\int_{\Theta_{KL}(\epsilon)} \frac{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)} \frac{\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta})}{\mathbb{P}_0(\Theta_{KL}(\epsilon))} \right) &= \log \left(\int_{\Theta_{KL}(\epsilon)} \prod_{t=1}^{T} \frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*)} \frac{\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta})}{\mathbb{P}_0(\Theta_{KL}(\epsilon))} \right) \\
&\geq \sum_{t=1}^{T} \int_{\Theta_{KL}(\epsilon)} \log \frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b}_t)|\boldsymbol{\theta}^*)} \frac{\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta})}{\mathbb{P}_0(\Theta_{KL}(\epsilon))},
\end{aligned} \tag{21}$$

We can prove (18) by combining (20) with (21).

Finally, with probability no less than $1 - \exp(-T\epsilon^2/4\bar{\kappa}^2)$,

$$\begin{aligned}
\int_{\Theta} \frac{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)} \mathrm{d}\mathbb{P}_0(\boldsymbol{\theta}) &\geq \int_{\Theta_{KL}(\epsilon)} \frac{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)} \mathrm{d}\mathbb{P}_0(\boldsymbol{\theta}) \\
&\geq \exp\left(-T\epsilon^2\right) \cdot \mathbb{P}_0(\Theta_{KL}(\epsilon)) \\
&\geq \left(\frac{2\epsilon^2}{\bar{\kappa}} \right)^n \cdot \exp(-T\epsilon^2),
\end{aligned}$$

where the first inequality is obtained by the non-negativity of the integrand, the second inequality comes from (18), and the last line comes from (17). □

B.3 Proof of Theorem 1

In this part, we combine three lemmas in the previous section and show Theorem 1.

Proof. Let

$$\epsilon_T = \max(4, 4\bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{\sqrt{T}}.$$

By Lemma 12, we have, with probability no less than $1-\exp(-T\epsilon^2/4\bar{\kappa}^2)$, the inequality

$$\int_\Theta \frac{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\mathrm{d}\mathbb{P}_0(\boldsymbol{\theta}) \geq \left(\frac{2\epsilon_T^2}{\bar{\kappa}}\right)^n \cdot \exp(-T\epsilon_T^2) \geq \exp(-2T\epsilon_T^2) \tag{22}$$

holds for all T satisfying $T \geq n\log T$. In (26), we will use this inequality to establish a lower bound for the denominator of the posterior distribution.

By Lemma 11, we have for all $T \geq 3$

$$\mathcal{N}\left(\frac{\epsilon_T}{2}, \Theta, D_{TV}\right) \leq \left(\frac{5T}{16n\cdot\log T}\right)^n \leq \exp(T\epsilon_T^2),$$

which gives an upper bound of the packing number and thus verifies the condition of Lemma 10. Then, by Lemma 10, for all $T \geq 4$, there exists a function ϕ_T mapping the data set $\mathcal{D}_T$ to $[0,1]$, which satisfies

$$\mathbb{E}_{\boldsymbol{\theta}^*}[\phi_T] \leq \frac{\exp(T\epsilon_T^2)\exp(-2T\epsilon_T^2)}{1-\exp(-2T\epsilon_T^2)} \leq 2\exp(-T\epsilon_T^2), \tag{23}$$

$$\sup_{D_{TV}(\boldsymbol{\theta},\boldsymbol{\theta}^*)>2\epsilon_T} \mathbb{E}_{\boldsymbol{\theta}}[1-\phi_T] \leq \exp(-4T\epsilon_T^2). \tag{24}$$

Recall

$$\Theta_T \coloneqq \left\{\boldsymbol{\theta}\in\Theta : D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}^*,\boldsymbol{A},\boldsymbol{b})|\boldsymbol{\theta}^*\right)\right) \leq \max\left(8, 8\bar{\kappa}\right)\frac{\sqrt{n}\cdot\log T}{\sqrt{T}}\right\}$$

We then have

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\theta}^*}\left[(1-\phi_T)\int_{\Theta_T^c}\prod_{t=1}^T\frac{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\mathbb{P}_0(\boldsymbol{\theta})\right] &= \int_{\Theta_T^c}\mathbb{E}_{\boldsymbol{\theta}^*}\left[(1-\phi_T)\prod_{t=1}^T\frac{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\right]\mathbb{P}_0(\boldsymbol{\theta}) \\
&= \int_{\Theta_T^c}\mathbb{E}_{\boldsymbol{\theta}}(1-\phi_T)\mathbb{P}_0(\boldsymbol{\theta}) \\
&\leq \exp(-4T\epsilon_T^2),
\end{aligned} \tag{25}$$

where the first line is obtained by Fubini's theorem, the second line is obtained directly by computing the inner integral, and the last line comes from the definition of Θ_T and inequality (24). Denote the low probability event corresponding to inequality (22) as $\mathcal{E}_T$. By combining (22) and (25), we have

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\theta}^*}\left[\mathbb{P}_T(\Theta_T^c)(1-\phi_T)I_{\mathcal{E}_T}\right] &= \mathbb{E}_{\boldsymbol{\theta}^*}\left[\frac{(1-\phi_T)I_{\mathcal{E}_T}\int_{\Theta_T^c}\frac{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\mathbb{P}(\boldsymbol{\theta})}{\int_\Theta\frac{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta})}{\mathbb{P}(\mathcal{D}_T|\boldsymbol{\theta}^*)}\mathbb{P}(\boldsymbol{\theta})}\right] \\
&\leq \exp(-4T\epsilon_T^2)\exp(2T\epsilon_T^2) = \exp(-2T\epsilon_T^2),
\end{aligned} \tag{26}$$

where the first equality comes from the definition of the posterior distribution, and the second line is obtained by plugging in (22) and (25).

Finally, we have

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\theta}^*}\left[\mathbb{P}_T(\Theta_T)\right] &\geq 1-\mathbb{E}_{\boldsymbol{\theta}^*}\left[(1-\phi_T)I_{\mathcal{E}_T}\mathbb{P}_T(\Theta_T^c)\right] - \mathbb{E}_{\boldsymbol{\theta}^*}\left[\phi_T\mathbb{P}_T(\Theta_T^c)\right] - \mathbb{E}_{\boldsymbol{\theta}^*}\left[I_{\mathcal{E}_T^c}\mathbb{P}_T(\Theta_T^c)\right] \\
&\geq 1-\mathbb{E}_{\boldsymbol{\theta}^*}\left[(1-\phi_T)I_{\mathcal{E}_T}\mathbb{P}_T(\Theta_T^c)\right] - \mathbb{E}_{\boldsymbol{\theta}^*}\left[\phi_T\right] - \mathbb{E}_{\boldsymbol{\theta}^*}\left[I_{\mathcal{E}_T^c}\right] \\
&\geq 1-2\exp(-T\epsilon_T^2) - \exp\left(\frac{T\epsilon_T^2}{4\bar{\kappa}^2}\right) \\
&\geq 1-\frac{3}{T},
\end{aligned} \tag{27}$$

where the first line comes from the fact that the posterior probability is bounded by 1, the second line comes from (23) and (26), and the last line comes from the definition of ϵ_T. The inequality above also indicates that $\mathbb{P}_T(\Theta_T)$ converges to 1 in L_1 norm, which implies the convergence in probability.

