On the Optimal Fixed-Price Mechanism in Bilateral Trade

Yang Cai*
Yale University, USA
yang.cai@yale.edu

Jinzhao Wu
Yale University, USA
jinzhao.wu@yale.edu

January 13, 2023

Abstract

We study the problem of social welfare maximization in bilateral trade, where two agents, a buyer and a seller, trade an indivisible item. The seminal result of Myerson and Satterthwaite [29] shows that no incentive compatible and *budget balanced* (i.e., the mechanism does not run a deficit) mechanism can achieve the optimal social welfare in bilateral trade. Motivated by this impossibility result, we focus on approximating the optimal social welfare. We consider arguably the simplest form of mechanisms – the fixed-price mechanisms, where the designer offers trade at a fixed price to the seller and buyer. Besides the simple form, fixed-price mechanisms are also the only *dominant strategy incentive compatible* and *budget balanced* mechanisms in bilateral trade [23].

We obtain improved approximation ratios of fixed-price mechanisms in both (i) the full information setting, where the designer knows the value distributions of both the seller and buyer; and (ii) the limited information settings. In the full information setting, we show that the optimal fixed-price mechanism can achieve at least 0.72 of the optimal welfare, and no fixed-price mechanism can achieve more than 0.7381 of the optimal welfare. Prior to our result the state of the art approximation ratio was $1-\frac{1}{e}+0.0001 \approx 0.632$ [24]. We further consider two limited information settings. In the first one, the designer is only given the mean of the buyer's value (or the mean of the seller's value). We show that with such minimal information, one can already design a fixed-price mechanism that achieves 0.65 of the optimal social welfare, which surpasses the previous state of the art ratio in the full information setting. In the second limited information setting, we assume that the designer has access to finitely many samples from the value distributions. Recent results show that one can already obtain a constant factor approximation to the optimal welfare using a single sample from the seller's distribution [3, 16, 24]. Our goal is to understand what approximation ratios are possible if the designer has more than one but still finitely many samples. This is usually a technically more challenging regime and requires tools different from the single-sample analysis. We propose a new family of sample-based fixed-price mechanisms that we refer to as the *order statistic mechanisms* and provide a complete characterization of their approximation ratios for any fixed number of samples. Using the characterization, we provide the optimal approximation ratios obtainable by order statistic mechanism for small sample sizes (no more than 10 samples) and observe that they significantly outperform the single sample mechanism.

*Yang Cai is supported by a Sloan Foundation Research Fellowship and the National Science Foundation Award CCF-1942583 (CAREER). Part of this work was done while the author was visiting the Simons Institute for the Theory of Computing.

1 Introduction

We study a fundamental problem in mechanism design – maximizing social welfare in bilateral trade, in which two agents, a seller and a buyer, trade an indivisible item. More specifically, we consider the Bayesian setting where the seller's private value S for the item that is drawn from distribution F_S, and the buyer's private value B for the item is drawn from distribution F_B. The social welfare is therefore defined as $\mathbb{E}_{B,S}[S+(B-S)\cdot x(B,S)]$, where $x(B,S)$ denotes the probability that the trade happens when the seller's value is S and the buyer's value is B.

Surprisingly, exactly maximizing the social welfare in bilateral trade is impossible. The seminal result by Myerson and Satterthwaite [29] shows that no mechanism can *simultaneously* be (i) incentive compatible (to the buyer and the seller), (ii) *budget balanced*, i.e., the mechanism does not run a deficit, and (iii) maximizes the social welfare. For example, the VCG mechanism is incentive compatible and maximizes the social welfare but is not budget balanced in general. Motivated by this impossibility result, our goal is to design incentive compatible and budget balanced mechanisms to approximate the optimal welfare. We focus on the fixed-price mechanisms, in which the designer offers trade at a fixed price to the seller and buyer. It is also known that fixed-price mechanisms are the only *dominant strategy* incentive compatible and budge balance mechanisms in bilateral trade [23].

1.1 Our Contributions

We make progress on this problem on multiple fronts. We first consider the *full information* setting, where the designer knows both F_S and F_B. We show how to use a *factor revealing min-max program* to improve the approximation ratio of achievable by a fixed-price mechanism.

Contribution 1: For any F_S, F_B, there exists a fixed-price mechanism whose welfare is at least $0.72\cdot\mathsf{OPT}$, where $\mathsf{OPT}=\mathbb{E}_{S,B}[\max\{S,B\}]$ is the optimal welfare. Moreover, there exists a F_S and F_B such that no fixed-price mechanism can attain welfare more than $0.7381\cdot\mathsf{OPT}$. The formal statement of our result can be found in Theorem 3.1.

We also have a "constant time" algorithm for computing the fixed-price mechanism that achieves the welfare guarantee above. More specifically, we construct a collection of numbers $p_1,\cdots,p_n$ in $[0,1]$, so that for any F_S, F_B, our algorithm chooses the best price in the set $\{p_1\cdot\mathsf{OPT},\cdots,p_n\cdot\mathsf{OPT}\}$. Clearly, the approximation ratio will be better when we increase n. We show that when $n=16$, our algorithm already computes a fixed-price mechanism that has welfare at least $0.72\cdot\mathsf{OPT}$.

Our result significantly improves on the state-of-the-art approximation $1-\frac{1}{e}+0.0001\approx 0.6322$ [24]. Our new hardness result also strengthens the previous best bound of 0.7385 [25]. Our upper and lower bounds are obtained by considering two discretized variants of an infinite-dimensional min-max optimization problem defined in Section 3.1. We show in Lemma 3.4 that, in the limit when the discretization accuracy approaches 0, the upper bound and lower bound obtainable by our method will converge to the optimal approximation ratio. Of course, the factor-revealing program become more expensive to solve with finer discretization. Our upper and lower bounds are derived using the finest discretization that we can computationally solve, but one could further close the gap with more computational resources.

Fixed-price mechanism based on only $\mathbb{E}[B]$ or $\mathbb{E}[S]$. Our first result requires the designer to know both F_S and F_B.[1] However, information about the underlying distributions of the agents' values is often scarce in practice, thus it is more desirable to design approximately optimal mechanisms using limited prior information. Our second contribution concerns the case where the designer does not have the full information of the underlying distributions but only knows the mean of F_S or F_B.

[1] Our first result uses F_S and F_B in two places: (1) to compute OPT and (2) to identify the best price in the set.

Contribution 2: Given access to $\mathbb{E}[B]$ or $\mathbb{E}[S]$, we can design a *randomized* fixed-price mechanism whose welfare is at least $0.65 \cdot \mathsf{OPT}$. See Theorem 4.1 for details.

Note that the ratio of 0.65 already exceeds the previous state-of-the-art approximation ratio in the full information setting. [5, 24] consider the setting where only F_S is known to the designer and show that a *quantile mechanism* (Mechanism 1), i.e., a fixed-price mechanism that chooses the trading price according to a distribution of quantiles of the seller's distribution, can obtain at least $1 - \frac{1}{e} \approx 0.6321$ fraction of the optimal welfare. [24] further shows that no quantile mechanism can obtain more than $1 - \frac{1}{e}$ fraction of the optimal welfare in the worst case. This result is sometimes interpreted as saying no mechanism can obtain an approximation ratio better than $1 - \frac{1}{e}$ with only information about the seller's value distribution. Our second result shows that there is a strictly better way to use the information about seller's value distribution. Indeed, with minimal information about F_S, i.e., its mean $\mathbb{E}[S]$, one can design a fixed-price mechanism that strictly outperforms the optimal quantile mechanism that requires full knowledge of F_S. Moreover, the quantile mechanism is asymmetric and only defined when we know the seller's value distribution. We show in Theorem D.1 that this is unavoidable, as no quantile mechanism over buyer's value distribution can guarantee a constant fraction of the optimal welfare.[2] In contrast, our second result holds when the designer only knows the mean of the buyer's value distribution F_B.

Mechanism 1: Quantile mechanism.

Input: A distribution Q over the interval $[0,1]$;
Randomly choose a quantile $x \in [0,1]$ according to Q;
Output the x-quantile of the seller's distribution as the price. Let F_S be the seller's distribution and $F_S^{-1}(\cdot)$ be seller's quantile function mapping any quantile to its corresponding value. The quantile mechanism outputs $F_S^{-1}(x)$ as the price;

Fixed-price mechanism using finitely many samples. Finally, we consider a different limited information model and initiate the study of approximating the optimal social welfare in using *finitely many samples*. Namely, we are given a finite and limited number of samples, e.g., 3 or 5 samples, and the goal is to design the best mechanism possible using these samples. It is important to distinguish this setting from the more standard *large sample* setting, where the goal is to determine the number of samples needed to design a $(1-\varepsilon)$-optimal mechanism (or optimal in a certain mechanism class) grows as a function of $\frac{1}{\varepsilon}$ and other parameters of the mechanism design environment. The sample complexity in large sample settings is usually stated using the big-O notation and ignores the accompanying constant. As a result, these bounds are often vacuous when apply to the *small sample* regime, where there are only a small finite number of samples available.

Contribution 3: We introduce a new family of mechanisms – *order statistic mechanisms* (Mechanism 2); and provide an exact characterization of the optimal order statistic mechanisms for any fixed number of samples (Theorem 5.1 and Theorem 5.2). Using our characterization, we can compute the optimal approximation ratio obtainable for any sample size.

Recent results show that one can already obtain a constant factor approximation to the optimal welfare using a single sample from the seller's distribution [3, 16, 24]. However, techniques from these papers are tailored for the single sample setting and are difficult to generalize to even the case when two samples are available. We provide a rich family of mechanisms that is well-defined for any number of samples and characterize their performance. Using the characterization, we manage to optimize within

[2] This asymmetry is due to the asymmetry of the initial allocation – the item is owned by the seller.

this family of mechanisms for any fixed number of samples.

Mechanism 2: Order statistic mechanism with N samples.

Input: A distribution P over $[N]$;
Randomly choose a number $i \in [N]$ according to the distribution P;
Given N samples from the seller, select the i-th smallest sample as the price, which is the i-th order statistic of these samples;

By numerically computing the optimal approximation ratios of order statistic mechanisms, we observe that the optimal order statistic mechanism with a small number of samples is usually sufficient to significantly boost the approximation ratio. For example, in the symmetric setting, i.e., $F_S = F_B$, *five samples* is sufficient to obtain an approximation ratio that is within 1% of the optimal ratio achievable by any fixed-price mechanism; in the asymmetric setting, i.e., $F_S \neq F_B$, the approximation ratio improves from 1/2 to 0.578 when the sample size increases from one to three. Another natural mechanism is the empirical risk minimization (ERM) mechanism, where one selects a price to maximize the social welfare w.r.t. the empirical distribution. We compare the performance of the optimal order statistic mechanism with ERM for sample size $N = 2, 3, 5, 10$ in the symmetric setting. In all cases, the order statistic mechanism substantially outperform the ERM. See Table 1 and 2 for our computed ratios in the symmetric and asymmetric cases respectively.

Our analysis of the order statistic mechanisms builds on an interesting connection between the order statistic mechanisms and the quantile mechanisms, that is, any order statistic mechanism is also a quantile mechanism. Note that the i-th order statistic over N samples drawn uniformly and independently from $[0,1]$ has density $f_N^i(x) = N\binom{N-1}{i-1} \cdot x^{i-1} \cdot (1-x)^{N-i}$. Suppose we use the i-th order statistic as the price, then it is equivalent to the quantile mechanism who selects a quantile corresponding to the density function $f_N^i(\cdot)$. More generally, if we choose the i-th order statistic with probability P_i, then the order statistic mechanism is equivalent to the quantile mechanism that chooses the quantile according to the density function $q(\cdot) = \sum_{i=1}^N P_i f_N^i(\cdot)$. With this connection, we can focus on quantile mechanisms, and we characterize the approximation ratio of any quantile mechanism as the solution of a minimization problem (Lemma 5.3). By applying this characterization for quantile mechanisms to order statistic mechanisms, we show that for any fixed sample size N, the ratio of the optimal order statistic mechanism is exactly the solution of a max-min optimization problem. Despite that the optimization problem seems intractable in general, we manage to solve it with sufficient numerical accuracy for $N \leq 10$. Although we only study approximating social welfare in bilateral trade in this paper, we believe this perspective of viewing sample-based mechanisms through the lens of quantile mechanisms is novel and has broader applications, especially in the small sample regime where the designer only has access to finitely many samples.

1.2 Related Work

Gains from Trade Maximization in Two-Sided Markets. Another important objective in two-sided markets is the gains from trade (GFT), which measures the increment of the welfare after the trade. Note that [29] also implies that optimal GFT is not achievable in bilateral trade. There has been increasing interest from the algorithmic mechanism design community to study the approximability of the optimal GFT [6, 8, 12, 2, 3, 10, 14]. It will be interesting to study the optimal approximation ratio obtainable for GFT maximization in both the full information and the limited information settings.

Sample-Based Mechanism Design. Sample-based mechanism design has become a central topic in algorithmic mechanism design as it provides an alternative model that weakens the classical but sometimes unrealistic Bayesian assumption. The results in this direction can be roughly partition into two groups: (1) Large sample results, where the goal is to determine the number of samples needed to design a $1-\varepsilon$-optimal mechanism (or optimal in a certain mechanism class) as a function of $\frac{1}{\varepsilon}$ and other parameters of the mechanism design environment, e.g., [17, 11, 26, 21, 31, 28, 27, 9, 7] or (2) Single sample results, where the goal is to determine the optimal approximation ratio obtainable using a single sample, e.g.,[18, 15, 19, 24, 20, 16]. Our result does not fit in to either of the groups. In particular, we study the regime where the designer has a small fixed number of samples, as a result, the machinery developed for large number of samples or a single sample does not apply to our setting. A recent line of works focus on the same regime as ours but for the monopolist pricing problem [4, 13, 1]. Due to the different nature of the studied problems, their techniques also do not apply here.

2 Preliminaries

Bilateral Trade. We study the bilateral trade problem. In this setting, there are two agents, a buyer and a seller, trade a single indivisible item. The seller owns the item and values it at S while the buyer values the item at B. Both S and B are non-negative and unknown to us but they are respectively drawn from distributions F_S and F_B independently. We assume that F_S and F_B are continous distributions. Actually, such assumption is w.l.o.g. and we discuss the reduction from distributions with point masses to continous ones in Appendix A.

Fixed-price Mechanism. We consider fixed-price mechanisms, which offer a price p to trade the item. The trade happens if and only if both the seller and the buyer accept the price, i.e., $B \geq p \geq S$. As shown by [23], fixed-price mechanism is the only dominant-strategy incentive-compatibility mechanism. In this paper, we consider (possibly randomized) fixed-price mechanisms. We abuse notation and use $\mathcal{M}(F_S, F_B)$ or $\mathcal{M}(\mathcal{I})$ where $\mathcal{I} = (F_S, F_B)$ to denote the distribution of prices p selected by mechanism $\mathcal{M}$ on instance $\mathcal{I} = (F_S, F_B)$.

Welfare and Approximation Ratio. We consider the objective of social welfare in this paper. For an instance $\mathcal{I} = (F_S, F_B)$, the optimal welfare is defined as:

$$\text{OPT-}\mathcal{W}(\mathcal{I}) = \mathop{\mathbb{E}}_{S\sim F_S, B\sim F_B}[\max(S, B)]$$

Similarly, for a fixed-price mechanism $\mathcal{M}$, the expected welfare on instance $\mathcal{I}$ can be written as:

$$\mathcal{W}(\mathcal{M}, \mathcal{I}) = \mathop{\mathbb{E}}_{\substack{S\sim F_S, B\sim F_B \\ p\sim \mathcal{M}(F_S, F_B)}}\left[S + \mathbb{1}[S \leq p \leq B] \cdot (B - S)\right]$$

Our goal is to maximize the approximiation ratio. That is, find some mechanism $\mathcal{M}$ maximize the following ratio.

$$\min_{\mathcal{I}=(F_S, F_B)} \frac{\mathcal{W}(\mathcal{M}, \mathcal{I})}{\text{OPT-}\mathcal{W}(\mathcal{I})}$$

Quantile Function. Suppose $F(\cdot)$ is the c.d.f. of a distribution, and we define $F^{-1}(\cdot)$ as the quantile function mapping the quantile to its corresponding value in this distribution. That is, $F^{-1}(x) = \inf\{y \mid F(y) = x\}$.

3 A Near-Optimal Mechanism in the Full Information Setting

In this section, we show a near-optimal fixed-price mechanism when given the full information of the buyer and the seller.

Theorem 3.1. *There exists a DSIC, individually rational, budget balanced mechanism that achieves at least 0.72 fraction of the optimal welfare for any instance $\mathscr{I} = (F_S, F_B)$. Moreover, no such mechanism has an approximation ratio better than 0.7381.*

To prove this, we first identify the best fixed-price mechanism when given the instance $\mathscr{I} = (F_S, F_B)$. Then, the approximation ratio is determined by the mechanism's performance on the worst-case instance. Such a worst-case instance could be characterized by an infinite dimensional quadratically constrained quadratic program (QCQP). However, the infinite dimensional program is hard to solve directly. Instead, we use two finite programs that can be solved numerically to upper bound and lower bound the infinite dimensional program. Additionally, we show that the optimal solutions of these two programs converge to the optimal solution of the infinite dimensional program as the number of variables tends to infinity.

3.1 Characterizing the Optimal Mechanism

We first characterize the optimal fixed-price mechanism via an infinite dimensional QCQP. Given any instance $\mathscr{I} = (F_S, F_B)$, we could assume that OPT-$\mathscr{W}(\mathscr{I}) = 1$ without loss of generality since we can always scale the instance so that this is true. The optimal fixed-price mechanism corresponds to choosing a price $p \in \arg\max_p \mathscr{W}(\mathscr{I}, p)$. The following program captures the worst-case instance for fixed-price mechanisms.

The Optimization Problem FullOp
$$\begin{aligned} \min_{\mu,\nu,r} \quad & r \\ \text{s.t.} \quad & \mu, \nu \text{ are probability measures defined on } \mathbb{R}_{\geq 0} \qquad (1) \\ & \text{OPT-}\mathscr{W}(\mathscr{I}) \stackrel{\text{def}}{=} \int_{\mathbb{R}_{\geq 0}} \int_{\mathbb{R}_{\geq 0}} \max(x, y)\, \nu(dy)\, \mu(dx) \geq 1 \\ & \mathscr{W}(\mathscr{I}, t) \stackrel{\text{def}}{=} \int_{\mathbb{R}_{\geq 0}} x\, \mu(dx) + \int_{\mathbb{R}_{\geq 0}} \int_{\mathbb{R}_{\geq 0}} (y - x) \cdot \mathbb{1}[x \leq t \leq y]\, \nu(dy)\, \mu(dx) \leq r \qquad \forall t \geq 0 \end{aligned}$$

Lemma 3.1. *The value of the optimal solution of* FullOp *is the* tight *worst-case approximation ratio achievable by a fixed-price mechanism.*

The proof of Lemma 3.1 is postponed to Appendix B.1. Since it is difficult to directly solve an infinite dimensional program like FullOp, we approximate FullOp from both above and below by constructing two families of finite programs which provide an upper bound and a lower bound respectively.

3.2 Factor Revealing Program for the Approximation Ratio under Full Information

We show that the approximation ratio of the optimal fixed price mechanism is at least 0.72, which significantly improves the previous state of the art bound of $1 - 1/e + \varepsilon$ with $\varepsilon \approx 10^{-4}$. Our approach is to find a fixed-price mechanism whose performance under the worst distribution is maximized. This is exactly captured by the optimization problem FullOp. However, it is an infinite-dimensional program. In this section, we consider a discretized version of FullOp. More specifically, we assume that OPT-$\mathscr{W}(\mathscr{I}) = 1$, and we restrict the mechanism to only choose price from a finite set $P = \{p_1, p_2, \ldots, p_k\}$. What we manage to show is that the optimal value of the optimization problem LowerOp is indeed a lower bound on the maximum approximation ratio one can obtain using prices from P for instance $\mathscr{I}$. We establish the following two crucial properties: (i) For any $\mathscr{I} = (F_S, F_B)$ satisfying OPT-$\mathscr{W}(\mathscr{I}) = 1$, we can carefully round

F_S and F_B to two discrete distributions supported on P, where $\{s_1,\dots,s_n\}$ and $\{b_1,\dots,b_n\}$ can be viewed as the corresponding "probability mass function" for the discretized distributions of the seller and the buyer.[3] Importantly, $\{s_1,\dots,s_n\}$ and $\{b_1,\dots,b_n\}$ satisfy inequalities (2) - (5). (ii) For any price p_t, the welfare from the corresponding fixed-price mechanism under $\mathscr{I}$ is at least the welfare under the rounded distributions $\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i)$. Therefore, if we choose r to be $\max_{t\in[n]} \mathscr{W}(\mathscr{I}, p_t)$, $\{s_1,\dots,s_n\}$, $\{b_1,\dots,b_n\}$, and r form a feasible solution of LowerOp, which implies that the optimal value of LowerOp is no greater than the constructed r. As the rounded distribution needs to satisfy a sequence of constraints (especially constraint (5)), the procedure we use to round F_S and F_B is subtle and does not simply round things up or down. See Appendix B.2 for details.

The Optimization Problem LowerOp

$$\min_{\substack{s_1,s_2\cdots,s_n\\ b_1,b_2,\cdots,b_n,r}} \quad r$$

$$\text{s.t.} \quad s_i, b_i \geq 0 \qquad \forall i \in [n] \tag{2}$$

$$\sum_{i=1}^n s_i \geq 1 \quad \text{and} \quad \sum_{i=1}^n b_i \geq 1 \tag{3}$$

$$\sum_{i=1}^n s_i \leq 1 + \frac{1}{p_n} \quad \text{and} \quad \sum_{i=1}^n b_i \leq 1 + \frac{1}{p_n} \tag{4}$$

$$\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j) \geq 1 \tag{5}$$

$$\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1} \sum_{j=t+1}^n s_i b_j (p_j - p_i) \leq r \qquad \forall t \in [n] \tag{6}$$

Lemma 3.2. *For any* $0 = p_1 < p_2 < \cdots < p_n$*, let* r^* *be the optimal value of* LowerOp*. Suppose* M *is the mechanism that chooses the best price from the set* $\{p_1 \cdot \mathbb{E}[\max(S,B)], p_2 \cdot \mathbb{E}[\max(S,B)], \cdots, p_n \cdot \mathbb{E}[\max(S,B)]\}$ *to maximize the welfare. The welfare obtained by* M *is at least* $r^* \cdot OPT$*.*

We defer the proof of the lemma to Appendix B.2.

3.3 Hardness Result under Full Information

In this section, our goal is to find a threshold and an instance such that no fixed-price mechanism has an approximation ratio better than the threshold on this instance. We focus on discrete distributions and consider an instance $\mathscr{I} = (F_S, F_B)$ where F_S is a discrete distribution supported on $\{p_1+\varepsilon, p_2+\varepsilon, \cdots, p_n+\varepsilon\}$, and F_B is a discrete distributions supported on $P = \{p_1, p_2, \cdots, p_n\}$. For such instance, the optimal price must also lie in the set $\{p_i+\varepsilon\}_{i\in[n]}$, as choosing a price x where $p_i+\varepsilon \leq x < p_{i+1}+\varepsilon$ is equivalent to choosing a price of $p_i+\varepsilon$. Therefore, any valid solution for the optimization problem below corresponds to a hard instance.

Lemma 3.3. *For any valid solution* $(s_1, s_2, \cdots, s_n, b_1, b_2, \cdots, b_n, r)$ *of* UpperOp *(defined in Appendix B.3) satisfying* $r = \max_{t\in[n]} \sum_{i=1}^n s_i p_i + \sum_{i=1}^t \sum_{j=t+1}^n s_i b_j (p_j - p_i)$ *and* $\varepsilon > 0$*, there exists an instance* $\mathscr{I} = (F_S, F_B)$ *such that no fixed-price mechanism can achieve more than* $(r+\varepsilon)$*-fraction of the optimal welfare on this instance.*

[3] For technical reasons, $\{s_1,\dots,s_n\}$ and $\{b_1,\dots,b_n\}$ do not exactly correspond to probability mass functions, but viewing them as the probability mass functions gives the right intuition.

The proof of Lemma 3.3 is deferred to Appendix B.3.

Proof of Theorem 3.1. With Lemma 3.2 and Lemma 3.3, we are now ready to prove Theorem 3.1. For the numerical results, our anonymous GitHub repository(`https://github.com/BilateralTradeAnonymous/On-the-Optimal-Fixed-Price-Mechanism-in-Bilateral-Trade`) provides all the certificates and codes and also carefully explains all the details.

For the lower bound, we choose $n = 16$. Using Gurobi [22], we obtain a lower bound of 0.72 for the optimization problem LowerOp for a carefully chosen set of price $\{p_1, p_2, \cdots, p_n\}$.[4] Therefore, by Lemma 3.2, there exists a 0.72-approximate fixed-price mechanism.

Things become much easier for the upper bound since we only need to find a feasible solution instead of proving a lower bound of the optimal value. We choose $n = 100$ and numerically solve UpperOp with a specific support $\{p_1, p_2, \cdots, p_n\}$ and find a feasible solution that satisfies the constraints in Lemma 3.3 where $r \le 0.7381$. Together with Lemma 3.3, we then find a hard instance such that no fixed-price mechanism attains a 0.7381-approximation of the optimal welfare. Please check our GitHub repository for the detailed specification of the distributions.

□

Finally, we would like to point out that the optimal value obtained by LowerOp and UpperOp will converge to the optimal value as the discretization accuracy tends to 0.

Lemma 3.4. *Let r^* be the optimal value of* FullOp*, i.e. the optimal approximation ratio. For any $\varepsilon > 0$, there exists two sets numbers $0 = p_1 < p_2 < \cdots < p_n$ and $\{p'_1, p'_2, \cdots, p'_{n'}\}$ such that the optimal value of* LowerOp *with respect to $\{p_1, p_2, \cdots, p_n\}$ is at least $r^* - \varepsilon$ and the optimal value of* UpperOp *w.r.t. $\{p'_1, p'_2, \cdots, p'_{n'}\}$ is at most $r^* + \varepsilon$.*

The proof of Lemma 3.4 is deferred to Appendix B.4

4 Breaking $1 - 1/e$ with Limited Information

We consider the limited information setting where we only knows the mean of the seller or the buyer. [24] shows that any mechanism that only uses quantile information from the seller can not achieve a better performance of $1 - 1/e$. However, we observe that even with minimal information of F_S such as its mean $\mathbb{E}[S]$ (or similarly $\mathbb{E}[B]$), we can break the $1 - 1/e$ barrier. We again provide a factor revealing program for this setting. Although it looks similar to LowerOp, there is some subtle differences in how we discretize a continuous distribution. See Appendix C for details.

Theorem 4.1. *Consider the following fixed-price mechanism: Given $\mathbb{E}[B]$ (or $\mathbb{E}[S]$), it randomly pick a number $x \sim P$ according to a distribution P, and selects $x \cdot \mathbb{E}[B]$ (or $x \cdot \mathbb{E}[S]$) as the price. There exists a distribution P_S for the seller and a distribution P_B for the buyer such that the corresponding fixed-price mechanism achieves at least $0.65 \cdot OPT$ welfare.*

The high level idea is as follows. Lemma B.1 shows us how to discretize a continuous distribution so that $\mathbb{E}[\max(S, B)]$ increases and $\mathbb{E}[\mathscr{W}(\mathscr{I}, p_t)]$ decreases. In other words, the discretization worsens the instance. Therefore, we could use a similar technique to derive a lower bound of the approximation ratio. The complete proof of Theorem 4.1 is in Appendix C

[4] We choose $\{p_1, p_2, \cdots, p_{16}\}$ to be $\{0.0, 0.1, 0.19, 0.27, 0.315, 0.355, 0.395, 0.44, 0.485, 0.535, 0.595, 0.665, 0.74, 0.875, 1.195, 1000.0\}$ to derive the 0.72. These numbers are chosen heuristically to provide good coverage between 0.3 to 0.5, which is the region with concentration of probability mass in some bad instances we encounter.

5 Fixed-Price Mechanism with Different Numbers of Samples

In this section, we consider the limited information setting where we only have sample access to the distributions. We focus on order statistic mechanisms which is defined in Section 1.1 and our results cover different number of samples for both symmetric and general instances. In the small sample regime, we are able to characterize the optimal order statistic mechanism with any fixed number of samples. When the number of samples goes to infinity, we show that the optimal quantile mechanism can be approximated by order statistics mechanism as closely as desired and also obtain an upper bound on the sample complexity. Finally, recall that we assume the distributions for the seller and the buyer are continuous. See Appendix A for details.

5.1 Order Statistic Mechanisms

To start with, we briefly discuss these two families of mechanisms that is used in the sample setting and give high level ideas on how to design the order statistic mechanisms. Order statistic mechanisms will be used when we only have samples from the distribution and quantile mechanisms will help us analyze the performance of order statistic mechanisms. Actually, we will point out that quantile mechanisms and order statistic mechanisms are equivalent in some sense.

5.1.1 Connection Between Two Mechanisms

Next we aim to show the connection between these two mechanisms. Such observations give us insights on designing mechanisms with small or large number of samples.

The order statistic mechanism is a special kind of quantile mechanisms First, we can see that the following two operations are equivalent:

- Draw a sample from distribution F.
- Uniformly sample a quantile x from $[0,1]$, and use $F^{-1}(x)$ as the sample.

Now suppose $f_N^i(x) = N\binom{N-1}{i-1} \cdot x^{i-1} \cdot (1-x)^{N-i}$ be the p.d.f. of the i-th order statistic over N samples drawn uniformly and independently from $[0,1]$ and let P_i to be $\Pr_{x\sim P}[x=i]$ for any distribution P over $[N]$. Using similar ideas above, it can be proved that any order statistic mechanism P is equivalent to a quantile mechanism Q with probability density function

$$q(x) = \sum_{i=1}^{N} P_i f_N^i(x)$$

Therefore, we can analyze the approximation ratio of quantile mechanism Q instead of order statistic mechanism P. If we are able to compute the approximation ratio of any quantile mechanism Q, it follows that we can also characterize the optimal order statistic mechanism exactly. When the number of samples are small, we can have a fine-grained analysis of the order statistic mechanisms and use these limited samples carefully. Section 5.2 actually follow such intuitions to characterize the best possible order statistic mechanism.

Quantile mechanisms can be approximated by order statistic mechanisms within any small error Our goal is that for any quantile mechanism Q with p.d.f. $q(x)$, we need to find some integer N and a distribution P over $[N]$, such that

$$q(x) \approx \sum_{i=1}^{N} P_i f_N^i(x)$$

Since $\sum_{i=1}^{N} P_i f_N^i(x)$ is a polynomial of degree $N-1$, this could be done for any continuous $q(x)$ on $[0,1]$ since the Weierstrass approximation theorem states that every continuous function defined on a closed interval can be uniformly approximated as closely as desired by a polynomial function. What's more interesting is that $\{f_N^i(x)\}_{i=1}^N$ are Bernstein basis polynomials and there are a series of work showing that (stochastic) Bernstein polynomials can efficiently and uniformly approximate to any continous function. Therefore, we can have an asymptotic analysis of the order statistic mechanism. What's more, such observation also shows that we have a block-box transformation from any quantile mechanism to mechanisms only using samples. Section 5.3 uses such techniques and ideas.

5.2 Small Sample Regime

In this section, we characterize the optimal order statistic mechanisms with any fixed number of samples for both symmetric and general instances. We first show that, in any setting, if we are able to give a tight analysis of the quantile mechanism, we could directly characterize the optimal order statistic mechanism with any fixed number of samples via an optimization problem. In the next, we show a tight analysis of the quantile mechanism on both symmetric and general instances, and thus we obtain the characterization of the optimal order statistic mechanism.

Recall that an order statistics mechanism with N samples randomly choose a number $i \in [N]$ according to a previously defined distribution P and select the i-th smallest sample as the price, and a quantile mechanism randomly choose a quantile $x \in [0,1]$ from a determined distribution Q and choose the x-quantile, i.e. $F_S^{-1}(x)$, as the price. Since every quantile mechanism and order statistic mechanism is determined by the previously defined distribution, we abuse the notation and use distribution P over $[N]$ denote its corresponding order statistic mechanism and distribution Q over $[0,1]$ denote its corresponding quantile mechanism.

Lemma 5.1. *Suppose $\mathscr{C} : \Delta([0,1]) \mapsto \mathbb{R}$ maps every quantile mechanism P to its exact approximation ratio. Let $\mathscr{P}(Q)$ be the corresponding quantile mechanism of the order statistic mechanism Q. Fixing the number of samples N, the optimal order statistic mechanism with N samples Q_N^* is characterized by the following optimization problem:*

$$Q_N^* = \arg\max_{Q \in \Delta_N} \mathscr{C}(\mathscr{P}(Q))$$

where $\Delta([0,1])$ is the set of all distributions over $[0,1]$, i.e. the set of all quantile mechanisms, and Δ_N is the set of all distributions over $[N]$, i.e. the set of all order statistic mechanisms with N samples.

The proof of Lemma 5.1 is quite straightforward and thus is postponed to Appendix D.2.

5.2.1 Symmetric Instances

Now we study the case when the distributions are symmetric, i.e., $F_S = F_B$, which means that the seller's value S and the buyer's value B are drawn from the same distribution. For simplification, we will use F to refer to their distributions in this setting.

In order to find out the optimal order statistic mechanism, we need to first give a tight analysis of the quantile mechanism.

Lemma 5.2. *For any quantile mechanism for symmetric instance with distribution Q over $[0,1]$, the approximation ratio is exactly*

$$\inf_{x \in [0,1)} \frac{\int_{[0,x]} t(1-x)\, dQ(t) + \int_{(x,1]} (1-t) x\, dQ(t) + (1-x)}{1-x^2}$$

where $Q(t)$ is the cumulative distribution function of distribution Q.

Therefore, combining Lemma 5.1 and Lemma 5.2, we could characterize the optimal order statistic mechanism via an optimization problem.

Theorem 5.1. *The optimal order statistic mechanism with N samples for symmetric instances is the solution to the following optimization problem:*

$$P_N^* = \max_{\substack{P_1,P_2,\cdots,P_N\geq 0 \\ \sum_{i=1}^N P_i=1}} \inf_{x\in[0,1)} \frac{\int_0^x p(t)t(1-x)\,dt + \int_x^1 p(t)(1-t)x\,dt + (1-x)}{1-x^2}$$

where $p(t) = \sum_{i=1}^N P_i f_N^i(x)$ *and* $f_N^i(x) = N\binom{N-1}{i-1}\cdot x^{i-1}\cdot(1-x)^{N-i}$ *is the p.d.f. of the i-th order statistic over N samples drawn uniformly and independently from* $[0,1]$.

The proof of Lemma 5.2 and Theorem 5.1 is in Appendix D.3 and D.4.

It turns out the optimization above is computationally tractable when N is not too large. We solve the optimization problem and find out the optimal order statistic mechanism numerally with different numbers of samples N.

To compare with the order statistic mechanisms, we also consider the most natural sample-based mechanism – the *Empirical Risk Minimization mechanism (ERM)*. We first provide the formal definition below.

Definition 5.1 (Empirical Risk Minimization Mechanism). *Given N samples* $X_1, X_2, \ldots, X_N$ *drawn from* F, *define* $\tilde{F}$ *be the empirical distribution of these N samples. That is to say,* $\tilde{F}$ *is the distribution with c.d.f.* $\tilde{F}(x)$ *satisfying:*

$$\tilde{F}(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}[x \geq X_i]$$

The Empirical Risk Minimization mechanism (ERM) is the mechanism that computes the optimal price according the empirical distribution $\tilde{F}$. *In particular, for N samples* $X_1, X_2, \ldots, X_N$,

$$\mathsf{ERM}(X_1, X_2, \ldots, X_N) = \arg\max_p \mathop{\mathbb{E}}_{S\sim\tilde{F}, B\sim\tilde{F}} [S + (B-S)\cdot\mathbb{1}[B \geq p \geq S]]$$

If there are multiple prices p that maximize the expected welfare, the ERM mechanism may select any of them.

For $N = 1,2,3,5,10$, we compute the approximation ratio of order statistic mechanisms and also show the upper bound of ERM. The results are listed below. To prove the upper bound, we use a counter example in [24] and show that ERM has a bad performance on this instance. We defer the complete proof of the upper bound of ERM to Appendix D.5 and the details of numerical results to Appendix E.1.

#Samples	Order Statistics Mechanism	ERM
1	0.75	0.5
2	0.821	≤ 0.67
3	0.822	≤ 0.75
5	0.847	≤ 0.76
10	0.852	≤ 0.80
∞	$\frac{2+\sqrt{2}}{4} \approx 0.8536$	/

Table 1: Approximation Ratios with Different Number of Samples in the Symmetric Setting.

5.2.2 General Instances

Now we consider the general setting, where the buyer's distribution may be different from the seller's. Recall that we only consider mechanisms over seller's information since there is no constant quantile or order statistic mechanism over seller's information. Using similar ideas, we first show a tight analysis regarding quantile mechanisms, which would guide us to discover the optimal order statistic mechanism.