□

B.4 Proof of Corollary 1

Proof. The convergence of the posterior distribution to the point mass distribution can be directly obtained by Doob's consistency theorem (Lemma 5). Here, we only show the second part, i.e., the upper bound of the posterior expectation $\mathbb{E}_T[\|\kappa_T\boldsymbol{\mu}_T - \kappa^*\boldsymbol{\mu}^*\|_2]$.

From the proof of Theorem 1, we have that, there exists a high probability event $\tilde{\mathcal{E}}_T$ satisfying

$$\mathbb{P}_{\boldsymbol{\theta}^*}(\tilde{\mathcal{E}}_T) \geq 1 - \frac{1}{T}$$

such that

$$\mathbb{E}_{\boldsymbol{\theta}^*}\left[\mathbb{P}_T(\Theta_T^c) I_{\tilde{\mathcal{E}}_T}\right] \leq \frac{2}{T}.$$

Then, by Markov's inequality, we have for all $T \geq 4$,

$$\begin{aligned}\mathbb{P}_{\boldsymbol{\theta}^*}\left(\mathbb{P}_T(\Theta_T^c) > \frac{\sqrt{n}\log T}{2M\sqrt{T}}\right) &\leq \mathbb{P}_{\boldsymbol{\theta}^*}(\tilde{\mathcal{E}}_T^c) + \mathbb{P}_{\boldsymbol{\theta}^*}\left(\tilde{\mathcal{E}}_T \cap \left(\mathbb{P}_T(\Theta_T^c) > \frac{\sqrt{n}\log T}{2M\sqrt{T}}\right)\right) \\ &\leq \frac{1}{T} + \frac{4M}{\sqrt{n}\log T \cdot \sqrt{T}} \leq \frac{1+4M}{\sqrt{T}}\end{aligned} \tag{28}$$

where the first line comes from a decomposition of the event, the second line comes from Markov's inequality, and the last comes from $n \geq 1, T \geq 4$.

Next, from the condition (3),

$$D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b})|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}_t^*, \boldsymbol{A}_t, \boldsymbol{b})|\boldsymbol{\theta}^*\right)\right) \geq M \cdot \|\kappa\boldsymbol{\mu} - \kappa^*\boldsymbol{\mu}^*\|_2.$$

Consequently, for any $\boldsymbol{\theta} = (\boldsymbol{\mu}, \kappa) \in \Theta_T$, we have

$$\|\kappa\boldsymbol{\mu} - \kappa^*\boldsymbol{\mu}^*\|_2 \leq \max(8, 8\bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{M \cdot \sqrt{T}}. \tag{29}$$

Combining (28) and (29), we have with probability no less than $1 - \frac{4M+1}{\sqrt{T}}$

$$\begin{aligned}\mathbb{E}_T\left[\|\kappa_T\boldsymbol{\mu}_T - \kappa^*\boldsymbol{\mu}^*\|_2\right] &\leq \max(8, 8\bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{M \cdot \sqrt{T}} \cdot \mathbb{P}_T(\Theta_T) + 2\bar{\kappa} \cdot \mathbb{P}_T(\Theta_T^c) \\ &\leq \max(8, 8\bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{M \cdot \sqrt{T}} + \max(1, \bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{M \cdot \sqrt{T}} = \max(9, 9\bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{M \cdot \sqrt{T}}.\end{aligned}$$

Here, the first inequality is obtained by the fact the the maximum distance between two different parameters are bounded by $2\bar{\kappa}$, and the second inequality is obtained by (29). □

B.5 Proof of Corollary 2

Proof. Recall the definition of Θ_T

$$\Theta_T = \left\{\boldsymbol{\theta} \in \Theta : D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*\right)\right) \leq \max(8, 8\bar{\kappa}) \frac{\sqrt{n} \cdot \log T}{\sqrt{T}}\right\}.$$

In a similar way as (28), we obtain

$$\mathbb{P}_{\boldsymbol{\theta}^*}\left(\mathbb{P}_T(\Theta_T^c) > \frac{\sqrt{n}\log T}{2\sqrt{T}}\right) \leq \frac{5}{\sqrt{T}}.$$

Then we utilize the total variation distance to bound the difference between optimal solutions. From the integral representation of the total variation distance, we have, for any $\boldsymbol{\theta}$,

$$D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}\right), \mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*\right)\right) = \frac{1}{2}\int_{(\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})} \left|1 - \frac{\mathrm{d}\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*\right)}{\mathrm{d}\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}\right)}\right| \mathrm{d}\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}\right). \quad (30)$$

On the right hand side, the ratio $\frac{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta})}{\mathrm{d}\mathbb{P}((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*)}$ can be viewed as the probability ratio that the optimal solutions (corresponding to $\boldsymbol{\theta}^*$ and $\boldsymbol{\theta}$) coincide for a fixed pair of $(\boldsymbol{A}, \boldsymbol{b})$. Thus, the integration calculates the probability that the the optimal solution corresponding to $\boldsymbol{\theta}$ is different from the optimal solution corresponding to $\boldsymbol{\theta}^*$. Let $\boldsymbol{\theta}_T$ be a random parameter drawn from the posterior distribution. Denote $\tilde{\boldsymbol{x}}^*$ as the optimal solution corresponding to $\boldsymbol{\theta}_T$ given $(\boldsymbol{A}, \boldsymbol{b})$, and $\boldsymbol{x}^*$ as the optimal solution corresponding to $\boldsymbol{\theta}^*$ given $(\boldsymbol{A}, \boldsymbol{b})$. We have

$$\mathbb{P}_T(\boldsymbol{x}^* \neq \tilde{\boldsymbol{x}}^*) \leq 2D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}_T\right), \mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*\right)\right). \quad (31)$$

Thus, with probability no less than $1 - \frac{5}{\sqrt{T}}$, we have

$$\begin{aligned}
\mathbb{E}[\|\boldsymbol{x}^* - \tilde{\boldsymbol{x}}^*\|_2] &\leq \sqrt{n}\mathbb{P}_T(\boldsymbol{x}^* \neq \tilde{\boldsymbol{x}}^*) \\
&\leq 2\sqrt{n}D_{TV}\left(\mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}_T\right), \mathbb{P}\left((\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})|\boldsymbol{\theta}^*\right)\right), \\
&\leq \max\left(16, 16\bar{\kappa}\right)\frac{\sqrt{n}\cdot\log T}{\sqrt{T}}
\end{aligned}$$

where the first line comes from the fact that the maximum distance between any two solutions is bounded by $\sqrt{n}$, the second line comes from (31), and the last line comes from the definition of Θ_T.