Lemma 5.3 (Theorem 4.1 of [5]). *For any quantile mechanism Q (over seller's distribution) with cumulative distribution function Q, its approximation ratio is exactly*

$$\min_{x\in[0,1]} \int_{[0,x]} t\, dQ(t) + 1 - x$$

Similarly, combining Lemma 5.1 and Lemma 5.3, we are able to charaterize the optimal order statistic mechanism over with N samples from seller's distribution by an optimization problem:

Theorem 5.2. *The optimal order statistic mechanism with N samples for symmetric instances is the solution to the following optimization problem:*

$$P_N^* = \max_{\substack{P_1,P_2,\cdots,P_N\geq 0\\ \sum_{i=1}^N P_i=1}} \min_{x\in[0,1]} \int_{[0,x]} t\cdot p(t)\,dt + 1 - x,$$

where $p(x) = \sum_{i=1}^N P_i \cdot f_N^i(x)$ and $f_N^i(x) = N\binom{N-1}{i-1}\cdot x^{i-1}\cdot(1-x)^{N-i}$ is the p.d.f. of the i-th order statistic over N samples drawn uniformly and independently from $[0,1]$.

The proof of Lemma 5.3 and Theorem 5.2 is deferred to Appendix D.6 and D.7.

Similarly, such optimization problem is easy to solve when the number of samples N is not to large. We solve the optimization problem numerically for $N = 1,2,3,5,10$. Note that we do not compare our mechanism to the Empirical Risk Minimization mechanism in the general setting. This is because we only have sample access to the seller's distribution, and the ERM can not be implemented without the buyer's samples. The details of numerical results is defered to Appendix E.2.

#Samples	Order Statistic Mechanism
1	0.5
2	0.531
3	0.578
5	0.601
10	0.615
∞	$1-\frac{1}{e}\approx 0.632$

Table 2: Approximation Ratios with Different Number of Samples
In the General Setting.

5.3 Asymptotic Analysis: From Quantile to Order Statistics

In this section, we turn to the case when the number of samples tends to infinity. As we show in section 5.1.1, we could approximate any quantile mechanisms by order statistic mechanisms within any small error. Using such ideas, we provide a "black-box" reduction that allows us to convert any quantile mechanism with continuous probability density function $q(x)$ to order statistic mechanism with N samples. Here N is usually a polynomial of $\frac{1}{\varepsilon}$, as long as the probability density function is not too crazy. We now formally write it down.

Lemma 5.4. *Let $\mathscr{C} : \Delta([0,1]) \mapsto \mathbb{R}$ be a function that maps every quantile mechanism Q with continuous probability density function to its approximation ratio such that for any quantile mechanism Q_1 with p.d.f. $q_1(x)$ and quantile mechanism Q_2 with p.d.f. $q_2(x)$, it holds that*

$$\mathscr{C}(Q_1) - \mathscr{C}(Q_2) \geq -c \cdot \left|q_1 - q_2\right|_\infty$$

where c is a constant.

Now let Q be any quantile mechanism with continuous probability density function $q(x)$. Define M as $\max_{x\in[0,1]} q(x)$. For any $\varepsilon > 0$, suppose n is a positive integer satisfying that

$$\omega\left(\frac{1}{\sqrt{n-1}}\right) \leq \varepsilon/100 \tag{7}$$

$$2n\exp\left(-\frac{\varepsilon^2}{8\omega^2\left(\frac{1}{\sqrt{n-1}}\right)}\right) \leq \varepsilon \tag{8}$$

$$\exp\left(-\frac{\varepsilon^2 n}{48M^3}\right) \leq \varepsilon/2 \tag{9}$$

where $w(h) = \sup_{\substack{x,y\in[0,1] \\ |x-y|\leq h}} \left|q(x) - q(y)\right|$. Then, there exists an order statistic mechanism with n samples that achieves an approximation ratio of $\mathscr{C}(Q) - c\cdot\varepsilon$.

The high level idea of the proof is as follows. Since we know that probability density functions of order statistics form Bernstein basis polynomials, we could approximate the p.d.f. of the quantile mechanism $q(x)$ within any small error. Inequality (7), (8) and (9) actually help us to get an order statistic mechanism whose corresponding distribution of quantile is close to the desired quantile mechanism Q. Finally, by the property of $\mathscr{C}$, we know that their approximation ratio is also close. The proof is postponed to Appendix D.8.

Finally, we show that we could apply lemma 5.4 to both the symmetric and general settings and convert the optimal quantile mechanism to order statistic mechanism within a error of at most ε using poly$\left(\frac{1}{\varepsilon}\right)$ samples. We leave the details of such applications to Appendix D.9.

References

[1] Amine Allouah, Achraf Bahamou, and Omar Besbes. Revenue Maximization from Finite Samples. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, EC '21, page 51, New York, NY, USA, July 2021. Association for Computing Machinery.

[2] Moshe Babaioff, Yang Cai, Yannai A. Gonczarowski, and Mingfei Zhao. The Best of Both Worlds: Asymptotically Efficient Mechanisms with a Guarantee on the Expected Gains-From-Trade. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, EC '18, page 373, New York, NY, USA, June 2018. Association for Computing Machinery.

[3] Moshe Babaioff, Kira Goldner, and Yannai A. Gonczarowski. Bulow-Klemperer-Style Results for Welfare Maximization in Two-Sided Markets. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 2452–2471. SIAM, 2020.

[4] Moshe Babaioff, Yannai A. Gonczarowski, Yishay Mansour, and Shay Moran. Are Two (Samples) Really Better Than One? In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, page 175, 2018.

[5] Liad Blumrosen and Shahar Dobzinski. (almost) efficient mechanisms for bilateral trading. *Games and Economic Behavior*, 130:369–383, 2021.

[6] Liad Blumrosen and Yehonatan Mizrahi. Approximating Gains-from-Trade in Bilateral Trading. In *Web and Internet Economics - 12th International Conference, WINE 2016, Montreal, Canada, December 11-14, 2016, Proceedings*, pages 400–413, 2016.

[7] Johannes Brustle, Yang Cai, and Constantinos Daskalakis. Multi-Item Mechanisms without Item-Independence: Learnability via Robustness. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20, pages 715–761, New York, NY, USA, July 2020. Association for Computing Machinery.

[8] Johannes Brustle, Yang Cai, Fa Wu, and Mingfei Zhao. Approximating Gains from Trade in Two-sided Markets via Simple Mechanisms. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 589–590, New York, NY, USA, June 2017. Association for Computing Machinery.

[9] Yang Cai and Constantinos Daskalakis. Learning Multi-Item Auctions with (or without) Samples. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 516–527, October 2017. ISSN: 0272-5428.

[10] Yang Cai, Kira Goldner, Steven Ma, and Mingfei Zhao. On Multi-Dimensional Gains from Trade Maximization. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Proceedings, pages 1079–1098. Society for Industrial and Applied Mathematics, January 2021.

[11] Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, 2014.

[12] Riccardo Colini-Baldeschi, Paul Goldberg, Bart de Keijzer, Stefano Leonardi, and Stefano Turchetta. Fixed price approximability of the optimal gain from trade. In *International Conference on Web and Internet Economics*, pages 146–160. Springer, 2017.

[13] Constantinos Daskalakis and Manolis Zampetakis. More revenue from two samples via factor revealing sdps. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 257–272, 2020.

[14] Yuan Deng, Jieming Mao, Balasubramanian Sivan, and Kangning Wang. Approximately efficient bilateral trade. *arXiv preprint arXiv:2111.03611*, 2021.

[15] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*, 2010.

[16] Paul Dütting, Federico Fusco, Philip Lazos, Stefano Leonardi, and Rebecca Reiffenhäuser. Efficient Two-Sided Markets with Limited Information. *arXiv:2003.07503 [cs]*, April 2021. arXiv: 2003.07503.

[17] Edith Elkind. Designing and learning optimal finite support auctions. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 736–745. Society for Industrial and Applied Mathematics, 2007.

[18] Hu Fu, Nicole Immorlica, Brendan Lucier, and Philipp Strack. Randomization Beats Second Price as a Prior-Independent Auction. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, page 323, 2015.

[19] Kira Goldner and Anna R Karlin. A prior-independent revenue-maximizing auction for multiple additive bidders. In *International Conference on Web and Internet Economics*, pages 160–173. Springer, 2016.

[20] Yannai A Gonczarowski and S Matthew Weinberg. The Sample Complexity of Up-to-ε Multi-Dimensional Revenue Maximization. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 416–426. IEEE, 2018.

[21] Chenghao Guo, Zhiyi Huang, and Xinzhi Zhang. Settling the Sample Complexity of Single-parameter Revenue Maximization. In *the 51st Annual ACM Symposium on Theory of Computing (STOC)*, 2019.

[22] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2022.

[23] Kathleen M Hagerty and William P Rogerson. Robust trading mechanisms. *Journal of Economic Theory*, 42(1):94–107, 1987.

[24] Zi Yang Kang, Francisco Pernice, and Jan Vondrák. Fixed-price approximations in bilateral trade. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2964–2985. SIAM, 2022.

[25] Zi Yang Kang and Jan Vondrák. Strategy-proof approximations of optimal efficiency in bilateral trade. 2018.

[26] Mehryar Mohri and Andres Munoz Medina. Learning Theory and Algorithms for revenue optimization in second price auctions with reserve. In *ICML*, pages 262–270, 2014.

[27] Jamie Morgenstern and Tim Roughgarden. Learning simple auctions. In *Proceedings of the 30th Annual Conference on Learning Theory (COLT)*, 2016.

[28] Jamie H Morgenstern and Tim Roughgarden. On the pseudo-dimension of nearly optimal auctions. In *Proceedings of the the 29th Annual Conference on Neural Information Processing Systems (NIPS)*, 2015.

[29] Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983. Publisher: Elsevier.

[30] Xingping Sun, Zongmin Wu, and Xuan Zhou. On probabilistic convergence rates of stochastic bernstein polynomials. *Math. Comput.*, 90(328):813–830, 2021.

[31] Vasilis Syrgkanis. A sample complexity measure with applications to learning optimal auctions. In *Advances in Neural Information Processing Systems*, pages 5352–5359, 2017.

A Tie Breaking

For distribution D with point masses, the following reduction will convert it to continuous one. We will overload the notation of D and think of it as a bivariate distribution with the first coordinate drawn from the previous single-variate distribution D and the second tie-breaker coordinate drawn independently and uniformly from $[0,1]$. And $(X_1, t_1) > (X_2, t_2)$ if and only if either $X_1 > X_2$, or $X_1 = X_2$ and $t_1 > t_2$. Since the tie-breaker coordinate is continuous, the probability of having $(X_1, t_1) = (X_2, t_2)$ for any two values during a run of any mechanism is zero. Therefore we could define the c.d.f. of D as

$$F_D(X,t) = \Pr_{(Y,u)\sim(D,U[0,1])}[(Y,u) < (X,t)]$$

Remind the second coordinate is only used to break ties, and it does not affect the calculation of welfare. After including the additional random variable, we can see that D has been converted into a continuous distribution since its second coordinate is continuous.

B Missing Proofs in Section 3

B.1 Proof of Lemma 3.1

We first show the proof of Lemma 3.1.

The approximation ratio of the optimal fixed-price mechanism could be written as

$$\min_{\mathcal{I}=(F_S,F_B)} \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})}.$$

We first show that for any instance $\mathcal{I} = (F_S, F_B)$, there is a valid solution (u, v, r) such that $r = \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})}$. We could first simply scale the instance by $\frac{1}{\text{OPT-}\mathcal{W}(\mathcal{I})}$ to $\mathcal{I}' = (F'_S, F'_B)$ where OPT-$\mathcal{W}(\mathcal{I}') = 1$. Such scaling means that $\mathcal{W}(\mathcal{I}, \text{OPT-}\mathcal{W}(\mathcal{I})\cdot p) = \text{OPT-}\mathcal{W}(\mathcal{I})\cdot\mathcal{W}(\mathcal{I}',p)$ for all $p \in \mathbb{R}_{\geq 0}$. This implies that

$$\max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})} = \max_{p\in\mathbb{R}} \mathcal{W}(\mathcal{I}',p).$$

Therefore, let u and v be the probability measures of F'_S and F'_B and r be $\max_{p\in\mathbb{R}}\mathcal{W}(\mathcal{I}',p)$. It is easy to verify that (u, v, r) is a valid solution. Let r^* be the optimal value of FullOp, this implies that $r^* \leq \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})}$ for any instance $\mathcal{I} = (F_S, F_B)$. Taking the minimum over all possible $\mathcal{I}$, we then get that

$$r^* \leq \min_{\mathcal{I}=(F_S,F_B)} \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})} \tag{10}$$

Next, let (u^*, v^*, r^*) be the optimal solution of FullOp. Since u^*, v^* are both probability measures, let F^*_S, F^*_B be the corresponding distributions of u and v and $\mathcal{I}^* = (F^*_S, F^*_B)$ be the instance. Now by the constraint of FullOp, we know that OPT-$\mathcal{W}(\mathcal{I}^*) \geq 1$. Besides, (u^*, v^*, r^*) is an optimal solution implies that $r^* = \max_{p\in\mathbb{R}}\mathcal{W}(\mathcal{I}^*,p)$. Therefore,

$$r^* = \max_{p\in\mathbb{R}} \mathcal{W}(\mathcal{I}^*,p) \geq \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I}^*,p)}{\text{OPT-}\mathcal{W}(\mathcal{I}^*)} \geq \min_{\mathcal{I}=(F_S,F_B)} \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})} \tag{11}$$

Now combining inequality (10) and (11), it follows that

$$r^* = \min_{\mathcal{I}=(F_S,F_B)} \max_{p\in\mathbb{R}} \frac{\mathcal{W}(\mathcal{I},p)}{\text{OPT-}\mathcal{W}(\mathcal{I})}$$

which completes the proof.

B.2 Proof of Lemma 3.2

Before we give the proof of Lemma 3.2, we first show prove a lemma that helps us discretize a continuous distribution.

Lemma B.1. *For any instance* $\mathscr{I} = (F_S, F_B)$*, and* $0 = p_1 < p_2 < \cdots < p_n$*, there exists a set of numbers* $\{s_i\}_{i\in[n]}, \{b_i\}_{i\in[n]}$ *satisfying the following equations.*

$$s_i, b_i \geq 0 \quad \forall i \in [n] \tag{12}$$

$$1 \leq \sum_{i=1}^{n} s_i \leq 1 + \frac{\mathbb{E}[S]}{p_n} \quad 1 \leq \sum_{i=1}^{n} b_i \leq 1 + \frac{\mathbb{E}[B]}{p_n} \tag{13}$$

$$\sum_{i=1}^{n-1} s_i \leq 1 \quad \sum_{i=1}^{n-1} b_i \leq 1 \tag{14}$$

$$\mathop{\mathbb{E}}_{S\sim F_S}[S] = \sum_{i=1}^{n} s_i p_i \tag{15}$$

$$\mathop{\mathbb{E}}_{B\sim F_B}[B] = \sum_{j=1}^{n} b_j p_j \tag{16}$$

$$OPT\text{-}\mathscr{W}(\mathscr{I}) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{\substack{S\sim F_S\\ B\sim F_B}}[\max(S,B)] \leq \sum_{i=1}^{n}\sum_{j=1}^{n} s_i b_j \cdot \max(p_i, p_j) \tag{17}$$

$$\begin{aligned}\mathscr{W}(\mathscr{I}, p_t) &\stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{S\sim F_S}[S] + \mathop{\mathbb{E}}_{\substack{S\sim F_S\\ B\sim F_B}}\Big[(B-S)\cdot \mathbb{1}[S \leq p_t \leq B]\Big] \\ &\geq \sum_{i=1}^{n} s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^{n} s_i b_j (p_j - p_i) \quad \forall t \in [n]\end{aligned} \tag{18}$$

Proof. We construct $(s_1, \cdots, s_n, b_1, \cdots, b_n)$ as follows. For the seller, define

$$q_{s,i} = \mathop{\Pr}_{S\sim F_S}[p_i \leq S < p_{i+1}] \text{ and } E_{s,i} = \mathop{\mathbb{E}}_{S\sim F_S}[S \cdot \mathbb{1}[p_i \leq S < p_{i+1}]],\ \forall i \in [n],$$

where we assume that $p_{n+1} = +\infty$. It is clear from the definition that $q_{s,i}\cdot p_i \leq E_{s,i} \leq q_{s,i}\cdot p_{i+1}$. Therefore, for any $i \in [n-1]$, there exists non-negative numbers $s_{i,\text{LEFT}}$ and $s_{i+1,\text{RIGHT}}$ such that

$$s_{i,\text{LEFT}} + s_{i+1,\text{RIGHT}} = q_{s,i} \text{ and } s_{i,\text{LEFT}}\cdot p_i + s_{i+1,\text{RIGHT}}\cdot p_{i+1} = E_{s,i}. \tag{19}$$

We further define $s_{n,\text{LEFT}}$ as $E_{s,n}/p_n$ and $s_{1,\text{RIGHT}} = 0$. Now set $s_i = s_{i,\text{LEFT}} + s_{i,\text{RIGHT}}$ for all $i \in [n]$. For the buyer, we define $\{b_i\}_{i\in[n]}$ similarly.

We now verify that $(\{s_i\}_{i\in[n]}, \{b_i\}_{i\in[n]})$ satisfies the properties above. The non-negativity of s_i and b_i is immediately derived from $s_{i,\text{LEFT}}, s_{i,\text{RIGHT}} \geq 0$. From our definition, it is clear that $\sum_{i=1}^{n} q_{s,i} = 1$, therefore

$$\sum_{i=1}^{n} s_i = \sum_{i=1}^{n} s_{i,\text{LEFT}} + s_{i,\text{RIGHT}} \geq \sum_{i=1}^{n} q_{s,i} = 1,$$

and

$$\sum_{i=1}^{n-1} s_i = \sum_{i=1}^{n-1} s_{i,\text{LEFT}} + s_{i,\text{RIGHT}} \leq \sum_{i=1}^{n} q_{s,i} = 1.$$

We could also see that

$$\sum_{i=1}^{n} s_i = \sum_{i=1}^{n} s_{i,\text{LEFT}} + s_{i,\text{RIGHT}} \leq \sum_{i=1}^{n} q_{s,i} + s_{n,\text{LEFT}} = 1 + \frac{E_{s,n}}{p_n} \leq 1 + \frac{\mathbb{E}[S]}{p_n}.$$

For the expectations, it holds that

$$\sum_{i=1}^{n} s_i p_i = \sum_{i=1}^{n} \left(s_{i,\text{LEFT}} + s_{i,\text{RIGHT}}\right) \cdot p_i = \sum_{i=1}^{n} E_{s,i} = \mathbb{E}[S]$$

By symmetry, similar inequalities also holds for for $\{b_i\}_{i\in[n]}$. So far, we have verified that properties (12), (13) , (14), (15) and (16) are satisfied. It only remains to show that (17) and(18) holds.