□

C Proof of Section 4

In this section, we prove the results in Section 4.

C.1 Proof of Proposition 3

Proof. The proof is a direct application of Hoeffding's inequality. Denote X_t as the indicator function of $\boldsymbol{c}_t \neq \boldsymbol{c}^*$. By Hoeffding's inequality, we have

$$\mathbb{P}\left(\frac{1}{T}\sum_{t=1}^{T} X_t \leq \delta + \frac{\log T}{\sqrt{T}}\right) \leq \frac{1}{T}.$$

Then, it is sufficient to show that

$$\max_{\boldsymbol{c}\in\mathcal{S}^{n-1}}\left|\frac{1}{T}\sum_{t=1}^{T} I_{\mathcal{C}_t}(\boldsymbol{c}) - \frac{1}{T}\sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t}(\boldsymbol{c})\right| \leq \frac{1}{T}\sum_{t=1}^{T} X_t.$$

To see this, we have

$$\begin{aligned}\max_{\boldsymbol{c}\in\mathcal{S}^{n-1}}\left|\frac{1}{T}\sum_{t=1}^{T}I_{\mathcal{C}_t}(\boldsymbol{c})-\frac{1}{T}\sum_{t=1}^{T}I_{\bar{\mathcal{C}}_t}(\boldsymbol{c})\right| &\le \max_{\boldsymbol{c}\in\mathcal{S}^{n-1}}\left|\frac{1}{T}\sum_{\{t:X_t=1\}}I_{\mathcal{C}_t}(\boldsymbol{c})-\frac{1}{T}\sum_{\{t:X_t=1\}}I_{\bar{\mathcal{C}}_t}(\boldsymbol{c})\right| \\ &\le \max_{\boldsymbol{c}\in\mathcal{S}^{n-1}}\frac{1}{T}\sum_{\{t:X_t=1\}}\left|I_{\mathcal{C}_t}(\boldsymbol{c})-I_{\bar{\mathcal{C}}_t}(\boldsymbol{c})\right| \\ &\le \max_{\boldsymbol{c}\in\mathcal{S}^{n-1}}\frac{1}{T}\sum_{\{t:X_t=1\}}1 \\ &= \frac{1}{T}\sum_{t=1}^{T}X_t,\end{aligned}$$

where the first inequality is obtained by the fact that $I_{\mathcal{C}_t}(\boldsymbol{c}) = I_{\bar{\mathcal{C}}_t}(\boldsymbol{c})$ if $X_t = 0$, the second line comes from Jensen's inequality for the absolute value function, and the last two lines come directly from the property of indicator functions. □

C.2 Proof of Proposition 4

Proof. In this part, we will show a stronger statement that, for all $\gamma > 0$,

$$\max_{\boldsymbol{c}:\|\boldsymbol{c}\|_2\le 1} -\mathrm{Acc}(\boldsymbol{c})+\frac{1}{T}\sum_{t=1}^{T}I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}) \le 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}}+4\sqrt{\frac{\log(T/\epsilon)}{T}},$$

where the uniform bound holds for all $\boldsymbol{c}$ in the unit ball (instead of the unit sphere).

Specifically, we utilize Lemma 6 to prove the above inequality. For any $\boldsymbol{c}_0$ in the unit ball, let $\mathbb{Q} = \mathcal{N}(0, \tau^2\boldsymbol{I}_n)$ and $\mathbb{Q}_0 = \mathcal{N}(\boldsymbol{c}_0, \tau^2\boldsymbol{I}_n)$ be two normal distributions over the estimated parameter space, where $\boldsymbol{I}_n$ is the n-dimensional identity matrix and τ is a constant to be determined. We remark that our choice of $\mathbb{Q}_0$ and $\mathbb{Q}$ will not affect the distribution of $(\boldsymbol{c}_t, \boldsymbol{A}_t, \boldsymbol{b}_t)$, which depends on $\mathcal{P}_{A,b}$, $\mathcal{P}_c$, δ, and $\boldsymbol{c}^*$. Thus, by Lemma 6, the following inequality holds for all $\boldsymbol{c}_0$ in the unit ball and $\tau \ge 0$ with probability no less than $1-\epsilon$,

$$\mathbb{E}_{\mathbb{Q}_0}\left[\mathbb{E}\left[I_{(\mathcal{C}(\gamma))^c}(\boldsymbol{c})\right]\right] \le \frac{1}{T}\mathbb{E}_{\mathbb{Q}_0}\left[\sum_{t=1}^{T}I_{(\mathcal{C}_t(\gamma))^c}(\boldsymbol{c})\right]+\sqrt{\frac{D_{KL}(\mathbb{Q},\mathbb{Q}_0)+\log\frac{T}{\epsilon}+2}{2T-1}},$$

where $\mathcal{E}^c$ denotes the complement of a set $\mathcal{E}$. We note that on the left-hand-side, the inner expectation is taken with respect to the indicator function (essentially the set $\mathcal{C}(\gamma)$), and the outer expectation is taken with respect to $\boldsymbol{c} \sim \mathbb{Q}_0$. For the right-hand-side, the expectation is taken with respect to $\boldsymbol{c} \sim \mathbb{Q}_0$.

Consequently, we have

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_0}\left[\mathbb{E}\left[I_{(\mathcal{C}(\gamma))^c}(\boldsymbol{c})\right]\right] &\le \frac{1}{T}\mathbb{E}_{\mathbb{Q}_0}\left[\sum_{t=1}^{T}I_{(\mathcal{C}_t(\gamma))^c}(\boldsymbol{c})\right]+\sqrt{\frac{\frac{\|\boldsymbol{c}_0\|_2^2}{2\tau^2}+\log\frac{T}{\epsilon}+2}{2T-1}} \\ &\le \frac{1}{T}\mathbb{E}_{\mathbb{Q}_0}\left[\sum_{t=1}^{T}I_{(\mathcal{C}_t(\gamma))^c}(\boldsymbol{c})\right]+2\sqrt{\frac{\frac{\|\boldsymbol{c}_0\|_2^2}{2\tau^2}+\log\frac{T}{\epsilon}}{T}}.\end{aligned} \tag{32}$$

Here, the first line is obtained by calculating of the KL-divergence between two Gaussian distributions $\mathbb{Q}$ and $\mathbb{Q}_0$, and the second line is a further simplification of the second line.