For any $i \neq j \in [n]$, w.l.o.g. we can assume that $i < j$. We could see that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[\max(S,B)\cdot \mathbb{1}[S\in[p_i,p_{i+1})]\mathbb{1}[B\in[p_j,p_{j+1})]\right]\\
&=\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[B\cdot \mathbb{1}[S\in[p_i,p_{i+1})]\mathbb{1}[B\in[p_j,p_{j+1})]\right]\\
&=\mathop{\mathbb{E}}_{B\sim F_B}\left[B\cdot \mathbb{1}[B\in[p_j,p_{j+1})]\right]\cdot \Pr_{S\sim F_S}\left[S\in[p_i,p_{i+1})\right]\\
&=E_{b,j}\cdot q_{s,i}\\
&=(b_{j,\text{LEFT}}\cdot p_j + b_{j+1,\text{RIGHT}}\cdot p_{j+1})\cdot(s_{i,\text{LEFT}}+s_{i+1,\text{RIGHT}})\\
&=s_{i,\text{LEFT}}\cdot b_{j,\text{LEFT}}\max(p_i,p_j)+s_{i+1,\text{RIGHT}}\cdot b_{j,\text{LEFT}}\max(p_{i+1},p_j)\\
&\quad + s_{i,\text{LEFT}}\cdot b_{j+1,\text{RIGHT}}\max(p_i,p_{j+1})+s_{i+1,\text{RIGHT}}\cdot b_{j+1,\text{RIGHT}}\max(p_{i+1},p_{j+1})
\end{aligned}\tag{20}$$

The second equality is due to the independence between S and B.

Now consider the case when $i = j \le n-1$. For any $x, y \in [p_i, p_{i+1}]$, we have

$$\max(x,y) \le \max(p_i,y)\cdot\frac{p_{i+1}-x}{p_{i+1}-p_i} + \max(p_{i+1},y)\cdot\frac{x-p_i}{p_{i+1}-p_i},$$

as $\frac{p_{i+1}-x}{p_{i+1}-p_i} + \frac{x-p_i}{p_{i+1}-p_i} = 1$ and $\frac{p_{i+1}-x}{p_{i+1}-p_i}\cdot p_i + \frac{x-p_i}{p_{i+1}-p_i}\cdot p_{i+1} = x$.

Based on the inequality above, for any fixed $y \in [p_i, p_{i+1})$, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{S\sim F_S}\left[\max(S,y)\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]\\
\le&\mathop{\mathbb{E}}_{S\sim F_S}\left[\left(\max(p_i,y)\cdot\frac{p_{i+1}-S}{p_{i+1}-p_i}+\max(p_{i+1},y)\cdot\frac{S-p_i}{p_{i+1}-p_i}\right)\mathbb{1}[S\in[p_i,p_{i+1})]\right]\\
=&\max(p_i,y)\mathop{\mathbb{E}}_{S\sim F_S}\left[\frac{p_{i+1}-S}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]+\max(p_{i+1},y)\mathop{\mathbb{E}}_{S\sim F_S}\left[\frac{S-p_i}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]\\
=&y\cdot s_{i,\text{LEFT}}+p_{i+1}\cdot s_{i+1,\text{RIGHT}}
\end{aligned}$$

The last equality is because of the following identities:

$$\mathop{\mathbb{E}}_{S\sim F_S}\left[\frac{p_{i+1}-S}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]+\mathop{\mathbb{E}}_{S\sim F_S}\left[\frac{S-p_i}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]=\Pr[S\in[p_i,p_{i+1})]]$$

$$p_i\cdot\mathop{\mathbb{E}}_{S\sim F_S}\left[\frac{p_{i+1}-S}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]+p_{i+1}\cdot\mathop{\mathbb{E}}_{S\sim F_S}\left[\frac{S-p_i}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]=\mathbb{E}[S\cdot\mathbb{1}[S\in[p_i,p_{i+1})]].$$

Hence, we can conclude that $\left(\mathbb{E}_{S\sim F_S}\left[\frac{p_{i+1}-S}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right],\mathbb{E}_{S\sim F_S}\left[\frac{S-p_i}{p_{i+1}-p_i}\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]\right)$ is the unique solution to (19). Thus, these two numbers respectively equal to $s_{i,\text{LEFT}}$ and $s_{i+1,\text{RIGHT}}$.

Due to the inequality above, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{B\sim F_B}\left[\mathop{\mathbb{E}}_{S\sim F_S}\left[\max(S,B)\cdot\mathbb{1}[S\in[p_i,p_{i+1})]\right]\cdot\mathbb{1}[B\in[p_i,p_{i+1})]\right]\\
\le&\mathop{\mathbb{E}}_{B\sim F_B}\left[\left(B\cdot s_{i,\textsc{Left}}+p_{i+1}\cdot s_{i+1,\textsc{Right}}\right)\cdot\mathbb{1}[B\in[p_i,p_{i+1})]\right] \qquad (21)\\
=&b_{i,\textsc{Left}}s_{i,\textsc{Left}}\cdot p_i+b_{i+1,\textsc{Right}}s_{i,\textsc{Left}}\cdot p_{i+1}+b_{i,\textsc{Left}}s_{i+1,\textsc{Right}}\cdot p_{i+1}+b_{i+1,\textsc{Right}}s_{i+1,\textsc{Right}}\cdot p_{i+1}
\end{aligned}$$

The last special case is when $i=j=n$.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[\max(S,B)\cdot\mathbb{1}[S\ge p_n]\mathbb{1}[B\ge p_n]\right]&\le\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[BS/p_n\cdot\mathbb{1}[S\ge p_n]\mathbb{1}[B\ge p_n]\right]\\
&=p_n\cdot\left(\mathop{\mathbb{E}}_{S\sim F_S}\left[S\cdot\mathbb{1}[S\ge p_n]\right]/p_n\right)\cdot\left(\mathop{\mathbb{E}}_{S\sim F_B}\left[B\cdot\mathbb{1}[B\ge p_n]\right]/p_n\right) \qquad (22)\\
&=s_{n,\textsc{Left}}\cdot b_{n,\textsc{Left}}\cdot p_n.
\end{aligned}$$

Combining inequality (20), (21) and (22), we have

$$\begin{aligned}
\sum_{i=1}^n\sum_{j=1}^n s_ib_j\cdot\max(p_i,p_j)&\ge\sum_{i=1}^n\sum_{j=1}^n\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[\max(S,B)\mathbb{1}[S\in[p_i,p_{i+1})]\mathbb{1}[B\in[p_j,p_{j+1})]\right]\\
&=\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}[\max(S,B)],
\end{aligned}$$

so inequality (17) is satisfied.

Finally, we are only left to show that property (18) holds. For any $t\in[n]$, it follows that

$$\begin{aligned}
\mathcal{W}(\mathcal{I},p_t)&=\mathop{\mathbb{E}}_{S\sim F_S}[S]+\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[(B-S)\cdot\mathbb{1}[S\le p_t\le B]\right]\\
&\ge\mathop{\mathbb{E}}_{S\sim F_S}[S]+\mathop{\mathbb{E}}_{\substack{S\sim F_S\\B\sim F_B}}\left[(B-S)\cdot\mathbb{1}[S<p_t\le B]\right]\\
&=\sum_{i=1}^n E_{s,i}+\mathop{\mathbb{E}}_{B\sim F_B}[B\cdot\mathbb{1}[B\ge p_t]]\cdot\Pr_{S\sim F_S}[S<p_t]-\mathop{\mathbb{E}}_{S\sim F_S}[S\cdot\mathbb{1}[S<p_t]]\cdot\Pr_{B\sim F_B}[B\ge p_t]\\
&\ge\sum_{i=1}^n s_i\cdot p_i+\left(\sum_{j=t+1}^n b_j\cdot p_j+b_{t,\textsc{Left}}\cdot p_t\right)\cdot\left(\sum_{i=1}^{t-1}s_i+s_{t,\textsc{Right}}\right) \qquad (23)\\
&\qquad-\left(\sum_{i=1}^{t-1}s_i\cdot p_i+s_{t,\textsc{Right}}\cdot p_t\right)\cdot\left(\sum_{j=t+1}^n b_j+b_{t,\textsc{Left}}\right)\\
&=\sum_{i=1}^n s_i\cdot p_i+\sum_{i=1}^{t-1}\sum_{j=t+1}^n s_ib_j(p_j-p_i)+\sum_{j=t+1}^n b_j\cdot s_{t,\textsc{Right}}\cdot(p_j-p_t)+\sum_{i=1}^{t-1}s_i\cdot b_{t,\textsc{Left}}\cdot(p_t-p_i)\\
&\ge\sum_{i=1}^n s_i\cdot p_i+\sum_{i=1}^{t-1}\sum_{j=t+1}^n s_ib_j(p_j-p_i).
\end{aligned}$$

where the second inequality follows from the fact that

$$\Pr[B\ge p_t]=\sum_{j=t}^{n-1}(b_{j,\textsc{Left}}+b_{j+1,\textsc{Right}})+\Pr[B\ge p_n]\le\sum_{j=t}^{n-1}(b_{j,\textsc{Left}}+b_{j+1,\textsc{Right}})+b_{n,\textsc{Left}}\le\sum_{j=t+1}^n b_j+b_{t,\textsc{Left}}$$

Therefore, we could see that inequality (18) holds. This finishes our proof. □

With the lemma above, we are ready to give the proof of Lemma 3.2.

Consider the following fixed-price mechanism: Given any instance $\mathscr{I} = (F_S, F_B)$, we first compute the optimal welfare of the instance. Suppose OPT-$\mathscr{W}(\mathscr{I}) = c$, we choose the fixed price p^* from $\{cp_1, \cdots, cp_n\}$ to maximizes the welfare, i.e., $p^* \in \arg\max_{p\in\{cp_1,\cdots,cp_n\}} \mathscr{W}(\mathscr{I}, p)$. In the following, we show that this mechanism is an r^*-approximation to the optimal welfare.

Note that the approximation ratio of our mechanism is independent of c.[5] To keep our analysis clean, we first assume that the instance $\mathscr{I} = (F_S, F_B)$ has optimal welfare 1. The approximation ratio of our mechanism could be written as

$$\min_{\substack{\mathscr{I}=(F_S,F_B)\\ \text{OPT-}\mathscr{W}(\mathscr{I})=1}} \max_{p\in\{p_1,p_2,\cdots,p_n\}} \mathscr{W}(\mathscr{I}, p).$$

Next, we argue that for any instance $\mathscr{I} = (F_S, F_B)$ satisfying OPT-$\mathscr{W}(\mathscr{I}) = 1$, there exists a valid solution $(s_1, \cdots, s_n, b_1, \cdots, b_n, r)$ of $\mathsf{LowerOp}$ such that $r \le \max_{p\in\{p_1,p_2,\cdots,p_n\}} \mathscr{W}(\mathscr{I}, p)$. This immediately implies that r^* is a lower bound of the approximation ratio.

Given an instance $\mathscr{I} = (F_S, F_B)$ s.t. OPT-$\mathscr{W}(\mathscr{I}) = 1$, the solution $(s_1, \cdots, s_n, b_1, \cdots, b_n, r)$ is constructed as follows. Let $(s_1, s_2, \cdots, s_n, b_1, b_2, \cdots, b_n)$ be the set of numbers that satisfies all the properties stated in Lemma B.1. Let r be

$$\max_{t\in[n]} \sum_{i=1}^{n} s_i p_i + \sum_{i=1}^{t-1} \sum_{j=t+1}^{n} s_i b_j (p_j - p_i).$$

We first verify that $(\{s_i\}_{i\in[n]}, \{b_i\}_{i\in[n]}, r)$ is a valid solution of $\mathsf{LowerOp}$. Notice that $\mathbb{E}[S] \le \mathbb{E}[\max(S,B)] = 1$ and $\mathbb{E}[B] \le \mathbb{E}[\max(B,S)] = 1$. Therefore, constraints (2), (3) and (4) directly follows from inequality (12) and (13). What's more, we could see (6) holds by the definition of r.

Now by property (17), we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} s_i b_j \cdot \max(p_i, p_j) \ge \sum_{i=1}^{n}\sum_{j=1}^{n} \mathop{\mathbb{E}}_{\substack{S\sim F_S\\ B\sim F_B}} \Big[\max(S,B)\mathbb{1}[S \in [p_i, p_{i+1})]\mathbb{1}[B \in [p_j, p_{j+1})]\Big] = 1,$$

so constraint (5) is satisfied.

Finally, we are only left to show that the best price in $\{p_1, p_2, \cdots, p_k\}$ must obtain an approximation ratio that is at least r on instance $\mathscr{I}$, i.e., $r \le \max_{p\in\{p_1,p_2,\cdots,p_k\}} \mathscr{W}(\mathscr{I}, p)$. Inequality (18) states that

$$\mathscr{W}(\mathscr{I}, p_t) \ge \sum_{i=1}^{n} s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^{n} s_i b_j (p_j - p_i).$$

Taking maximum over $t \in [n]$, we then get that

$$r = \max_{t\in[n]} \sum_{i=1}^{n} s_i \cdot p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^{n} s_i b_j (p_j - p_i) \le \max_{t\in[n]} \mathscr{W}(\mathscr{I}, p_t)$$

which finishes our proof.

B.3 Proof of Lemma 3.3

In the following, we complete the proof of Lemma 3.3.

[5] The price p^* depends on c, but the approximation ratio to the optimal welfare does not.

The Optimization Problem UpperOp
$$\begin{aligned} \min_{\substack{s_1,s_2\cdots,s_n\\ b_1,b_2,\cdots,b_n,r}} \quad & r \\ \text{s.t.} \quad & s_i, b_i \geq 0 & \forall i \in [n] \\ & \sum_{i=1}^{n} s_i = 1 \quad \text{and} \quad \sum_{i=1}^{n} b_i = 1 \\ & \sum_{i=1}^{n}\sum_{j=1}^{n} s_i b_j \max(p_i, p_j) \geq 1 \\ & \sum_{i=1}^{n} s_i p_i + \sum_{i=1}^{t}\sum_{j=t+1}^{n} s_i b_j (p_j - p_i) \leq r & \forall t \in [n] \end{aligned}$$

For any fixed support $0 = p_1 < p_2 < \cdots < p_n$ and a valid solution $(s_1, s_2, \cdots, s_n, b_1, b_2, \cdots, b_n, r)$, define an instance $\mathscr{I} = (F_S, F_B)$ satisfying

$$S \sim F_S, S = \begin{cases} p_1 + \varepsilon & w.p. \quad s_1 \\ p_2 + \varepsilon & w.p. \quad s_2 \\ \cdots \\ p_n + \varepsilon & w.p. \quad s_n \end{cases} \qquad B \sim F_B, B = \begin{cases} p_1 & w.p. \quad b_1 \\ p_2 & w.p. \quad b_2 \\ \cdots \\ p_n & w.p. \quad b_n \end{cases}$$

where $\varepsilon > 0$ is a constant that small enough.

It is easy to see that both F_S and F_B are valid distributions since the UpperOp requires the non-negativity of s_i, b_i and $\sum_{i=1}^n s_i = \sum_{i=1}^n b_i = 1$. Next, we aim to show that no fixed-price mechanism have an approximation ratio of $r + \varepsilon$ on this instance $\mathscr{I} = (F_S, F_B)$. For any $x \in \mathbb{R}_{\geq 0}$, we could first see that $x < \varepsilon$ would never be a optimal price. Thus let p_i be the largest $p \in \{p_1, p_2, \cdots, p_n\}$ that is not greater than $x - \varepsilon$. Notice that both F_S is a distribution on support $\{p_i + \varepsilon\}_{i \in [n]}$ and F_B is a discrete distribution on support $\{p_i\}_{i \in [n]}$. This means choosing $p_i + \varepsilon$ instead of x would never become worse. Therefore, we could see that the optimal fixed-price mechanism on this instance is simply choosing one $p_t \in \{p_1, p_2, \cdots, p_k\}$ that maximizes $\mathscr{W}(\mathscr{I}, p_t + \varepsilon)$. Again, by the fact that F_S and F_B are discrete distributions, $\mathscr{W}(\mathscr{I}, p_t + \varepsilon)$ could be written as:

$$\begin{aligned} \mathscr{W}(\mathscr{I}, p_t + \varepsilon) &= \sum_{i=1}^{n} (p_i + \varepsilon) s_i + \sum_{i=1}^{t}\sum_{j=t+1}^{n} s_i b_j (p_j - p_i - \varepsilon) \\ &\leq \sum_{i=1}^{n} p_i s_i + \sum_{i=1}^{t}\sum_{j=t+1}^{n} s_i b_j (p_j - p_i) + \varepsilon. \end{aligned}$$

Also notice that the constraints of UpperOp guarantee that

$$\text{OPT-}\mathscr{W}(\mathscr{I}) = \sum_{i=1}^{n}\sum_{j=1}^{n} s_i b_j \max(p_i + \varepsilon, p_j) \geq \sum_{i=1}^{n}\sum_{j=1}^{n} s_i b_j \max(p_i, p_j) \geq 1.$$

Therefore, the approximation ratio of the optimal fixed-price mechanism on instance $\mathscr{I} = (F_S, F_B)$ is upper bounded by

$$\frac{\max_{t \in [n]} \mathscr{W}(\mathscr{I}, p_t + \varepsilon)}{\text{OPT-}\mathscr{W}(\mathscr{I})} \leq \max_{t \in [n]} \sum_{i=1}^{n} p_i s_i + \sum_{i=1}^{t}\sum_{j=t}^{n} s_i b_j (p_j - p_i) + \varepsilon = r + \varepsilon.$$

And this finishes our proof.

B.4 Proof of Lemma 3.4

In this section, we assume that $\varepsilon > 0$ is a small enough constant such that $\varepsilon^2 \ll \varepsilon$.

We first show that, for any $\varepsilon > 0$, there exists a set of support $\{p_1, p_2, \cdots, p_n\}$ such that UpperOp has an optimal value of at most $r^* + \varepsilon$.

As we show before, we could assume that the instance $\mathscr{I} = (F_S, F_B)$ has optimal welfare 1. Thus, the approximation ratio of the optimal fixed-price mechanism is

$$r^* = \min_{\substack{\mathscr{I}=(F_S,F_B)\\ \text{OPT-}\mathscr{W}(\mathscr{I})=1}} \max_{p\in\mathbb{R}} \mathscr{W}(\mathscr{I}, p).$$

Suppose r^* is attained at $\mathscr{I}^* = (F_S^*, F_B^*)$. Now define $n = 1/\varepsilon^4$, and $p_i = i \cdot (1/\varepsilon^2) + \varepsilon/2$ for $i \in [n+1]$. Our idea is to construct a valid solution $\{s_i, b_i\}_{i\in[n+1]}$ by rounding up $\mathscr{I}^*$ to p_i and show that this solution has an objective value that is close to r^*.

Suppose $p_0 = 0$. Now we define

$$s_i = \Pr_{S\sim F_S^*}\left[S \in [(i-1)\varepsilon^2, i\varepsilon^2)\right] \text{ and } b_i = \Pr_{B\sim F_B^*}\left[B \in [(i-1)\varepsilon^2, i\varepsilon^2)\right]$$

for $i \in [n]$. Especially, let

$$s_{n+1} = \mathop{\mathbb{E}}_{S\sim F_S^*}\left[S \cdot \mathbb{1}[S \ge n\varepsilon^2]\right] / \left(n\varepsilon^2\right) \text{ and } b_{n+1} = \mathop{\mathbb{E}}_{B\sim F_B^*}\left[B \cdot \mathbb{1}[B \ge n\varepsilon^2]\right] / \left(n\varepsilon^2\right).$$

Since $\mathbb{E}[S]$ and $\mathbb{E}[B]$ are upper bounded by 1, we could see that $s_{n+1}, b_{n+1} \le \varepsilon^2$. In the last, let $s = \sum_{i=1}^{n+1} s_i$ and $b = \sum_{i=1}^{n+1}$ be the normalization factors. It's also straightforward to see that $s \le \sum_{i=1}^{n} s_i + s_{n+1} \le 1 + \varepsilon^2$. Following the same argument, it also holds that $b \le 1 + \varepsilon^2$. Now define

$$r = \max_{t\in[n]} \sum_{i=1}^{n} (s_i/s) p_i + \sum_{i=1}^{t} \sum_{j=t+1}^{n} (s_i/s)(b_j/b)(p_j - p_i).$$

We aim to verify that $(s_1/s, s_2/s, \cdots, s_{n+1}/s, b_1/b, b_2/b, \cdots, b_{n+1}/b, r)$ is a valid solution of UpperOp.

It is easy to see the non-negativity of s_i, b_i and $\sum_{i=1}^{n+1} s_i/s = \sum_{i=1}^{n+1} b_i/b = 1$. What's more, from the definition of r, we could see the last constraint holds. Now we only need to check the third constraint. For any $i, j \in [n]$, it holds that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{S\sim F_S^*\\ B\sim F_B^*}}\left[\max(S,B)\mathbb{1}[S \in [(i-1)\varepsilon^2, i\varepsilon^2)]\mathbb{1}[B \in [(j-1)\varepsilon^2, j\varepsilon^2)]\right]\\
&\le (\max(p_i, p_j) - \varepsilon/4) s_i b_j
\end{aligned}$$

When one of i, j equals to $n+1$(we can assume $i = n+1$ w.l.o.g.), it is true that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{S\sim F_S^*\\ B\sim F_B^*}}\left[\max(S,B)\mathbb{1}[S \ge n\varepsilon^2]\mathbb{1}[B \in [(j-1)\varepsilon^2, j\varepsilon^2)]\right]\\
&= \mathop{\mathbb{E}}_{S\sim F_S^*}[S \cdot \mathbb{1}[S \ge n\varepsilon^2]] \Pr_{B\sim F_B^*}[B \in [(j-1)\varepsilon^2, j\varepsilon^2)]\\
&= \left(n\varepsilon^2\right) s_{n+1} b_j\\
&\le (p_{n+1} - \varepsilon/4) s_i b_j
\end{aligned}$$

Finally, for the special case that $i = j = n+1$, we could see that

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{S\sim F_S^* \\ B\sim F_B^*}} \left[\max(S,B) \mathbb{1}[S \geq n\varepsilon^2] \mathbb{1}[B \geq n\varepsilon^2] \right] \\
&\leq \mathop{\mathbb{E}}_{\substack{S\sim F_S^* \\ B\sim F_B^*}} \left[BS/\left(n\varepsilon^2\right) \cdot \mathbb{1}[S \geq n\varepsilon^2] \mathbb{1}[B \geq n\varepsilon^2] \right] \\
&= \left(n\varepsilon^2\right) s_{n+1} b_{n+1} \\
&\leq (p_{n+1} - \varepsilon/4) s_{n+1} b_{n+1}
\end{aligned}
$$

Summing up all the inequalities above, we then get that

$$
\mathop{\mathbb{E}}_{\substack{S\sim F_S^* \\ B\sim F_B^*}} [\max(S,B)] \leq \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \left(\max(p_i, p_j) - \varepsilon/4\right) s_i b_j
$$

This implies that

$$
\begin{aligned}
&\sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (s_i/s)(b_j/b) \max(p_i, p_j) \\
&\geq \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} s_i b_j \max(p_i, p_j) \cdot \left(1+\varepsilon^2\right)^{-2} \\
&\geq \left(\mathop{\mathbb{E}}_{\substack{S\sim F_S^* \\ B\sim F_B^*}} [\max(S,B)] + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \varepsilon/4 s_i b_j \right) \cdot (1-\varepsilon^2)^2 \\
&\geq 1 + \varepsilon/8.
\end{aligned}
$$

which means that $(s_1/s, s_2/s, \cdots, s_{n+1}/s, b_1/b, b_2/b, \cdots, b_{n+1}/b, r)$ is truly a valid solution.

Next, we give an upper bound of r. To start with, notice that

$$
\begin{aligned}
\sum_{i=1}^{n+1} s_i p_i &= \sum_{i=1}^{n} \Pr\left[S \in [(i-1)\varepsilon^2, i\varepsilon^2)\right] \cdot \left((i-1)\varepsilon + \varepsilon^2 + \varepsilon/4\right) + \mathbb{E}[S \cdot \mathbb{1}[S \geq n\varepsilon^2]] + s_{n+1} \cdot (\varepsilon^2 + \varepsilon/4) \\
&\leq \sum_{i=1}^{n} \mathbb{E}\left[S \cdot \mathbb{1}\left[S \in [(i-1)\varepsilon^2, i\varepsilon^2)\right]\right] + \mathbb{E}[S \cdot \mathbb{1}[S \geq n\varepsilon^2]] + \sum_{i=1}^{n+1} s_i(\varepsilon^2 + \varepsilon/4) \\
&\leq \mathbb{E}[S] + \varepsilon/2.
\end{aligned} \tag{24}
$$

For the term of gain from trade, it holds that

$$
\begin{aligned}
\sum_{i=1}^{t} \sum_{j=t+1}^{n+1} s_i b_j (p_j - p_i) &= \sum_{i=1}^{t} \sum_{j=t+1}^{n} s_i b_j (j\varepsilon^2 - i\varepsilon^2) + \sum_{i=1}^{t} s_i b_{n+1} \left(n\varepsilon^2 - (i-1)\varepsilon^2\right) \\
&\leq \sum_{i=1}^{t} \sum_{j=t+1}^{n} s_i b_j ((j-1)\varepsilon^2 - i\varepsilon^2) + \sum_{i=1}^{t} s_i b_{n+1} \left(n\varepsilon^2 - i\varepsilon^2\right) + \varepsilon^2(1+\varepsilon^2)^2 \\
&\leq \sum_{i=1}^{t} \sum_{j=t+1}^{n} \mathbb{E}[(B-S) \cdot \mathbb{1}[S \in [(i-1)\varepsilon^2, i\varepsilon^2)]] \mathbb{1}[B \in [(j-1)\varepsilon^2, j\varepsilon^2]] \\
&\quad + \mathbb{E}[(B-S) \cdot \mathbb{1}[S \in [0, t \cdot \varepsilon^2)] \mathbb{1}[B \geq t \cdot \varepsilon^2]] + \varepsilon/2 \\
&\leq \mathbb{E}[(B-S) \cdot \mathbb{1}[S \in [0, t \cdot \varepsilon^2)] \mathbb{1}[B \geq t \cdot \varepsilon^2]] + \varepsilon/2.
\end{aligned} \tag{25}
$$

Combining (24) and (25), we know that for any $t \in [n+1]$,

$$\begin{aligned}
&\sum_{i=1}^{n+1}(s_i/s)p_i + \sum_{i=1}^{t}\sum_{j=t+1}^{n+1}(s_i/s)(b_j/b)(p_j-p_i) \\
&\le \sum_{i=1}^{n+1} s_i p_i + \sum_{i=1}^{t}\sum_{j=t+1}^{n+1} s_i b_j (p_j - p_i) \\
&= \mathbb{E}[S] + \mathbb{E}[(B-S)\cdot \mathbb{1}[S\in[0,t\cdot\varepsilon^2)]\mathbb{1}[B\ge t\cdot\varepsilon^2]] + \varepsilon \\
&\le \mathscr{W}(\mathscr{I}, t\cdot\varepsilon^2) + \varepsilon.
\end{aligned}$$

Taking maximum over $[n+1]$, we then get that

$$r = \max_{t\in[n+1]} \sum_{i=1}^{n+1}(s_i/s)p_i + \sum_{i=1}^{t}\sum_{j=t}^{n+1}(s_i/s)(b_j/b)(p_j-p_i) \le \max_{t\in[n+1]} \mathscr{W}(\mathscr{I}, t\cdot\varepsilon^2) + \varepsilon \le r^* + \varepsilon.$$

This means that the optimal value of UpperOp with respect to $\{p_i\}_{i\in[n+1]}$ is at most $r^*+\varepsilon$, and this finishes our proof.

Next, we aim to show that for any $\varepsilon > 0$, there exists $0 = p_0 < p_1 < p_2 < \cdots p_n$ such that LowerOp has an optimal value of at least $r^* - \varepsilon$. Now define $n = \lceil \varepsilon^{-6} \rceil + 1$, and $p_i = (i-1)\cdot\varepsilon^3$ for $i \in [n]$. Let $(s_1,\cdots,s_n,b_1,\cdots,b_n,r)$ be the optimal solution of the optimization problem LowerOp with respect to $\{p_i\}_{i\in n}$. It is equivalent to show that there exists an instance $\mathscr{I} = (F_S, F_B)$ such that the optimal approximation ratio of $\mathscr{I}$, i.e. $\max_{x\in\mathbb{R}} \frac{\mathscr{W}(\mathscr{I},x)}{\text{OPT-}\mathscr{W}(\mathscr{I})}$, is at most $r+\varepsilon$.

We construct the instance as follows. Let $n' = n + \lceil \frac{4}{\varepsilon} \rceil$, and $s_i = b_i = 0$ for $n < i \le n'$. Now define $\{s'_i\}$ where

$$s'_j = \sum_{i=\max\left(1, j-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{j} s_i \Big/ \left\lceil \frac{4}{\varepsilon} \right\rceil.$$

It follows that

$$\sum_{j=1}^{n'} s'_j = \sum_{j=1}^{n}\sum_{i=\max\left(1, j-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{j} s_i \Big/ \left\lceil \frac{4}{\varepsilon} \right\rceil = \sum_{i=1}^{n'} s_i.$$

Let $s = \sum_{i=1}^{n'} s'_i$ and $b = \sum_{i=1}^{n'} b_i$ be the normalization factors. LowerOp guarantees that $1 \le s, b \le 1+\varepsilon^3$. What's more, we could also see that

$$s'_j = \sum_{i=\max\left(1, j-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{j} s_i \Big/ \left\lceil \frac{4}{\varepsilon} \right\rceil \le (1+\varepsilon^3)\left\lceil \frac{4}{\varepsilon} \right\rceil \le \varepsilon/3.$$

holds for all $j \in [n']$.

Consider the following instance $\mathscr{I} = (F_S, F_B)$:

$$S \sim F_S, S = \begin{cases} p_1 + \varepsilon^4 & w.p. \quad s'_1/s \\ \cdots & \\ p_{n'} + \varepsilon^4 & w.p. \quad s'_{n'}/s \end{cases} \qquad B \sim F_B, B = \begin{cases} p_1 & w.p. \quad b_1/b \\ \cdots & \\ p_{n'} & w.p. \quad b_{n'}/b \end{cases}$$

First, it is straight forward to verify this is a valid distribution. We first calculate OPT-$\mathcal{W}(\mathcal{I})$:

$$
\begin{aligned}
\text{OPT-}\mathcal{W}(\mathcal{I}) &= \sum_{i=1}^{n'}\sum_{j=1}^{n'} \max(p_i+\varepsilon^4, p_j)\cdot(s'_i/s)(b_j/b) \\
&\geq \sum_{i=1}^{n'}\sum_{j=1}^{n'} \max(p_i, p_j)\cdot(s'_i/s)(b_j/b) - \varepsilon^4 \\
&\geq \sum_{i=1}^{n'}\sum_{j=1}^{n'} \max(p_i, p_j)\left(\sum_{k=\max\left(1, i-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{i}\left(s_k/\left\lceil\frac{4}{\varepsilon}\right\rceil\right)/s\right)\cdot(b_j/b) - \varepsilon^4 \\
&\geq \sum_{i=1}^{n'}\sum_{j=1}^{n'} \max(p_i, p_j)\cdot s_i b_j/(bs) - \varepsilon^4 \\
&\geq (1+\varepsilon^3)^{-2} - \varepsilon^4 \\
&\geq 1-\varepsilon^2
\end{aligned}
$$

Now consider the optimal fixed-price mechanism for the instance. As we have shown in the proof of Lemma 3.3, the optimal mechanism only need to choose price from the support of the discrete distribution. This implies that

$$
\max_{p\in\mathbb{R}} \mathcal{W}(\mathcal{I}, p) = \max_{t\in[n]} \sum_{i=1}^{n'}(p_i+\varepsilon^4)(s'_i/s) + \sum_{i=1}^{t}\sum_{j=t+1}^{n'}(s'_i/s)(b_j/b)(p_j-p_i-\varepsilon^4)
$$

For any $t\in[n]$, one could see that

$$
\begin{aligned}
\sum_{i=1}^{n'} p_i s'_i &= \sum_{i=1}^{n'}\left(\sum_{k=\max\left(1, i-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{i}\left(s_k/\left\lceil\frac{4}{\varepsilon}\right\rceil\right)\right)p_i \\
&\leq \sum_{i=1}^{n'}\left(\sum_{k=\max\left(1, i-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{i}\left(s_k/\left\lceil\frac{4}{\varepsilon}\right\rceil\right)\cdot\left(p_k+\left\lceil\frac{4}{\varepsilon}\right\rceil\cdot\varepsilon^3\right)\right) \\
&\leq \sum_{i=1}^{n'}(p_i+5\varepsilon^2)s_i
\end{aligned}
\tag{26}
$$

For the term of gain from trade, it follows that

$$
\begin{aligned}
\sum_{i=1}^{t}\sum_{j=t+1}^{n'} s'_i b_j(p_j-p_i) &= \sum_{i=1}^{t-1}\sum_{j=t+1}^{n'} s'_i b_j(p_j-p_i) + \sum_{j=t+1}^{n'} s'_t b_j(p_j-p_i) \\
&\leq \sum_{i=1}^{t-1}\sum_{j=t+1}^{n'}\left(\sum_{k=\max\left(1, i-\lceil\frac{4}{\varepsilon}\rceil+1\right)}^{i}\left(s_k/\left\lceil\frac{4}{\varepsilon}\right\rceil\right)\right)b_j(p_j-p_i) + s'_t\sum_{j=1}^{n'} b_j p_j \\
&\leq \sum_{i=1}^{t-1}\sum_{j=t+1}^{n'} s_i b_j(p_j-p_i) + \varepsilon/3.
\end{aligned}
\tag{27}
$$

where we use the fact that $s_k b_j(p_j-p_i) \leq s_k b_j(p_j-p_k)$ for $j > i \geq k$, $s'_t \leq \varepsilon/3$, and $\sum_{i=1}^{n'} b_j p_j \leq 1$.

Again by combining the two inequalities above, we know that

$$
\begin{aligned}
&\sum_{i=1}^{n'}(p_i+\varepsilon^4)(s_i'/s)+\sum_{i=1}^{t}\sum_{j=t+1}^{n'}(s_i'/s)(b_j/b)(p_j-p_i-\varepsilon^4)\\
&\le \sum_{i=1}^{n'} p_i s_i' + \sum_{i=1}^{t}\sum_{j=t+1}^{n'} s_i' b_j(p_j-p_i)+\varepsilon^4\\
&\le \sum_{i=1}^{n'}(p_i+5\varepsilon^2)s_i+\sum_{i=1}^{t-1}\sum_{j=t+1}^{n'} s_i b_j(p_j-p_i)+\varepsilon/3\\
&\le \sum_{i=1}^{n'} p_i s_i+\sum_{i=1}^{t-1}\sum_{j=t+1}^{n'} s_i b_j(p_j-p_i)+\varepsilon/2
\end{aligned}
$$

where we apply (26) and (27) in the second inequality.

We could see that the optimal solution $(s_1, s_2, \cdots, s_n, b_1, b_2, \cdots, b_n, r)$ of LowerOp must satisfy that $r = \max_{t\in[n]} \sum_{i=1}^{n} p_i s_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^{n} s_i b_j(p_j - p_i)$.

Therefore, taking the maximum over $t \in [n']$, we then get that

$$
\begin{aligned}
\max_{p\in\mathbb{R}} \mathscr{W}(\mathscr{I}, p) &= \max_{t\in[n']}\sum_{i=1}^{n'}(p_i+\varepsilon^4)(s_i'/s)+\sum_{i=1}^{t}\sum_{j=t+1}^{n'}(s_i'/s)(b_j/b)(p_j-p_i-\varepsilon^4)\\
&\le \max_{t\in[n']}\sum_{i=1}^{n'} p_i s_i+\sum_{i=1}^{t-1}\sum_{j=t+1}^{n'} s_i b_j(p_j-p_i)+\varepsilon/2\\
&= \max_{t\in[n]}\sum_{i=1}^{n} p_i s_i+\sum_{i=1}^{t-1}\sum_{j=t+1}^{n} s_i b_j(p_j-p_i)+\varepsilon/2\\
&= r+\frac{\varepsilon}{2}.
\end{aligned}
$$

where the second equation follows from $s_i = b_i = 0$ when $i > n$.

Therefore, on instance $\mathscr{I}$, it holds that

$$
\frac{\max_{p\in\mathbb{R}} \mathscr{W}(\mathscr{I}, p)}{\text{OPT-}\mathscr{W}(\mathscr{I})} \le \frac{r+\varepsilon/2}{1-\varepsilon^2} \le r+\varepsilon.
$$

And this completes our proof.

C Proof of Theorem 4.1

We start with the case when we only know $\mathbb{E}[S]$.