Now we analyze the left-hand-side and show that

$$I_{(\mathcal{C}(0))^c}(\boldsymbol{c}_0) - \exp\left(-\frac{\gamma^2}{2\tau^2(1+m\sigma^2)}\right) \le \mathbb{E}_{\mathbb{Q}_0}\left[I_{(\mathcal{C}(\gamma))^c}(\boldsymbol{c}_0)\right] \le I_{(\mathcal{C}(2\gamma))^c}(\boldsymbol{c}_0) + \exp\left(-\frac{\gamma^2}{2\tau^2(1+m\sigma^2)}\right) \tag{33}$$

where τ is the standard deviation of the distributions $\mathbb{Q}$ and $\mathbb{Q}_0$, and σ comes from Assumption 3.

To see this, recall that N is the index set of non-basic variables as defined in the proof of Proposition 1. For a fixed $(\boldsymbol{x}^*, \boldsymbol{A}, \boldsymbol{b})$, if $I_{\mathcal{C}(0)^c}(\boldsymbol{c}) = 1$, at least on entry of (9) is violated. Thus, if (9) does not holds for $\boldsymbol{c}_0$ while $I_{(\mathcal{C}(\gamma))^c}(\boldsymbol{c}) = 0$, there exists at least one index $i \in N$ such that

$$(\boldsymbol{c}_0)_i - (\boldsymbol{A}_i)^\top (\boldsymbol{A}_B)^{-1}(\boldsymbol{c}_0)_B \ge 0,\ c_i - (\boldsymbol{A}_i)^\top (\boldsymbol{A}_B)^{-1}\boldsymbol{c}_B < -\gamma, \tag{34}$$

where $\boldsymbol{A}_i$ is the i-th column of $\boldsymbol{A}$, and $\boldsymbol{A}_B$ is the optimal basis as defined in the proof of Proposition 1. The corresponding probability of such an event is no less than (by taking difference of the two inequalities in (34))

$$\mathbb{P}\left((\boldsymbol{c}_0)_i - c_i - (\boldsymbol{A}_i)^\top (\boldsymbol{A}_B)^{-1}(\boldsymbol{c}_0 - \boldsymbol{c})_B \ge \gamma\right).$$

Note that $\boldsymbol{c}_0$ follows a normal distribution. The random variable in the left-hand side above is a mean-zero Gaussian random variable with variance of

$$\tau^2(1 + \boldsymbol{A}_i^\top (\boldsymbol{A}_B)^{-1}(\boldsymbol{A}_B^\top)^{-1}\boldsymbol{A}_i) \le \tau^2(1+m\sigma^2).$$

where the inequality comes from Assumption 3.

Thus, (34) holds with probability no more than $\exp\left(-\frac{\gamma^2}{2\tau^2(1+m\sigma^2)}\right)$, which gives the left part of (33). The right part follows the same analysis.

Thus, let $\tau^2 = \frac{\gamma^2}{2(1+m\sigma^2)\log T}$. From (33) and (32), we have, with probability no less than $1-\epsilon$, the following inequalities hold simultaneously for all $\boldsymbol{c}_0$ in the unit ball,

$$\begin{aligned}
\mathbb{E}[I_{\mathcal{C}^c}(\boldsymbol{c}_0)] &\le \frac{1}{T}\sum_{t=1}^{T} I_{(\mathcal{C}(2\gamma))^c}(\boldsymbol{c}_0) + 2\exp\left(-\frac{\gamma^2}{2\tau^2(1+m\sigma^2)}\right) + 2\sqrt{\frac{\frac{\|\boldsymbol{c}_0\|_2^2}{2\tau^2} + \log\frac{T}{\epsilon}}{T}} \\
&\le \frac{1}{T}\sum_{t=1}^{T} I_{(\mathcal{C}(2\gamma))^c}(\boldsymbol{c}_0) + \frac{2}{T} + 2\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T} + \frac{\log(T/\epsilon)}{T}} \\
&\le \frac{1}{T}\sum_{t=1}^{T} I_{(\mathcal{C}(2\gamma))^c}(\boldsymbol{c}_0) + 2\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} + 4\sqrt{\frac{\log(T/\epsilon)}{T}},
\end{aligned} \tag{35}$$

where the first inequality is obtained by plugging (33) into both sides of (32), the second line is obtained by plugging the value of τ into the inequality, and the last line is obtained by the convexity of the square root function. Finally, with probability no less than $1-\epsilon$

$$\begin{aligned}
\sup_{\{\boldsymbol{c}:\|\boldsymbol{c}\|_2\le 1\}} -\mathrm{Acc}(\boldsymbol{c}) + \frac{1}{T}\sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}) &= \sup_{\{\boldsymbol{c}:\|\boldsymbol{c}\|_2\le 1\}} \mathbb{E}[I_{\mathcal{C}^c}(\boldsymbol{c})] - \frac{1}{T}\sum_{t=1}^{T} I_{(\mathcal{C}_t(\gamma))^c}(\boldsymbol{c}) \\
&\le 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} + 4\sqrt{\frac{\log(T/\epsilon)}{T}},
\end{aligned}$$

where the first line comes from the fact that $I_{\mathcal{E}}(\boldsymbol{c}) = 1 - I_{\mathcal{E}^c}(\boldsymbol{c})$ holds for any point $\boldsymbol{c}$ and set $\mathcal{E}$, and the second line comes from (35) with replacing 2γ by γ. □

C.3 Proof of Theorem 2

Proof. First, let

$$\bar{\mathcal{C}} \coloneqq \left\{ \boldsymbol{c} \in \mathcal{S}^{n-1} : \bar{\boldsymbol{x}}^* \text{ is an optimal solution of LP}(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}) \right\},$$

where $\bar{\boldsymbol{x}}^*$ is an optimal solution of LP$(\boldsymbol{c}^*, \boldsymbol{A}, \boldsymbol{b})$. Similar to the usage of Hoeffding's inequality in the proof of Proposition 4, we have the following inequality holds under Assumption 4

$$\mathbb{P}\left(\left| \sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t(\gamma)}(\boldsymbol{c}^*) - \sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t}(\boldsymbol{c}^*) \right| > L\gamma + \frac{\log T}{\sqrt{T}} \right) \leq 2\exp\left(-\frac{T\log T}{T} \right) = \frac{2}{T}, \tag{36}$$

where the randomness in the above inequalities comes only from $(\boldsymbol{A}_t, \boldsymbol{b}_t)$ (or equivalently $\mathcal{C}_t$).