We consider discrete distribution P_S. Suppose $x \sim P_S$ equals to p_i with probabiltiy w_i for $i \in [n]$ where $\sum_{i=1}^{n} w_i = 1$. This means that our mechanism would choose p_i with probability w_i and use $p_i \cdot \mathbb{E}_{S\sim F_S}[S]$ as the price. Fixing $\{p_i, w_i\}_{i\in[n]}$, we claim that the optimal solution of the following program lower bounds the approximation ratio of this mechanism.

$$\min_{\substack{s_1,s_2\cdots,s_n\\ b_1,b_2,\cdots,b_n}} \quad \frac{\sum_{t=1}^n w_t\left(\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i)\right)}{\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j)}$$

$$\text{s.t.} \quad s_i, b_i \geq 0 \qquad \forall i \in [n] \tag{28}$$

$$\sum_{i=1}^n s_i \geq 1 \quad \text{and} \quad \sum_{i=1}^n s_i \leq 1 + \frac{1}{p_n} \quad \text{and} \quad \sum_{i=1}^{n-1} s_i \leq 1 \tag{29}$$

$$\sum_{i=1}^n b_i \geq 1 \quad \text{and} \quad \sum_{i=1}^{n-1} b_i \leq 1 \tag{30}$$

$$\sum_{i=1}^n s_i \cdot p_i = 1 \tag{31}$$

Similar to the proof of Lemma 3.2, we could assume that $\mathbb{E}[S] = 1$ without loss of generality. The approximation ratio r^* of our mechanism could be written as

$$r^* = \min_{\substack{\mathcal{I}=(F_S,F_B)\\ \mathbb{E}[S]=1}} \frac{\sum_{i=1}^n w_i \cdot \mathcal{W}(\mathcal{I}, p_i)}{\text{OPT-}\mathcal{W}(\mathcal{I})}.$$

Suppose r^* is attained at $\mathcal{I}^* = \left(F_S^*, F_B^*\right)$. Applying Lemma B.1 with $\mathcal{I}^*$, we know that there exists $\{s_1, s_2, \cdots, s_n, b_1, b_2, \cdots, b_n\}$ satisfying all the properties in the statement. Notice that $\mathbb{E}_{S\sim F_S^*}[S] = 1$. Therefore, we could directly verify that constraints (28), (29), (30) and (31) are satisfied by all the properties in Lemma B.1. Again, by Lemma B.1, it holds that $\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i) \leq \mathcal{W}(\mathcal{I}', p_t)$ for all $t \in [n]$ and $\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j) \geq \text{OPT-}\mathcal{W}(\mathcal{I}')$. This implies that

$$\frac{\sum_{t=1}^n w_t\left(\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i)\right)}{\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j)} \leq \frac{\sum_{i=1}^n w_i \cdot \mathcal{W}(\mathcal{I}^*, p_i)}{\text{OPT-}\mathcal{W}(\mathcal{I}^*)} = r^*.$$

Since $\{s_i, b_i\}_{i\in[n]}$ is a valid solution of the optimization problem above, we then show that it lower bounds the approximation ratio.

Finally, we solve this optimization problem numerically and show that there exists $\{p_i, w_i\}_{i\in[n]}$ such that the optimal solution is at least by 0.65. The details of the numerical result could be found at our GitHub repository.

Now we turn to the case when we know $\mathbb{E}[B]$. The proof uses similar ideas and is almost identical.

Consider the following mechanism: it picks p_i with probability w_i for $i \in [n]$ where $\sum_{i=1}^n w_i = 1$, and chooses $p_i \cdot \mathbb{E}[B]$ as the price. Again, for fixed $\{p_i, w_i\}_{i\in[n]}$, we aim to show that the following optimization problem give a lower bound of the approximation ratio.

$$\min_{\substack{s_1,s_2\cdots,s_n\\ b_1,b_2,\cdots,b_n}} \quad \frac{\sum_{t=1}^n w_t\left(\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i)\right)}{\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j)}$$

$$\text{s.t.} \quad s_i, b_i \geq 0 \qquad \forall i \in [n] \tag{32}$$

$$\sum_{i=1}^n b_i \geq 1 \quad \text{and} \quad \sum_{i=1}^n b_i \leq 1 + \frac{1}{p_n} \quad \text{and} \quad \sum_{i=1}^{n-1} b_i \leq 1 \tag{33}$$

$$\sum_{i=1}^n s_i \geq 1 \quad \text{and} \quad \sum_{i=1}^{n-1} s_i \leq 1 \tag{34}$$

$$\sum_{i=1}^n b_i \cdot p_i = 1 \tag{35}$$

Without loss of generality, we could assume that $\mathbb{E}[B] = 1$. Therefore, the approximation ratio of the mechanism is exactly

$$r^* = \min_{\substack{\mathscr{I}=(F_S,F_B)\\ \mathbb{E}[B]=1}} \frac{\sum_{i=1}^n w_i \cdot \mathscr{W}(\mathscr{I}, p_i)}{\text{OPT-}\mathscr{W}(\mathscr{I})}.$$

Suppose r^* is attained at $\mathscr{I}^* = \left(F_S^*, F_B^*\right)$. Again by Lemma B.1 , we could discretize the instance $\mathscr{I}^*$ to $\{s_1, s_2, \cdots, s_n, b_1, b_2, \cdots, b_n\}$ that satisfies all the properties in the lemma. Since $\mathbb{E}_{B\sim F_B^*}[B] = 1$, it is easy to verify that constraints (32), (33), (34) and (35) holds due to Lemma B.1. Again, by Lemma B.1, it holds that $\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i) \le \mathscr{W}(\mathscr{I}', p_t)$ for all $t \in [n]$ and $\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j) \ge$ OPT-$\mathscr{W}(\mathscr{I}')$. This implies that

$$\frac{\sum_{t=1}^n w_t \left(\sum_{i=1}^n s_i p_i + \sum_{i=1}^{t-1}\sum_{j=t+1}^n s_i b_j (p_j - p_i)\right)}{\sum_{i=1}^n \sum_{j=1}^n s_i b_j \max(p_i, p_j)} \le \frac{\sum_{i=1}^n w_i \cdot \mathscr{W}(\mathscr{I}^*, p_i)}{\text{OPT-}\mathscr{W}(\mathscr{I}^*)} = r^*.$$

Since $\{s_i, b_i\}_{i\in[n]}$ is a valid solution of the optimization problem above, this means that the optimization problem is a lower bound of the approximation ratio.

Finally, we solve this optimization problem numerically and find a set of numbers $\{p_i, w_i\}_{i\in[n]}$ such that the optimal solution is at least by 0.65. The details of the numerical result could be found at our GitHub repository.

D Missing Proofs in Section 5

D.1 Mechanisms over Buyer's information

In the sample setting, we only consider mechanisms over seller's information. We do not consider quantile or order statistics mechanisms over buyer's information since it is impossible to get any constant approximation with these family of mechanisms.

Theorem D.1. *No quantile mechanism over buyer's distribution or order statistic mechanism over only buyer's samples can achieve a constant fraction of the optimal welfare.*

Proof. We first show that there is no constant approximation quantile mechanism Q over buyer's distribution. Remind that F_B and F_S respectively stand for the distribution of the buyer and the seller, and we will also use Q to denote the corresponding distribution over the buyer's quantile.

To start with, we can assume that distribution Q does not have point mass at 1. That's because if we set the 1-quantile of the buyer's distribution, i.e. $F_B^{-1}(1)$, as the price, we have $\Pr_{B\sim F_B}[B \ge F_B^{-1}(1)] = 0$. This means that the trade will never happen under such price and thus this price will not increase the welfare. Therefore, if we move this probability mass to other values, the welfare and also the approximation ratio will not decrease, and we prove that this assumption is with out loss of generality.

Now, for an arbitarily small $\varepsilon > 0$, we will show that there is no ε-approximation quantile mechanism over buyer's distribution. For any quantile mechanism Q over buyer's distribution, we construct the following set

$$\mathscr{X} = \{t \in [0,1] \mid \Pr_{x\sim Q}[x \ge t] \le \varepsilon/2\}$$

Since there is no point mass at 1, this set will contain some $t \in \mathscr{X}$ and $t \ne 1$. Consider the following instance $\mathscr{I} = (F_S, F_B)$:

$$B \sim F_B, B = \begin{cases} 0 & w.p. \quad t \\ H & w.p. \quad 1-t \end{cases} \qquad S \sim F_S, S = \varepsilon/2 \qquad w.p. \quad 1$$

where $H = \frac{1}{1-t}$ is a large enough number.

In this instance, the intuition is that all the welfare is hide at some very little probability of the buyer, and we must make sure that the trade is very likely to happen when the buyer has a very high value. However, since we don't know the value of the seller, it is hard for us to make sure that $p \geq S$ which means that this trade will not happen.

Remind that we define OPT-$\mathcal{W}(\mathcal{I})$ as the optimal welfare, i.e., $\mathbb{E}_{S\sim F_S, B\sim F_B}[\max(S,B)]$ against instance $\mathcal{I} = (F_S, F_B)$ and $\mathcal{W}(Q,\mathcal{I})$ as the welfare of mechanism Q against instance $\mathcal{I}$. Formally speaking, we have that

$$\text{OPT-}\mathcal{W}(\mathcal{I}) \geq H \cdot (1-t) = 1$$

and also

$$\begin{aligned} \mathcal{W}(Q,\mathcal{I}) &= \mathbb{E}[S] + \mathop{\mathbb{E}}_{p\sim F_B^{-1}(Q)}[(B-S)\cdot \mathbb{1}[B \geq p \geq S]] \\ &\leq \varepsilon/2 + \mathop{\mathbb{E}}_{p\sim F_B^{-1}(Q)}[B \cdot \mathbb{1}[p \geq S]] \\ &\leq \varepsilon/2 + H\cdot(1-t)\cdot \varepsilon/2 = \varepsilon \end{aligned}$$

where the last inequality holds since $p = F_B^{-1}(x) \geq S$ is equivlent to $x \geq t$ where x is drawn from Q, and this happends w.p. at most $\varepsilon/2$ by the definition of t. So for every distribution Q over buyer's quantile, we find an instance $\mathcal{I}$ so that $\mathcal{W}(Q,\mathcal{I}) \leq \varepsilon \cdot \text{OPT-}\mathcal{W}(\mathcal{I})$, which completes the first part of our proof.

Next we aim to show that for any $\varepsilon > 0, N > 0$, there is no ε-approximation mechanisms using only N samples from the buyer.

First, for any mechanisms $\mathcal{M}$ using N samples from the buyer, it can be formallized as a mapping

$$f : \mathbb{R}_{\geq 0}^N \mapsto \Delta(\mathbb{R}_{\geq 0})$$

where $f(x_1, x_2, \cdots, x_N)$ stands for the distribution of price selected by this mechanism after receiving N samples $(x_1, x_2, \cdots, x_N)$. Let D be $f(0,0,\ldots,0)$, which is the distribution of the price if this mechanism sees N samples all with value 0. Similarly, we consider the following set:

$$\mathcal{H}' = \{t \in \mathbb{R}_{\geq 0} \mid \Pr_{x\sim D}[x \geq t] \leq \varepsilon/2\}$$

Again we know this set is non-empty, so let t be any real positive number in the set $\mathcal{H}'$. Therefore, we could construct an instance $\mathcal{I} = (F_S, F_B)$ satisfying

$$B \sim F_B, B = \begin{cases} 0 & w.p. \quad (1-\varepsilon/4)^{1/N} \\ H & w.p. \quad 1-(1-\varepsilon/4)^{1/N} \end{cases} \qquad S \sim F_S, S = t+1 \qquad w.p. \quad 1$$

where $H > \frac{t+1}{\varepsilon/4\cdot(1-(1-\varepsilon/4)^{1/N}}$ is a large enough number.

In this instance, we can see that with just N samples, no mechanism can distinguish this instance with another instance whose buyer always have a value of 0. Therefore, it can not get the welfare hidden at the buyer. Formally speaking:

$$\text{OPT-}\mathcal{W}(\mathcal{I}) \geq H \cdot \left(1-(1-\varepsilon/4)^{1/N}\right) > \frac{t+1}{\varepsilon/4}$$

To calculate $\mathscr{W}(\mathscr{M},\mathscr{I})$, we consider the case when all the samples are zero and the case when there is at least one non-zero number in the samples. In the latter case, the probability that at least one sample is non-zero is at most $1-\left((1-\varepsilon/4)^{1/N}\right)^N = \varepsilon/4$ which is negligible. In the former case, since $\Pr_{p\sim f(0,0,\ldots,0)}[p \ge t] \le \varepsilon/2$, the trade happends w.p. at most $\varepsilon/2$. Therefore, we could expand $\mathscr{W}(\mathscr{M},\mathscr{I})$ into:

$$\begin{aligned}\mathscr{W}(\mathscr{M},\mathscr{I}) &\le \mathop{\mathbb{E}}_{p\sim f(0,0,\ldots,0)}[S+(B-S)\cdot\mathbb{1}[B\ge p\ge S]]\cdot\Pr[\text{All } N \text{ samples are } 0]\\ &\quad + \text{OPT-}\mathscr{W}(\mathscr{I})\cdot\Pr[\text{at least 1 sample is not 0}]\\ &\le (t+1+\mathbb{E}[B\cdot\mathbb{1}[p\ge S]])\cdot 1 + \text{OPT-}\mathscr{W}(\mathscr{I})\cdot(\varepsilon/4)\\ &\le \text{OPT-}\mathscr{W}(\mathscr{I})\cdot(\varepsilon/4+\varepsilon/2+\varepsilon/4)\\ &= \varepsilon\cdot\text{OPT-}\mathscr{W}(\mathscr{I})\end{aligned}$$

where the second inequality holds since $t+1 \le (\varepsilon/4)\cdot\text{OPT-}\mathscr{W}(\mathscr{I})$, $\Pr_{p\sim f(0,0,\ldots,0)}[p\ge S]\le\varepsilon/2$ and $\mathbb{E}_{B\sim F_B}[B] \le \mathbb{E}_{S\sim F_S,B\sim F_B}[\max(S,B)] = \text{OPT-}\mathscr{W}(\mathscr{I})$.

And this finishes our proof.

D.2 Proof of Lemma 5.1

The proof here is quite straight forward. As we show in Section 5.1.1, each order statistic mechanism corresponds to a quantile mechanism. Thus $\mathscr{C}(\mathscr{P}(Q))$ is exactly the approximation ratio of the order statistic mechanism Q. What's more, we could see that the Δ_N enumerates all possible order statistic mechanisms with N samples. Therefore, this directly implies that $\text{argmax}_{Q\in\Delta_N}\mathscr{C}(\mathscr{P}(Q))$ is the optimal order statistic mechanism with N samples. □

D.3 Proof of Lemma 5.2

Fix an instance $\mathscr{I}=(F,F)$, recall that S and B are the random variables respectively indicating the value of the seller and the buyer. Define ALG to be the random variable which indicates the welfare of our mechanism in the realization, which is $S+(B-S)\cdot\mathbb{1}[B\ge p\ge S]$ where p is the price chosen by our quantile mechanism Q. Similarly, let OPT be the random variable which indicates the optimal welfare in the realization, which is $\max(B,S)$.

To prove Lemma 5.2, we introduce the following lemma.

Lemma D.1. *For any quantile mechanism Q, let* ALG *and* OPT *respectively be the random variables indicating the welfare of the mechanism Q and the optimal welfare in the realization. Let r be*

$$\min_{\mathscr{I}=(F,F)}\inf_{x\in[0,1)}\frac{\Pr[\text{ALG}\ge F^{-1}(x)]}{\Pr[\text{OPT}\ge F^{-1}(x)]}$$

where $F(x)$ is the cumulative distribution function of distribution F, and $F^{-1}(x)$ is the quantile function. The quantile mechanism Q is at least r-approximate.

Proof. We have

$$\Pr[\text{ALG}\ge F^{-1}(x)] \ge r\cdot\Pr[\text{OPT}\ge F^{-1}(x)]$$

for all $x\in[0,1]$ and quantile function $F^{-1}(x)$.

Without loss of generality, we could assume the distribution has a support over $[0,a]$. Notice that since we assume the distribution is continuous w.l.o.g. in the sample setting, $F^{-1}(x)$ is a continuous and increasing function over $[0,1]$ and $F^{-1}(0)=0, F^{-1}(1)=a$, so we have

$$\begin{aligned}
\mathscr{W}(\mathscr{I}, Q) &= \mathbb{E}[\text{ALG}] \\
&= \int_0^a \Pr[\text{ALG} \geq x]\, dx \\
&= \int_0^1 \Pr[\text{ALG} \geq F^{-1}(z)]\, dF^{-1}(z) \\
&\geq \int_0^1 r \cdot \Pr[\text{OPT} \geq F^{-1}(z)]\, dF^{-1}(z) \\
&= r \cdot \int_0^a \Pr[\text{OPT} \geq x]\, dx = r \cdot \mathbb{E}[\text{OPT}] \\
&= r \cdot \text{OPT-}\mathscr{W}(\mathscr{I})
\end{aligned}$$

holds for any instance $\mathscr{I} = (F, F)$, which implies that quantile mechanism Q is at least r-approximate.

□

With Lemma D.1, we are able to give a lower bound of approximation ratio for any quantile function Q.

Fixing the buyer and seller's distribution F, we only need to calculate the term $\Pr[\text{ALG} \geq F^{-1}(x)]$ and $\Pr[\text{OPT} \geq F^{-1}(x)]$. The event $\text{OPT} \geq F^{-1}(x)$ happens if and only if either B or S is greater than $F^{-1}(x)$. Thus,

$$\Pr[\text{OPT} \geq F^{-1}(x)] = 1 - x^2 \tag{36}$$

The event $\text{ALG} \geq F^{-1}(x)$ happens if and only if one of the following conditions is satisfied:

- $S \geq F^{-1}(x)$
- $p \leq F^{-1}(x)$, $S \leq p$ and $B \geq F^{-1}(x)$. Here $S \leq p \leq B$, thus the trade takes place, and $B \geq F^{-1}(x)$.
- $p > F^{-1}(x)$, $S \leq F^{-1}(x)$ and $B \geq p$. Since $S \leq p \leq B$, the seller trades the item to the buyer,and we have $B \geq p \geq F^{-1}(x)$.

Note that these three events are disjoint, so we could calculate the probability for each event to happen and add them up.

For the first event, $\Pr[S \geq F^{-1}(x)] = 1 - x$.

For the second event, we just enumerate the quantile of p. Suppose the quantile of p is t, which means that $F^{-1}(t) = p$. Then, we have $\Pr[S \leq p] = t$ and $\Pr[B \geq F^{-1}(x)] = 1 - x$. Thus, this event takes place w.p. $\int_{[0,x]} t(1-x)\, dQ(t)$ where $Q(t)$ is the c.d.f. of distribution Q.

For the third event, we use the same idea. Suppose the quantile of p is t, we have $\Pr[S \leq F^{-1}(x)] = x$ and $\Pr[B \geq p] = 1 - t$. Therefore, this event happens w.p. $\int_{(x,1]} (1-t)x\, dQ(t)$.

By adding the terms above up, we have:

$$\Pr[\text{ALG} \geq F^{-1}(x)] = \int_{[0,x]} t(1-x)\, dQ(t) + \int_{(x,1]} (1-t)x\, dQ(t) + (1-x) \tag{37}$$

Therefore, combining Lemma D.1 and Equation (36) and (37), we have that for any quantile mechanism Q with c.d.f. $Q(x)$, the minimum of the following optimization problem lower bounds the approximation ratio of the quantile mechanism Q.

$$\begin{aligned}
&\min_{\mathscr{I}=(F,F)} \inf_{x \in [0,1)} \frac{\int_0^x q(t) \cdot t(1-x)\, dt + \int_x^1 q(t) \cdot (1-t)x\, dt + (1-x)}{1-x^2} \\
&= \inf_{x \in [0,1)} \frac{\int_0^x q(t) \cdot t(1-x)\, dt + \int_x^1 q(t) \cdot (1-t)x\, dt + (1-x)}{1-x^2}
\end{aligned}$$

where the equality holds since we could see that the term is independent from $F(x)$.

We now left to show that the approximation ratio of Q is also upper bounded by r. It suffices to show that for any $\varepsilon > 0$ there exists some instance $\mathscr{I} = (F, F)$ such that

$$\mathscr{W}(Q, I) \le (r+\varepsilon)\cdot \text{OPT-}\mathscr{W}(\mathscr{I})$$

First, Recall Equation (36) and (37). We could see that both the term $\Pr[\text{ALG} \ge F^{-1}(x)]$ and the term $\Pr[\text{OPT} \ge F^{-1}(x)]$ are independent of the distribution F. Thus,

$$r = \min_F \inf_{x\in[0,1)} \frac{\Pr[\text{ALG} \ge F^{-1}(x)]}{\Pr[\text{OPT} \ge F^{-1}(x)]} = \inf_{x\in[0,1)} \frac{\int_0^x q(t)\cdot t(1-x)\,dt + \int_x^1 q(t)\cdot(1-t)x\,dt + (1-x)}{1-x^2}$$

Suppose the optimum of $\inf_{x\in[0,1)} \frac{\int_0^x q(t)\cdot t(1-x)\,dt+\int_x^1 q(t)\cdot(1-t)x\,dt+(1-x)}{1-x^2}$ is attained at x^*. Consider the following instance $\mathscr{I} = (F, F)$ satisfying

$$v \sim F, v = \begin{cases} U[0, r\cdot(1-x^*)\cdot\varepsilon/2] & w.p. \quad x^* \\ U[1, 1+\varepsilon/2] & w.p. \quad 1-x^* \end{cases}$$

Notice that in this instance, The event $\mathbb{1}[\text{ALG} \ge F^{-1}(x^*)]$ is equivalent to $\mathbb{1}\left[\text{ALG} \in [1, 1+\varepsilon/2]\right]$. Such argument also holds for OPT. Thus, we can see that

$$\begin{aligned}
\mathscr{W}(Q,\mathscr{I}) &= \mathbb{E}[\text{ALG}] \\
&\le \Pr[\text{ALG} \in [1, 1+\varepsilon/2]]\cdot(1+\varepsilon/2) + \Pr[\text{ALG} \in [0, r\cdot(1-x^*)\cdot\varepsilon/2]]\cdot(r\cdot(1-x^*)\cdot\varepsilon/2) \\
&\le \Pr[\text{ALG} \ge F^{-1}(x^*)]\cdot(1+\varepsilon/2) + r\cdot(1-x^*)\cdot\varepsilon/2 \\
&= r\cdot\Pr[\text{OPT} \ge F^{-1}(x^*)]\cdot(1+\varepsilon/2) + r\cdot(1-x^*)\cdot\varepsilon/2 \\
&\le r\cdot\mathbb{E}[\text{OPT}]\cdot(1+\varepsilon/2) + r\cdot(1-x^*)\cdot\varepsilon/2 \\
&\le (r+\varepsilon)\cdot\mathbb{E}[\text{OPT}] = (r+\varepsilon)\cdot\text{OPT-}\mathscr{W}(\mathscr{I})
\end{aligned}$$

where the second equation uses the fact that $r\cdot\Pr[\text{OPT} \ge F^{-1}(x^*)] = \Pr[\text{ALG} \ge F^{-1}(x^*)]$ since r is attained at x^*.

Since the above holds for every $\varepsilon > 0$, this completes our proof of Lemma 5.2.

D.4 Proof of Theorem 5.1

The proof of Theorem 5.1 is directly a combination of Lemma 5.1 and Lemma 5.2. From Lemma 5.2, we know that the approximation ratio of a quantile mechanism Q with c.d.f. $Q(x)$ is exactly

$$\inf_{x\in[0,1)} \frac{\int_{[0,x]} q(t)t(1-x)\,dt + \int_{(x,1]} q(t)(1-t)x\,dt + (1-x)}{1-x^2}.$$

where we use $q(x)\,dx$ instead of $dQ(x)$ since we have a continuous probability density function.

We could see that the set of all distributions over $[N]$ is actually $\left\{\{P_i\}_{i\in[n]} \mid P_i \ge 0 \text{ and } \sum_{i=1}^n P_i = 1\right\}$ where distribution P would choose i w.p. P_i. What's more, the probabiltiy density function of the quantile of such order statistic mechanism is exactly $p(x) = \sum_{i=1}^n P_i f_N^i(x)$ where $f_N^i(x) = N\binom{n-1}{i-1}\cdot x^{i-1}\cdot(1-x)^{N-i}$. Therefore, combining it with Lemma 5.1, it follows that Theorem 5.1 holds.

D.5 Analysis of Empirical Risk Minimization Mechanism

In this section, we give the upper bounds of the Empirical Risk Minimization mechanism.

When $N = 1$, the empirical distribution is a one-point distribution, so any price $x \in \mathbb{R}_{\geq 0}$ is optimal. Therefore, we consider the following instance $\mathscr{I} = (F, F)$ s.t.

$$x \sim F, x = \begin{cases} 0 & w.p. \quad 1 - 1/H \\ H & w.p. \quad 1/H \end{cases}$$

Since any price is optimal for ERM, we can assume that it will always select $H + 1$ so that the trade will never take place. Therefore taking $H \to \infty$ in this instance, the approximation ratio tends to 0.5.

When $N = 2$, the empirical distribution is a two-point distribution. Suppose the two samples are $X_1 \leq X_2$, we can see that any price in $[X_1, X_2)$ is optimal for this empirical distribution. Thus, we can assume that it will always select X_1. So, ERM is equivalent to an order statistic mechanism that always selects the smallest sample. Therefore, we can solve the following optimization problem:

$$\inf_{x \in [0,1)} \frac{\frac{1}{3}x^3 - x^2 - \frac{1}{3}x + 1}{1 - x^2} = \frac{2}{3}$$

Then, by Lemma 5.2, we know that there exists an instance such that ERM achieves exactly $\frac{2}{3}$ approximation.

Now we are in the case that $N = 3$. By some calculations, we know that the second smallest sample will always be an optimal choice. Therefore, ERM is equivalent to an order statistic mechanism that always selects the second smallest sample when $N = 3$. Similarly, we can calculate that

$$\inf_{x \in [0,1)} \frac{\frac{1}{2}x^4 - x^3 - \frac{1}{2}x + 1}{1 - x^2} = \frac{3}{4}$$

Applying Lemma 5.2 again, there is an instance such that ERM has an approximation ratio of exactly $\frac{3}{4}$.

Our proof strategy changes when the number of samples is greater than 5. We consider a particular instance $\mathscr{I} = (F, F)$ and calculate the performance of ERM on such instance.

$$x \sim F, x = \begin{cases} 0 & w.p. \quad (\sqrt{2} - 1) \cdot (1 - \frac{1}{n}) \\ \frac{1}{n} & w.p. \quad 2 - \sqrt{2} \\ 1 & w.p. \quad \frac{1}{n}(\sqrt{2} - 1) \end{cases}$$

Actually, this is a counterexample appeared in [24]. They show that

$$\text{OPT-}\mathscr{W}(\mathscr{I}) = \frac{4(\sqrt{2} - 1)}{n} - \frac{4\sqrt{2} - 1}{n^2}$$

Since we will let $n \to \infty$, we will ignore the $O(\frac{1}{n^2})$ terms in the following calculation. Notice that $\Pr[x = 1] = O(\frac{1}{n})$, the probability that there is at least 1 sample with value 1 is negligible, so we will also assume that all samples are 0 or $\frac{1}{n}$. Recall that there will be a tie-breaker coordinate drawn uniformly from $[0, 1]$ for each variable, and we will compare the tie-breaker coordinate if they have the same value. Now suppose there are k_1 samples with value 0 and k_2 samples with value $\frac{1}{n}$. We know that the largest 0 or the smallest $\frac{1}{n}$ is an optimal price for the empirical distribution when $k_1, k_2 \neq 0$. Now, if we choose the largest 0, as the price p, the expected welfare is:

$$\begin{aligned} W_1(k_1) =& \Pr[S=0]\cdot\Pr[S<p]\cdot E[B]+\Pr\left[S=\frac{1}{n}\right]\cdot\frac{1}{n}+\Pr[S=1]\cdot 1 \\ =&(\sqrt{2}-1)\cdot\frac{k_1}{k_1+1}\left((2-\sqrt{2}+\sqrt{2}-1)\cdot\frac{1}{n}\right)+(2-\sqrt{2})\cdot\frac{1}{n}+\frac{1}{n}\cdot(\sqrt{2}-1) \end{aligned} \tag{38}$$

if we choose the smallest $\frac{1}{n}$, as the price p, the expected welfare is:

$$\begin{aligned} W_2(k_2) =& \Pr[S=0]\cdot E[B \,\&\, B>p]+\Pr\left[S=\frac{1}{n}\right]\cdot\left(\frac{1}{n}+\Pr[S<p]\cdot\mathbb{E}[B \,\&\, B=1]\right)+\Pr[S=1]\cdot 1 \\ =&(\sqrt{2}-1)\cdot\left((2-\sqrt{2})\cdot\frac{1}{n}\cdot\frac{k_2}{k_2+1}+\frac{1}{n}(\sqrt{2}-1))\right)+(2-\sqrt{2})\cdot\left(\frac{1}{n}+\frac{1}{k_2+1}\cdot(\sqrt{2}-1)\frac{1}{n}\right)+\frac{1}{n}\cdot(\sqrt{2}-1) \end{aligned} \tag{39}$$

When all the samples have the same value, any price is optimal. Similar to the case when $N=1$, the trade may never happen, so the expected welfare is $\frac{1}{n}$ in this case.

Now, when there are 5 samples, suppose the ERM will choose the largest 0 when there are $1\sim 3$ samples with value 0, and choose the smallest $\frac{1}{n}$ when there are 1 samples with value $\frac{1}{n}$. The expected welfare of ERM when $N=5$ is :

$$\sum_{i=1}^{3}(\sqrt{2}-1)^i(2-\sqrt{2})^{5-i}\binom{5}{i}\cdot W_1(i)+\sum_{i=1}^{1}(\sqrt{2}-1)^{5-i}(2-\sqrt{2})^i\binom{5}{i}\cdot W_2(i)+(p^5+(1-p)^5)\cdot\frac{1}{n}$$

Now compare it to the optimal:

$$\text{OPT-}\mathcal{W}(\mathcal{I})=\frac{4(\sqrt{2}-1)}{n}-\frac{4\sqrt{2}-1}{n^2}$$

By numerical calculations, the ratio is ≈ 0.76 as $n\to\infty$.

Similarly, when there are 10 samples, suppose the ERM will choose the largest 0 when there are $1\sim 6$ samples with value 0, and choose the smallest $\frac{1}{n}$ when there are $1\sim 3$ samples with value $\frac{1}{n}$. The expected welfare of ERM when $N=10$ is :

$$\sum_{i=1}^{6}(\sqrt{2}-1)^i(2-\sqrt{2})^{10-i}\binom{10}{i}\cdot W_1(i)+\sum_{i=1}^{3}(\sqrt{2}-1)^{10-i}(2-\sqrt{2})^i\binom{10}{i}\cdot W_2(i)+(p^{10}+(1-p)^{10})\cdot\frac{1}{n}$$

By calculations, the approximation ratio is ≈ 0.80 as $n\to\infty$.

D.6 Proof of Lemma 5.3

Now we fix the quantile mechanism Q and suppose its c.d.f. is $Q(x)$. Let r be

$$\min_{x\in[0,1]}\int_{[0,x]} t\,dQ(t)+1-x.$$

[5] already prove that r is the lower bound the approximation ratio. We are only left to show that it is also the upper bound. We aim to show that for any $\varepsilon>0$, there exists an instance $\mathcal{I}=(F_S,F_B)$ such that

$$\mathcal{W}(\mathcal{I},Q)\le(r+\varepsilon)\cdot\text{OPT-}\mathcal{W}(\mathcal{I})$$

Now suppose the optimization problem above achieves its minimum at x^*. Consider the following instance $\mathcal{I}=(F_S,F_B)$.

$$S \sim F_S, S = \begin{cases} U[0,\varepsilon] & w.p. \quad x^* \\ H+\varepsilon & w.p. \quad 1-x^* \end{cases} \qquad B \sim F_B, B = H \qquad w.p. \quad 1$$

where $H > 1$ is a sufficiently large number. We can see that in this instance,

$$\text{OPT-}\mathscr{W}(\mathscr{I}) \geq H$$

Now we compute the expected welfare for our mechanism. When its price has a quantile t smaller than or equal to x^*, the trade will happen with probability exactly t. When the quantile of its price is greater than x^*, the trade will never happen.

$$\begin{aligned}
\mathscr{W}(\mathscr{I}) &= \mathbb{E}[S] + \mathbb{E}[(B-S)\mathbb{1}[B \leq p \leq S]] \\
&\leq x^*\varepsilon + (H+\varepsilon)(1-x^*) + \int_{[0,x^*]} tH\,dQ(t) \\
&\leq H \cdot \left(1 - x^* + \int_{[0,x^*]} t\,dQ(t)\right) + \varepsilon \\
&\leq H \cdot r + \varepsilon \leq H \cdot (r+\epsilon) \\
&\leq (r+\epsilon) \cdot \text{OPT-}\mathscr{W}(\mathscr{I})
\end{aligned}$$

where the third inequality follows from $r = \int_{[0,x^*]} t\,dQ(t) + (1-x^*)$.

Therefore, we could find an instance $\mathscr{I} = (F_S, F_B)$ such that $\mathscr{W}(Q,\mathscr{I}) \leq (r+\varepsilon) \cdot \text{OPT-}\mathscr{W}(\mathscr{I})$ for any small enough $\varepsilon > 0$, and this concludes our proof.

D.7 Proof of Theorem 5.2

The proof of Theorem 5.2 is nearly identical to Theorem 5.1. From Lemma 5.3, we know that the approximation ratio of a quantile mechanism Q with a continuous p.d.f. $q(x)$ is exactly

$$\mathscr{C}(p) = \min_{x \in [0,1]} \int_0^x q(t)\,t\,dt + 1 - x.$$

where we use $q(x)\,dx$ instead of $dQ(x)$ since we have a continuous probability density function.

Again, we know that the set of all distributions over $[N]$ is actually $\left\{\{P_i\}_{i\in[n]} \mid P_i \geq 0 \text{ and } \sum_{i=1}^n P_i = 1\right\}$ where distribution P would choose i w.p. P_i. What's more, the probabiltiy density function of the quantile of such order statistic mechanism is exactly $p(x) = \sum_{i=1}^n P_i f_N^i(x)$ where $f_N^i(x) = N\binom{n-1}{i-1} \cdot x^{i-1} \cdot (1-x)^{N-i}$. Therefore, combining it with Lemma 5.1, it follows that Theorem 5.2 holds.

D.8 Proof of Lemma 5.4

Before we give the proof, we first introduce some notations and lemmas about Bernstein that may useful to our proof.

Definition D.1 (Stochastic Bernstein Polynomials)**.** *The stochastic Bernstein polynomial of degree n for a continous function f on $[0,1]$ is defined as*

$$\left(B_n^X f\right)(t) = \sum_{k=0}^{n} f(X_k)\,p_{n,k}(t)$$

in which $X_0, X_1, \cdots X_n$ *are the order statistics of* $(n+1)$ *independent copies of the random variable uniformly distributed in* $[0,1]$*, and,*

$$p_{n,k}(t) = \binom{n}{k} x^k (1-x)^{n-k}, 0 \le k \le n, 0 \le t \le 1$$

Now fix the continous function $q(x)$ we aim to approximate, define $\omega(h)$ as the following function:

$$\omega(h) = \sup_{\substack{0 \le x,y \le 1 \\ |x-y| \le h}} |q(x) - q(y)|$$

Now we can introduce the lemma in [30] that help us approximate the function $q(x)$ by order statistics.

Lemma D.2 (Theorem 2.11 In [30])**.** *Let* $\varepsilon > 0$ *and* $f \in C[0,1]$ *be given. Suppose that* $\omega\left(\frac{1}{\sqrt{n}}\right) < \varepsilon/6.2$*. Then the following inequality holds true:*

$$\Pr\left[\left\|B_n^X f - f\right\|_\infty > \varepsilon\right] \le 2(n+1)\exp\left(-\frac{2\varepsilon^2}{\omega^2\left(\frac{1}{\sqrt{n}}\right)}\right)$$

We are now ready to prove Lemma 5.4.

We first present our mechanism P. For some instance $\mathscr{I} = (F_S, F_B)$, suppose there are n samples $X_1 \le X_2 \le \cdots \le X_n$ drawn from the distribution. We draw another n samples $Y_1 \le Y_2 \le \cdots \le Y_n$ uniformly and independently from $[0,1]$. Let $s = \sum_{i=1}^n q(Y_i)$ be the sum. Then our mechanism will choose X_i with probability $q(Y_i)/s$

Now, let

$$g(x) = \sum_{i=1}^n q(Y_i) f_n^i(x)/s \quad \text{where } f_n^i(x) = n\binom{n-1}{i-1} \cdot x^{i-1}(1-x)^{n-i}$$

be the corresponding probability density function of the order statistic mechanism P. It suffices to prove that with high probability

$$|g(x) - q(x)| \le \varepsilon \quad \forall x \in [0,1]$$

To prove this, we introduce an intermediate function $h(x)$:

$$h(x) = \sum_{i=1}^n q(Y_i) \cdot \binom{n-1}{i-1} x^{i-1}(1-x)^{N-i} = \sum_{i=1}^n q(Y_i) p_{n-1,i-1}(x)$$

As we can see, h is the stochastic Bernstein polynomial of q with degree $n-1$ and $\omega\left(\frac{1}{\sqrt{n-1}}\right) \le \varepsilon/100$. Applying lemma D.2, we know that

$$\Pr\left[\left\|h - q\right\|_\infty > \varepsilon/4\right] \le 2n\exp\left(-\frac{\varepsilon^2}{8 \cdot \omega^2\left(\frac{1}{\sqrt{n-1}}\right)}\right) \le \varepsilon \tag{40}$$

where the last inequality comes from the assumption in the statement of lemma.

Thus, we only need to show that the difference between h and g is small. First we have

$$\begin{aligned} g(x) &= \sum_{i=1}^{N} \frac{q(Y_i)}{S} f_n^i(x) \\ &= \sum_{i=1}^{N} q(Y_i) \binom{n-1}{i-1} x^{i-1}(1-x)^{n-i} \frac{n}{s} \\ &= \frac{n}{s} \cdot h(x) \end{aligned}$$

So it is equivalent to prove that n and s are close w.h.p. We have the following lemma.