Then, we combine the above inequality and Proposition 4. Let $\hat{\boldsymbol{c}}$ be the optimal solution of the problem $OPT_\delta(\gamma)$. We have, for any γ, with probability no more than $1 - \epsilon - \frac{4}{T}$,

$$\begin{aligned}
\text{Acc}(\hat{\boldsymbol{c}}) &\geq \sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\hat{\boldsymbol{c}}) - 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} - 4\sqrt{\frac{\log(T/\epsilon)}{T}} \\
&\geq \sum_{t=1}^{T} I_{\mathcal{C}_t(\gamma)}(\boldsymbol{c}^*) - 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} - 4\sqrt{\frac{\log(T/\epsilon)}{T}} \\
&\geq \sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t(\gamma)}(\boldsymbol{c}^*) - \delta - 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} - 5\sqrt{\frac{\log(T/\epsilon)}{T}} \\
&\geq \sum_{t=1}^{T} I_{\bar{\mathcal{C}}_t}(\boldsymbol{c}^*) - \delta - L\gamma - 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} - 6\sqrt{\frac{\log(T/\epsilon)}{T}} \\
&\geq 1 - \delta - L\gamma - 4\sqrt{\frac{(1+m\sigma^2)\log T}{\gamma^2 T}} - 6\sqrt{\frac{\log(T/\epsilon)}{T}},
\end{aligned} \tag{37}$$

where the first inequality is obtained by Proposition 4, the second line is obtained by the optimality of $\hat{\boldsymbol{c}}$, the third line is obtained by a similar statement as Proposition 3, the fourth line comes from (36), and the last line is obtained by the fact that $I_{\bar{\mathcal{C}}_t}(\boldsymbol{c}^*) = 1$ for all t. Finally, we plug $\epsilon = \frac{1}{T}$ and $\gamma = \frac{2(1+m\sigma^2)^{1/4}}{L^{1/2}T^{1/4}}$ into (37). We have, with probability no less than $1 - \frac{5}{T}$,

$$\text{Acc}(\hat{\boldsymbol{c}}) \geq 1 - \delta - 4\sqrt{\frac{L(1+m\sigma^2)^{1/2}\log T}{T^{1/2}}} - 9\sqrt{\frac{\log T}{T}} \geq 1 - \delta - \frac{(9 + 4L^{1/2}(1+m\sigma^2)^{1/4})\log T}{T^{1/4}}.$$

□

C.4 Proof of Proposition 5

Proof. Denote the realized objective coefficient vector from law (4) as $\boldsymbol{c}$. Then if the realized utility vector $\boldsymbol{c} = \boldsymbol{c}^*$, we have

$$\left(\max_{\boldsymbol{x} \geq \boldsymbol{0}: \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}} (\boldsymbol{c}^*)^\top \boldsymbol{x} \right) - (\boldsymbol{c}^*)^\top \boldsymbol{x}^* = 0$$

where $\boldsymbol{x}^*$ is the optimal solution of LP$(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$.

Otherwise the optimality gap can still be bounded by $2D$ since

$$\begin{aligned}
\left| \left(\max_{\boldsymbol{x} \geq \boldsymbol{0}: \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}} (\boldsymbol{c}^*)^\top \boldsymbol{x} \right) - (\boldsymbol{c}^*)^\top \boldsymbol{x}^* \right| &\leq \|\boldsymbol{c}^*\|_2 \left\| \left(\arg \max_{\boldsymbol{x} \geq \boldsymbol{0}: \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}} (\boldsymbol{c}^*)^\top \boldsymbol{x} \right) - \boldsymbol{x}^* \right\|_2 \\
&\leq 2D,
\end{aligned} \tag{38}$$

where the first line is obtained by Cauchy's inequality, and the second line is obtained by the boundedness of the feasible region $\{\boldsymbol{x} \geq 0 : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$.

Thus, we have

$$\begin{aligned}\mathbb{E}\left[\left(\max_{\boldsymbol{x}\geq \boldsymbol{0}:\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}} (\boldsymbol{c}^*)^\top \boldsymbol{x}\right) - (\boldsymbol{c}^*)^\top \boldsymbol{x}^*\right] &\leq 0 \cdot \mathbb{P}(\boldsymbol{c} = \boldsymbol{c}^*) + 2D\mathbb{P}(\boldsymbol{c} \neq \boldsymbol{c}^*) \\ &\leq 2D\delta,\end{aligned}$$

where the first inequality is obtained by dividing the probability space into two events and the bound in (38), and the second line is obtained by $\mathbb{P}(\boldsymbol{c} \neq \boldsymbol{c}^*) \leq \delta$ □

C.5 Proof of Theorem 3

Proof. The idea of the proof is a straightforward application of using Rademacher complexity for generalization bound.

We first introduce the definition of Rademacher complexity. For any function class $\mathcal{F}$ that contains functions mapping the LP's inputs $(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$ to real numbers, the Rademacher complexity of the function class with t samples is defined by

$$R_t(\mathcal{F}) \coloneqq \frac{1}{t}\mathbb{E}\left[\sup_{f\in\mathcal{F}} \sum_{s=1}^{t} \sigma_s f(\boldsymbol{c}_s, \boldsymbol{A}_s, \boldsymbol{b}_s)\right],$$

where σ_s's are independent random signs distributed uniformly on $\{-1, 1\}$, for all $s = 1, ..., t$. Here, we define

$$\mathcal{F} = \left\{(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}) \to \left(\max_{\boldsymbol{x}\geq \boldsymbol{0}:\boldsymbol{A}\boldsymbol{x}\leq\boldsymbol{b}} \boldsymbol{c}^\top \boldsymbol{x}\right) - \boldsymbol{c}^\top \boldsymbol{x}^* : \|\boldsymbol{c}\|_2 = 1\right\},$$

where $\boldsymbol{x}^*$ is the optimal solution of $\mathrm{LP}(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$. Then, we have

$$\begin{aligned}R_t(\mathcal{F}) &= \frac{1}{t}\mathbb{E}\left[\sup_{\boldsymbol{c}:\|\boldsymbol{c}\|_2=1} \boldsymbol{c}^\top \left(\sum_{s=1}^{t} \sigma_s \left(\operatorname*{arg\,max}_{\boldsymbol{x}_s\geq \boldsymbol{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s} \boldsymbol{c}^\top \boldsymbol{x}_s - \boldsymbol{x}_s^*\right)\right)\right] \\ &\leq \frac{1}{t}\mathbb{E}\left[\sup_{\boldsymbol{c}:\|\boldsymbol{c}\|_2=1} \|\boldsymbol{c}\|_2 \left\|\sum_{s=1}^{t} \sigma_s \left(\operatorname*{arg\,max}_{\boldsymbol{x}_s\geq \boldsymbol{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s} \boldsymbol{c}^\top \boldsymbol{x}_s - \boldsymbol{x}_s^*\right)\right\|_2\right] \\ &\leq \frac{1}{t}\mathbb{E}\left\|\sum_{s=1}^{t} \sigma_s \left(\operatorname*{arg\,max}_{\boldsymbol{x}_s\geq \boldsymbol{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s} \boldsymbol{c}^\top \boldsymbol{x}_s - \boldsymbol{x}_s^*\right)\right\|_2 \\ &\leq \frac{1}{t}\sqrt{\mathbb{E}\left\|\sum_{s=1}^{t} \sigma_s \left(\operatorname*{arg\,max}_{\boldsymbol{x}_s\geq \boldsymbol{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s} \boldsymbol{c}^\top \boldsymbol{x}_s - \boldsymbol{x}_s^*\right)\right\|_2^2} \\ &= \frac{1}{t}\sqrt{\mathbb{E}\left(\sum_{s=1}^{t} \left\|\operatorname*{arg\,max}_{\boldsymbol{x}_s\geq \boldsymbol{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s} \boldsymbol{c}^\top \boldsymbol{x}_s - \boldsymbol{x}_s^*\right\|_2^2\right)} \\ &\leq \frac{2D}{\sqrt{t}},\end{aligned}$$

where the first line comes from the definition of $\mathcal{F}$, the second line comes from Cauchy's inequality, the third line comes from $\|\boldsymbol{c}\|_2 = 1$, the fourth line comes from Jensen's inequality, the fifth line comes from the elimination of cross terms since those random signs are independent, and the last line comes from the boundedness of the feasible region from Assumption 5. Similarly, from the boundedness of the feasible region, we have that,

$$0 \leq f(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}) \leq 2D, \text{ for all } f \in \mathcal{F}.$$

Next, by Lemma 8, we have with probability no less than $1-\frac{1}{T^2}$,

$$\sup_{f\in\mathcal{F}}\left|\sum_{s=1}^{t}\left(\left(\max_{\boldsymbol{x}_s\geq\mathbf{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s}\boldsymbol{c}^\top\boldsymbol{x}_s\right)-\boldsymbol{c}^\top\boldsymbol{x}_s^*\right)-\mathbb{E}\left[\left(\max_{\boldsymbol{x}\geq\mathbf{0}:\boldsymbol{A}\boldsymbol{x}\leq\boldsymbol{b}}\boldsymbol{c}^\top\boldsymbol{x}\right)-\boldsymbol{c}^\top\boldsymbol{x}^*\right]\right|\leq\frac{4D}{\sqrt{t}}+\frac{2D\log T}{\sqrt{t}}. \tag{39}$$

Recall $\hat{\boldsymbol{c}}_{t+1}$ is an optimal solution of (7) at time $t+1$. Denote $\hat{\boldsymbol{x}}_{t+1}$ as any arbitrary optimal solution of $\mathrm{LP}(\hat{\boldsymbol{c}}_{t+1},\boldsymbol{a}_{t+1},b_{t+1})$. Then, at each time $t=1,...,T-1$, we have with probability no less than $1-\frac{1}{T^2}$

$$\begin{aligned}\mathbb{E}\left(\hat{\boldsymbol{c}}_{t+1}^\top\hat{\boldsymbol{x}}_{t+1}-\hat{\boldsymbol{c}}_{t+1}^\top\boldsymbol{x}_{t+1}^*\right)&\leq\sum_{s=1}^{t}\left(\left(\max_{\boldsymbol{x}_s\geq\mathbf{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s}\hat{\boldsymbol{c}}_{t+1}^\top\boldsymbol{x}_s\right)-\hat{\boldsymbol{c}}_{t+1}^\top\boldsymbol{x}_s^*\right)+\frac{6D\log T}{\sqrt{t}}\\&\leq\sum_{s=1}^{t}\left(\left(\max_{\boldsymbol{x}_s\geq\mathbf{0}:\boldsymbol{A}_s\boldsymbol{x}_s\leq\boldsymbol{b}_s}(\boldsymbol{c}^*)^\top\boldsymbol{x}_s\right)-(\boldsymbol{c}^*)^\top\boldsymbol{x}_s^*\right)+\frac{6D\log T}{\sqrt{t}}\\&\leq\mathbb{E}\left[\left(\max_{\boldsymbol{x}\in\mathcal{X}}(\boldsymbol{c}^*)^\top\boldsymbol{x}\right)-(\boldsymbol{c}^*)^\top\boldsymbol{x}^*\right]+\frac{12D\log T}{\sqrt{t}}\\&\leq 2D\delta+\frac{12D\log T}{\sqrt{t}},\end{aligned} \tag{40}$$

where the first and third line comes from (39), the second line comes from the optimality of $\hat{\boldsymbol{c}}_{t+1}$, and the last line comes from Proposition 5. Thus, combining (40) and the boundedness of the feasible set, we have

$$\mathbb{E}\left(\hat{\boldsymbol{c}}_{t+1}^\top\hat{\boldsymbol{x}}_{t+1}-\hat{\boldsymbol{c}}_{t+1}^\top\boldsymbol{x}_{t+1}^*\right)\leq 2D\delta+\frac{12D\log T}{\sqrt{t}}+\frac{2D}{T^2}. \tag{41}$$

Finally, we take a summation with respect to both sides of (41) and obtain

$$\mathrm{Reg}_T(\pi_1)\leq\sum_{t=1}^{T}\mathbb{E}\left(\hat{\boldsymbol{c}}_t^\top\hat{\boldsymbol{x}}_t-\hat{\boldsymbol{c}}_t^\top\boldsymbol{x}_t^*\right)\leq 2D\delta+\frac{12D\log T}{\sqrt{T}}+\frac{3D}{T}.$$

□

C.6 Proof of Theorem 4

Proof. Let $f_t(\boldsymbol{c})=\boldsymbol{c}^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*)$ for $t=1,...,T$, which is a sequence of linear functions with respect to $\boldsymbol{c}$. By Assumption 5, we have

$$\|\nabla f_t\|_2=\|\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*\|_2\leq 2D$$

for all $\boldsymbol{c}$ with $\|\boldsymbol{c}\|_2\leq 1$. Then,

$$\begin{aligned}6D\sqrt{T}&\geq\sum_{t=1}^{n}\hat{\boldsymbol{c}}_t^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*)-\min_{\boldsymbol{c}\in\mathcal{K}}\sum_{t=1}^{n}\boldsymbol{c}^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*)\\&\geq\sum_{t=1}^{T}\hat{\boldsymbol{c}}_t^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*)-\sum_{t=1}^{T}(\boldsymbol{c}^*)^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*),\end{aligned} \tag{42}$$

where the first line comes from Lemma 9, and the second line comes from

$$\min_{\boldsymbol{c}\in\mathcal{K}}\sum_{t=1}^{n}\boldsymbol{c}^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*)\leq\sum_{t=1}^{n}(\boldsymbol{c}^*)^\top(\hat{\boldsymbol{x}}_t-\boldsymbol{x}_t^*).$$

Moreover, by Hoeffding's inequality, we have with probabiliy no less than $1-\frac{1}{T^2}$,

$$|\{t : \boldsymbol{c}_t \neq \boldsymbol{c}^*,\ t = 1, ..., T\}| \leq \delta T + 2\sqrt{T}\log T, \tag{43}$$

which implies

$$\begin{aligned}\sum_{t=1}^{T} \hat{\boldsymbol{c}}_t^\top (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t^*) &\leq \sum_{t=1}^{T} (\boldsymbol{c}^*)^\top (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t^*) + 6D\sqrt{T} \\ &\leq \sum_{t:\boldsymbol{c}_t \neq \boldsymbol{c}^*} \|\boldsymbol{c}^*\|_2 \|\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t^*\|_2 + 6D\sqrt{T} \\ &\leq 2D\,|\{t : \boldsymbol{c}_t \neq \boldsymbol{c}^*,\ t = 1, ..., T\}| + 6D\sqrt{T} \\ &\leq 2D\delta T + 4D\sqrt{T}\log T + 6D\sqrt{T},\end{aligned} \tag{44}$$

with probability no less than $1-\frac{1}{T^2}$. Here, the first line comes from (42), the second line comes from

$$(\boldsymbol{c}^*)^\top (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t^*) \leq 0, \text{ if } \boldsymbol{c}_t = \boldsymbol{c}^*,$$

the second line comes from Cauchy inequality, the third line comes from boundedness of $\boldsymbol{c}^*, \boldsymbol{c}_t, \boldsymbol{x}_t, \boldsymbol{x}_t^*$, and the last line comes from inequality (43).

Finally, taking expectation of (44), we have

$$\frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} \hat{\boldsymbol{c}}_t^\top (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t^*)\right] \leq 2D\delta + \frac{10D\log T}{\sqrt{T}} + \frac{2D}{T^2}.$$

□

D The Case of Single Constraint – Revealed Preference

D.1 Interpretation of Assumption 4

Here we provide an interpretation of the stability parameter L for the single-constraint case, which is also known as the revealed preference problem. Specifically, for the revealed preference problem, the decision maker can observe a sequence of actions of a consumer, and the goal is to recover the utility of the consumer. If the utility function of the customers is expressed by a linear function, then the observed action is the optimal solution of the following LP:

$$\begin{aligned}\max_{\boldsymbol{x}}\ & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t. } & \boldsymbol{a}^\top \boldsymbol{x} = b, \quad \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{1},\end{aligned}$$

where $\boldsymbol{c} = (c_1, ..., c_n) \in \mathbb{R}^n$ denotes the utility vector for n different products, $\boldsymbol{x} \in \mathbb{R}^n$ denotes the action, $\boldsymbol{a} = (a_1, ..., a_n) \in \mathbb{R}^n$ denotes the corresponding prices of those products, and $b \in \mathbb{R}$ denotes the consumer's budget. This LP is also a special case of the standard-form LP (1), and it encodes the decision-making process of the consumer, where the consumer wants to maximize the utility with the given constraint, and partial purchase is allowed. The following lemma say that if the price vector is positive and bounded from above by 1, i.e. $\boldsymbol{a} \in (0, 1]^n$, Assumption 4 holds with $L = \frac{4n^2}{\min_{i:c_i^* \neq 0} |c_i^*|}$.

Lemma 13. *Suppose the distribution of the coefficient pair $\mathcal{P}_{a,b}$ is a continuous distribution on $\boldsymbol{a} \in (0, 1]^n$, and its density is upper bounded by $\bar{p}$ almost surely. Then, with probability no less than* $1-$

$\frac{4n^2}{\min_{i:c_i^*\neq 0}|c_i^*|}\gamma$,

$$I_{\bar{\mathcal{C}}(\gamma)}(\boldsymbol{c}^*) = I_{\bar{\mathcal{C}}}(\boldsymbol{c}^*), \text{ for all } b,\ \gamma \leq \min_{i:c_i^*\neq 0}|c_i^*|, \text{ and fixed } \boldsymbol{c}^* \in \mathcal{S}^{n-1}.$$

Proof. For the proof, we assume that the optimal solution $\boldsymbol{x}_t^*$ satisfies

$$(\boldsymbol{x}_t^*)_i = 0 \text{ if } (\boldsymbol{c}_t)_i \leq 0 \text{ for all } t = 1, ..., T \text{ and } i = 1, ..., n.$$

This assumption is without loss of generality in that if purchasing the i-th item brings no positive utility, the customer will not purchase the item. Moreover, recall

$$\bar{\mathcal{C}} \coloneqq \left\{\boldsymbol{c} \in \mathcal{S}^{n-1} : \bar{\boldsymbol{x}}^* \text{ is an optimal solution of LP}(\boldsymbol{c}^*, \boldsymbol{a}, b)\right\},$$

where $\bar{\boldsymbol{x}}^*$ is an optimal solution of LP$(\boldsymbol{c}^*, \boldsymbol{a}, b)$. In this special LP, the optimality conditions can be simplified, and we have $\boldsymbol{c} \in \bar{\mathcal{C}}$ if and only if

$$\begin{aligned} -\min_{\{i:x_i^*>0\}}\left(\frac{c_i}{a_i}\right) &\leq 0, \\ \max_{\{i:x_i^*=0, c_i\neq 0\}}\frac{c_i}{a_i} &\leq 0, \text{if } \boldsymbol{a}^\top \boldsymbol{x}^* < b \\ \max_{\{i:x_i^*=0, c_i>0\}}\frac{c_i}{a_i} - \min_{\{i:x_i>0\}}\frac{c_i}{a_i} &\leq 0, \text{ if } \boldsymbol{a}^\top \boldsymbol{x}^* = b. \end{aligned}$$

Similarly, we define $\bar{\mathcal{C}}(\gamma)$ be the set of utility vectors $\boldsymbol{c}$ that satisfies the following conditions:

$$\begin{aligned} -\min_{\{i:x_i^*>0\}}\left(\frac{c_i}{a_i}\right) &\leq -\gamma, \\ \max_{\{i:x_i^*=0, c_i\neq 0\}}\frac{c_i}{a_i} &\leq -\gamma, \text{if } \boldsymbol{a}^\top \boldsymbol{x}^* < b \\ \max_{\{i:x_i^*=0, c_i>0\}}\frac{c_i}{a_i} - \min_{\{i:x_i>0\}}\frac{c_i}{a_i} &\leq -\gamma, \text{ if } \boldsymbol{a}^\top \boldsymbol{x}^* = b. \end{aligned}$$

Then, by definition of $\bar{\mathcal{C}}$ and the assumption $a_i \leq 1$ for all $i = 1, ..., n$, we have $\boldsymbol{c}^* \in \bar{\mathcal{C}}$ and

$$\begin{aligned} -\min_{\{i:\bar{x}_i^*>0\}}\left(\frac{c_i^*}{a_i}\right) &\leq -\min_{i:c_i^*\neq 0}|c_i^*|, \\ \max_{\{i:\bar{x}_i^*=0, u_i^*\neq 0\}}\left(\frac{c_i^*}{a_i}\right) &\leq -\min_{i:c_i^*\neq 0}|c_i^*|, \end{aligned}$$

for any $\boldsymbol{a}$, b and $\boldsymbol{c}^*$. Thus, $\boldsymbol{c}^* \in \bar{\mathcal{C}}(\gamma)$ if $\gamma \leq \min_{i:c_i^*\neq 0}|c_i^*|$ and $\left|\frac{c_j^*}{a_j} - \frac{c_i^*}{a_i}\right| \leq \gamma$ for all $i, j = 1, ..., n$. Moreover, for any $i, j = 1, ..., n$,

$$\left|\frac{c_j^*}{a_j} - \frac{c_i^*}{a_i}\right| \leq \gamma \Leftrightarrow \frac{c_i^* a_j^*}{c_j^*} - \frac{a_i a_j}{c_j^*}\gamma \leq a_i \leq \frac{c_i^* a_j^*}{c_j^*} + \frac{a_i a_j}{c_j^*}\gamma,$$

which happens with probability no more than $\frac{4\bar{p}\gamma}{\min_{i:c_i^*\neq 0}|u_i^*|}$. Since there are at most n^2 index pairs, with probability no less than $1 - \frac{4n^2\bar{p}\gamma}{\min_{i:u_i^*\neq 0}|u_i^*|}$, we have

$$I_{\bar{\mathcal{C}}(\gamma)}(\boldsymbol{c}^*) = I_{\bar{\mathcal{C}}}(\boldsymbol{c}^*), \text{ for all } b,\ \gamma \leq \min_{i:c_i^*\neq 0}|c_i^*|, \text{ and fixed } \boldsymbol{c}^* \in \mathcal{S}^{n-1}.$$

□

D.2 Gaussian setting with Known Concentration Parameter

In this section, we revisit the Gaussian setting for the revealed preference problem and discuss two methods to learn the mean vector $\boldsymbol{\mu} \in \mathcal{S}^{n-1}$ when the concentration parameter κ is known. Specifically, we consider the following two cases: (i) We can design the constraint pair $(\boldsymbol{a}_t, b_t)$ for all $t = 1, ..., T$, (ii) We can only design $\boldsymbol{a}_t \in [1, \infty)^n$ for all t while we know a lower bound $\underline{b}$ such that $b_t \geq \underline{b}$ for all t. The basic idea is to estimate $\mathbb{P}(c_i > 0|\boldsymbol{\mu})$ for every $i = 1, ..., n$ and to use the idea of "moment matching" to identify $\boldsymbol{\mu}$.

Constraint Design for the First Case

For the first case, we set $\boldsymbol{a}_t = (1, ..., 1) \in \mathbb{R}^n$ and $b_t = n$ for all $t = 1, ..., T$. Then, at each time t, we have $(\boldsymbol{x}_t^*)_i = 1$ if and only if $(\boldsymbol{c}_t)_i > 0$. We can then estimate $\mathbb{P}(c_i > 0|\boldsymbol{\mu})$ by the sample average mean

$$\hat{p}_i \coloneqq \frac{1}{T} \cdot \#\{t = 1, ..., T : (\boldsymbol{x}_t^*)_i = 1\}.$$

Next, we estimate $\boldsymbol{\mu}$ based on $\hat{p}_i$.

For any $i = 1, .., n$, let ϕ be an angle in $[0, \pi]$ satisfying $\cos\phi = \mu_i$. By Lemma 1 from Romanazzi (2014), we have

$$\mathbb{P}(c_i > 0|\boldsymbol{\mu}) = \int_0^{\pi/2} \left(\frac{\kappa}{2\pi}\right)^{1/2} \frac{\sin^{(n-1)/2}\psi \cdot \tilde{I}_{(n-3)/2}(\kappa \sin\phi \sin\psi)}{\sin^{(n-3)/2}\phi \cdot \tilde{I}_{n/2-1}(\kappa)} \cdot \exp\left(\kappa \cos\phi \cos\psi\right) \psi, \tag{45}$$

which depends only on μ_i. Thus, with slight abuse of notation, we denote $\mathbb{P}(c_i > 0|\mu_i)$ as the probability in (45). Moreover, if $n = 2$, we can have that the above function is strictly increasing. Then, by Lemma 1 from Romanazzi (2014) with the induction method, we can show that the above function is strictly increasing with respect to μ_i for every fixed n and κ. For any fixed n and κ, we first numerically compute (45). Then, a natural estimate of μ_i is

$$\hat{\mu}_i = \underset{\mu_i' \in [-1,1]}{\arg\min} |\mathbb{P}(c_i > 0|\mu_i') - \hat{p}_i|.$$

Then, $\hat{\boldsymbol{\mu}} = \{\hat{\mu}_1, ..., \hat{\mu}_n\}$ is our estimation of $\boldsymbol{\mu}$. We can further normalize $\hat{\boldsymbol{\mu}}$ in case that $\hat{\boldsymbol{\mu}} \notin \mathcal{S}^{n-1}$.

We remark that this method can be hardly generalized to the case that κ is unknown. The reason is that we do not have a similar strictly increasing structure, and that the bijection between the probability (45) and the parameter $\boldsymbol{\theta}$ might not exist.

Constraint Design for the Second Case

Now we discuss how to design $\boldsymbol{a}$ if we cannot control b. One difficulty that prevents us applying the same method as in the previous section is that we might have $(\boldsymbol{x}_t^*)_i = 0$ while $(\boldsymbol{c}_t)_i > 0$ due to an insufficient budget. However, if we know a lower bound of $\{b_t\}_{t=1}^T$, we can still estimate $\mathbb{P}(c_i > 0|\boldsymbol{\mu})$ by dismantling the high-dimensional estimation problem into a number of low-dimensional problems.

Denote the lower bound as $\underline{b}$. We set the first $\lceil \underline{b} \rceil$ entries in $\boldsymbol{a}$ be 1 and others be ∞, where $\lceil \cdot \rceil$ denotes the ceiling function. In this case, we have

$$(\boldsymbol{x}_t^*)_i > 0 \Leftrightarrow (\boldsymbol{c}_t)_i > 0 \text{ for all } i \leq \lceil \underline{b} \rceil \text{ and } t = 1, ..., T.$$

In this way, we can estimate $\mathbb{P}(u_i|\boldsymbol{\mu})$ for $i = 1, ..., \lceil \underline{b} \rceil$ following the previous case. To estimate the probability for other $i > \lceil \underline{b} \rceil$, we can divide the problem into $n/\underline{b}$ parts and estimate the probabilities separately.