Lemma D.3.

$$\Pr\left[\left(1-\frac{\varepsilon}{4M}\right)\cdot n \le s \le \left(1+\frac{\varepsilon}{4M}\right)\cdot n\right] \ge 1-\varepsilon$$

We first use the lemma to continue our proof, before proving the lemma itself. We know that $|g(x) - h(x)| \le \varepsilon/2\ \forall x \in [0,1]$ if $n\left(1-\frac{\varepsilon}{4M}\right) \le s \le n\left(1+\frac{\varepsilon}{4M}\right)$ and $h(x) < 2M\ \forall x \in [0,1]$. Therefore,

$$\begin{aligned} \Pr[|g(x)-h(x)| \ge \varepsilon/2\ \exists x \in [0,1]] &\le \Pr\left[s > n\left(1+\frac{\varepsilon}{4M}\right)\right] + \Pr\left[s < n\left(1-\frac{\varepsilon}{4M}\right)\right] + \Pr[\exists_{x\in[0,1]} h(x) > 2M] \\ &\le 2\varepsilon \end{aligned} \tag{41}$$

where the second inequality is from Lemma D.3 and the fact that with probability at least $1-\varepsilon$, $\|h - q\| \le \varepsilon/4$ and q has a maximum of M on $[0,1]$.

Now combining inequality (40) and (41), we know that with probability at least $1-3\varepsilon$, we have:

$$|g(x)-q(x)| \le |g(x)-h(x)| + |h(x)-q(x)| \le \varepsilon$$

Since such probability is strictly greater than 0, we know that there exists some order statistic mechanism P over N samples such that $|g(x)-q(x)| \le \varepsilon\ \forall x \in [0,1]$. Also we show that our construction will find such order statistics with high probability. Finally, as we assumed in the statement, it holds that

$$\mathscr{C}(P) \ge \mathscr{C}(Q) - c\cdot|g-q|_\infty \ge \mathscr{C}(Q) - c\varepsilon$$

This completes the proof of Lemma 5.4.

Proof of Lemma D.3. We know that $\mathbb{E}[s] = n$. Notice that s is the sum of n i.i.d. random variables ranging in $[0, M]$. Therefore, by Chernoff bound, it holds that

$$\begin{aligned} \Pr\left[\left(1-\frac{\varepsilon}{4M}\right)\cdot n \le s \le \left(1+\frac{\varepsilon}{4M}\right)\cdot n\right] &\ge 1 - \exp\left(-\frac{\varepsilon^2 n}{48M^2}\right) - \exp\left(-\frac{\varepsilon^2 n}{32M^2}\right) \\ &\ge 1-\varepsilon \end{aligned}$$

where the last inequality is from the property in the statement. □

D.9 Application of Lemma 5.4

D.9.1 Symmetric Instance

We first study the case when the distributions are symmetric. [24] provide a mechanism that chooses the mean of the distribution F as the price. They show that in the symmetric setting, this is the optimal fixed price mechanism and achieves an approximation ratio of $\frac{2+\sqrt{2}}{4}$. However, what we want here is a quantile mechanism and we could not convert such mean-based mechanism directly into an order statistic mechanism. We show that quantile mechanisms can also reach the optimal $\frac{2+\sqrt{2}}{4}$-approximation ratio. After that, we use the technique in Lemma 5.4 to produce an order statistic mechanism that achieves an approximation ratio of $\frac{2+\sqrt{2}}{4} - \varepsilon$ with $\text{poly}\left(\frac{1}{\varepsilon}\right)$ samples.

To start with, we first show our optimal order statistic mechanism.

Theorem D.2. *There is a $\frac{2+\sqrt{2}}{4}$-approximation quantile mechanism in the symmetric setting.*

Proof. Our quantile mechanism Q runs as following:

- Let F be the c.d.f. of the distribution.
- Output $F^{-1}\left(\frac{\sqrt{2}}{2}\right)$, i.e. $\frac{\sqrt{2}}{2}$-quantile of the distribution, as the price.

The approximation ratio could be directly calculated using Lemma 5.2. One could see that

$$\begin{aligned}
&\inf_{x\in[0,1)} \frac{\int_{[0,x]} t(1-x)\,dQ(t) + \int_{(x,1]}(1-t)x\,dQ(t) + (1-x)}{1-x^2} \\
&= \min\left(\min_{x\in[0,\frac{\sqrt{2}}{2}]} \frac{x\cdot(1-\frac{\sqrt{2}}{2})+1-x}{1-x^2}, \inf_{x\in[\frac{\sqrt{2}}{2},1)} \frac{\frac{\sqrt{2}}{2}\cdot(1-x)+(1-x)}{1-x^2}\right) \\
&= \frac{2+\sqrt{2}}{4}.
\end{aligned}$$

which completes the proof. ◻

Now we aim to convert it to an order statistic mechanism.

Theorem D.3. *There exists an order statistic mechanism P with $N = O\left(\frac{1}{\varepsilon^7}\right)$ samples that achieves $\frac{2+\sqrt{2}}{4} - \varepsilon$ approximation.*

To start with, we may notice that it is impossible to directly apply Lemma 5.4 to the optimal quantile mechanism a since it is does not have a continiuous probability density function. So our first step is to provide a quantile mechanism with continuous distribution.

Lemma D.4. *For any $\varepsilon > 0$, let Q' be the quantile mechanism with p.d.f. $\tilde{q}(x)$ satisfying*

$$\tilde{q}(x) = \begin{cases} 0 & x \in \left[0, \frac{\sqrt{2}}{2} - \frac{1}{32}\varepsilon\right] \bigcup \left[\frac{\sqrt{2}}{2} + \frac{1}{32}\varepsilon, 1\right] \\ (32/\varepsilon)^2 \cdot \left(x - \frac{\sqrt{2}}{2} + \frac{1}{32}\varepsilon\right) & x \in \left[\frac{\sqrt{2}}{2} - \frac{1}{32}\varepsilon, \frac{\sqrt{2}}{2}\right] \\ -(32/\varepsilon)^2 \cdot \left(x - \frac{\sqrt{2}}{2} - \frac{1}{32}\varepsilon\right) & x \in \left[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} + \frac{1}{32}\varepsilon\right] \end{cases}$$

The quantile mechanism Q' has an apprximation ratio of $\frac{\sqrt{2}+2}{4} - \varepsilon/2$.

Our last step is to make sure that the approximation ratio would not differ to much for two probability density functions that are close to each other.

Lemma D.5. *Suppose $\mathscr{C}(p)$ is the function that maps every quantile mechanism Q for symmetric instances with a continuous probabilitiy density function $q(x)$ to its approximation ratio where*

$$\mathscr{C}(p) = \inf_{x\in[0,1)} \frac{\int_{[0,x]} q(t)t(1-x)\,dt + \int_{(x,1]} q(t)(1-t)x\,dt + (1-x)}{1-x^2}.$$

For any quantile mechanism Q_1 with continuous p.d.f $q_1(x)$ and Q_2 with continuous p.d.f. $q_2(x)$, it holds that

$$\mathscr{C}(p_1) - \mathscr{C}(p_2) \geq -\left|q_1 - q_2\right|_\infty.$$

We first use these lemmas to give a proof of Theorem D.3, and leave the proof of Lemma D.4 and Lemma D.5 to Appendix D.10 and D.11.

Proof Of Theorem D.3. Let $n = c\cdot\frac{1}{\varepsilon^7}+1$ where c is a large enough constant. We could see that $\omega\left(\frac{1}{\sqrt{n-1}}\right) \leq 32^2 c^{-0.5}\cdot\varepsilon^{1.5}$. This means that $2n\exp\left(-\frac{\varepsilon^2}{8\omega^2\left(\frac{1}{\sqrt{n-1}}\right)}\right) \leq \varepsilon/2$. Besides, define $M = \max_{x\in[0,1]}\tilde{q}(x) = 32/\varepsilon$, it is also easy to verify that $\exp\left(-\frac{\varepsilon^2 n}{48M^3}\right) \leq \varepsilon/4$. Combining the properties above with Lemma D.5, we could apply Lemma 5.4, and see that there exists an order statistic mechanism with n samples with an approximation of at least $\mathscr{C}(Q') - \varepsilon/2$. Together with Lemma D.4, it follows that this order statistic mechanism is $\frac{2+\sqrt{2}}{4} - \varepsilon$-approximate. This concludes our proof.

We would like to comment that if we always choose the $\left\lfloor\frac{\sqrt{2}}{2}\cdot n\right\rfloor$-th order statistic as the price, there is an argument to prove that we could achieve an approximation ratio of $\frac{2+\sqrt{2}}{4} - \varepsilon$ with $\tilde{O}\left(\frac{1}{\varepsilon^2}\right)$ samples. □

D.9.2 General Instance

We now consider the general instance. [5] provides a $1-1/e$ approximation quantile mechanism that is also shown to be optimal by [24]. Using the block-box reduction shown in Lemma 5.4, we show that the optimal quantile mechanism can be approximated by order statistics mechanism as closely as desired and also obtain an upper bound on the sample complexity.

Theorem D.4. *There exists an order statistic mechanism P with $N = O\left(\frac{1}{\varepsilon^5}\right)$ samples that achieves $1-\frac{1}{e}-\varepsilon$ approximation.*

In the following proof, we will use order statistic mechanism to approximate the optimal quantile mechanism Q with p.d.f. $q(x) = \frac{1}{x}$ on $[1/e, 1]$. Similarly, $q(x)$ is not a continous function on $[0,1]$. Thus, we need to first convert it to a continous function $\tilde{q}(x)$ on $[0,1]$ and then apply Lemma 5.4.

Similarly, we introduce the following two lemmas first.

Lemma D.6. *For any $\varepsilon > 0$, let $\tilde{Q}$ be the quantile mechanism with p.d.f. $\tilde{q}(x)$ satisfying*

$$\tilde{q}(x) = \begin{cases} 0 & x \in \left[0, \frac{1}{e} - \frac{\left(1-e^{-1}\right)}{\left(e-\frac{1}{2}\varepsilon\right)}\varepsilon\right] \\ \frac{(e-\frac{1}{2}\varepsilon)^2}{\varepsilon(1-e^{-1})}\left(x - \left(\frac{1}{e} - \frac{\left(1-e^{-1}\right)}{\left(e-\frac{1}{2}\varepsilon\right)}\varepsilon\right)\right) & x \in \left[\frac{1}{e} - \frac{\left(1-e^{-1}\right)}{\left(e-\frac{1}{2}\varepsilon\right)}\varepsilon, \frac{1}{e}\right] \\ \frac{1}{x} - \frac{1}{2}\varepsilon & x \in \left[\frac{1}{e}, 1\right] \end{cases}$$

The quantile mechanism Q' has an apprximation ratio of $1-\frac{1}{e}-\varepsilon/2$.

Lemma D.7. *Suppose $\mathscr{C}(p)$ is the function that maps every quantile mechanism Q for general instances with a continuous probabilitiy density function $q(x)$ to its approximation ratio where*

$$\mathscr{C}(p) = \min_{x\in[0,1]} \int_0^x q(t)t\,dt + 1 - x.$$

For any quantile mechanism Q_1 with continuous p.d.f $q_1(x)$ and Q_2 with continuous p.d.f. $q_2(x)$, it holds that

$$\mathscr{C}(p_1) - \mathscr{C}(p_2) \geq -\left|q_1 - q_2\right|_\infty.$$

The proofs of Lemma D.6 and Lemma D.7 are respectively in Appendix D.12 and D.13.

Proof Of Theorem D.4. We follow the same argument to prove Theorem D.4.

Let $n = c\cdot\frac{1}{\varepsilon^5}+1$ where c is a large enough constant. Again it is easy to see that $\omega\left(\frac{1}{\sqrt{n}}\right) \leq c^{-0.5}e^2(1-e^{-1})^{-1}\cdot\varepsilon^{1.5}$. Thus we have $2n\exp\left(-\frac{\varepsilon^2}{8\omega^2\left(\frac{1}{\sqrt{n-1}}\right)}\right) \leq \varepsilon/2$.

Besides, define $M = \max_{x\in[0,1]}\tilde{q}(x) \leq 2e$, we could also verify that $\exp\left(-\frac{\varepsilon^2 n}{48M^3}\right) \leq \varepsilon/4$. Combining the properties above with Lemma D.7, we could see the existence of an order statistic mechanism with n samples with an approximation of at least $\mathscr{C}(Q')-\varepsilon/2$ by applying Lemma 5.4. Together with Lemma D.6, we know that the approximation ratio of this order statistic mechanism is $1-1/e-\varepsilon$. This finishes our proof. ◻

D.10 Proof of Lemma D.4

One could see that the quantile mechanism $\tilde{Q}$ would choose a price with its quantile in $\left[\frac{\sqrt{2}}{2}-\frac{1}{32}\varepsilon, \frac{\sqrt{2}}{2}+\frac{1}{32}\varepsilon\right]$.

Fix a price p and an instance $\mathscr{I} = (F,F)$, ALG is the random variable indicating the welfare for the fixed price p in the realzation, i.e. $S+(B-S)\mathbb{1}[B\geq p\geq S]$ where B, S are drawn independently from F. Thus, it suffices to show that for any instance $\mathscr{I} = (F,F)$, the following holds when $\frac{\sqrt{2}}{2}-\frac{1}{32}\varepsilon \leq t \leq \frac{\sqrt{2}}{2}+\frac{1}{32}\varepsilon$.

$$\mathbb{E}\left[\mathsf{ALG} \mid p = F^{-1}(t)\right] \geq \left(\frac{2+\sqrt{2}}{4} - \varepsilon/2\right)\cdot \text{OPT-}\mathscr{W}(\mathscr{I})$$

To prove the approximation ratio, we need to use Lemma 5.2. Notice that the term $\mathbb{E}[\text{ALG} \mid p = F^{-1}(t)]$ could be understood as the expected welfare of a quantile mechanism that always selects the t-quantile as the price, so Lemma 5.2 could also be applied to analyze the ratio between $\mathbb{E}\left[\mathsf{ALG} \mid p = F^{-1}(t)\right]$ and OPT-$\mathscr{W}(\mathscr{I})$. For $x \geq t$, we know that

$$\inf_{x\in[t,1)} \frac{(1-x)+t(1-x)}{1-x^2} = \frac{1+t}{2}$$

When $x < t$, it holds that

$$\min_{x\in[0,t]} \frac{(1-x)+x(1-t)}{1-x^2} = \frac{1+\sqrt{1-t^2}}{2}$$

As we can see, $\frac{1+t}{2} \geq \frac{1+\sqrt{1-t^2}}{2}$ when $1 \geq t \geq \frac{\sqrt{2}}{2}$ and $\frac{1+t}{2} \leq \frac{1+\sqrt{1-t^2}}{2}$ when $\frac{\sqrt{2}}{2} \geq t \geq 0$. By Lemma 5.2, we know that for any instance $\mathscr{I} = (F,F)$,

$$\mathbb{E}\left[\mathsf{ALG} \mid p = F^{-1}(t)\right] \geq \begin{cases} \dfrac{1+\sqrt{1-t^2}}{2} \cdot \text{OPT-}\mathscr{W}(\mathscr{I}) & t \in \left[\dfrac{\sqrt{2}}{2}, 1\right] \\ \dfrac{1+t}{2} \cdot \text{OPT-}\mathscr{W}(\mathscr{I}) & t \in \left[0, \dfrac{\sqrt{2}}{2}\right] \end{cases}$$

Now if $\frac{\sqrt{2}}{2} - \frac{1}{32}\varepsilon \leq t \leq \frac{\sqrt{2}}{2}$, it holds that

$$\begin{aligned} \mathbb{E}\left[\mathsf{ALG} \mid p = F^{-1}(t)\right] &\geq \frac{1+t}{2} \cdot \text{OPT-}\mathscr{W}(\mathscr{I}) \\ &\geq \left(\frac{2+\sqrt{2}}{4} - \varepsilon/2\right) \cdot \text{OPT-}\mathscr{W}(\mathscr{I}) \end{aligned}$$

.

For $\frac{\sqrt{2}}{2} \leq t \leq \frac{\sqrt{2}}{2} + \frac{1}{32}\varepsilon$, we have

$$\begin{aligned} \mathbb{E}\left[\mathsf{ALG} \mid p = F^{-1}(t)\right] &\geq \frac{1+\sqrt{1-t^2}}{2} \cdot \text{OPT-}\mathscr{W}(\mathscr{I}) \\ &\geq \left(\frac{2+\sqrt{2}}{4} - \varepsilon/2\right) \cdot \text{OPT-}\mathscr{W}(\mathscr{I}) WW \end{aligned}$$

which concludes the proof.

D.11 Proof of Lemma D.5

Suppose $|q_1(x) - q_2(x)| \leq c$ for all $x \in [0,1]$. We could see that

$$\begin{aligned} &\frac{\int_0^x (q_1(t) - q_2(t)) t(1-x)\, dt + \int_x^1 (q_1(t) - q_2(t))(1-t) x\, dt}{1-x^2} \\ &\geq -\frac{\int_0^x ct(1-x)\, dt + \int_x^1 c(1-t)x\, dt}{1-x^2} \\ &\geq -\frac{\int_0^x ct(1-x)\, dt + \int_x^1 c(1-x)x\, dt}{1-x^2} \\ &= -\frac{\int_0^x ct\, dt + \int_x^1 cx\, dt}{1+x} \geq -c. \end{aligned}$$

holds for all $x \in [0,1]$.

Taking the infimum over $[0,1)$, this directly implies that $\mathscr{C}(Q_1) - \mathscr{C}(Q_2) \geq -c$.

D.12 Proof of Lemma D.6

By Lemma 5.3, the approximation ratio of $\tilde{Q}$ can be computed as

$$\begin{aligned} \min_{x\in[0,1]} \int_0^x \tilde{q}(t) \cdot t\, dt + 1 - x &\geq \min_{x\in[0,1]} \int_{\frac{1}{e}}^x \left(\frac{1}{t} - \frac{1}{2}\varepsilon\right) \cdot t\, dt + 1 - x \\ &\geq \min_{x\in[0,1]} \int_{\frac{1}{e}}^x \frac{1}{t} \cdot t\, dt + 1 - x - \frac{1}{2}\varepsilon \\ &= 1 - \frac{1}{e} - \frac{1}{2}\varepsilon. \end{aligned}$$

D.13 Proof of Lemma D.7

Suppose $|q_1(x) - q_2(x)| \le c$ for all $x \in [0,1]$. We could see that

$$\left(\int_0^x q_1(t)\,t\,dt + (1-x)\right) - \left(\int_0^x q_2(t)\,t\,dt + (1-x)\right)$$
$$\ge -\int_0^x c\,t\,dt \ge -c.$$

holds for all $x \in [0,1]$.

Taking the minimum over $[0,1)$, this directly implies that $\mathscr{C}(Q_1) - \mathscr{C}(Q_2) \ge -c$.

E Details of Numerical Experiments

E.1 Symmetric Instance

We now present the details of numerical experiments for the symmetric instances $\mathscr{I} = (F,F)$. We first formally write down the optimization problem that indicates the optimal order statistic mechanisms. As we proved in Section 5.2.1, suppose the following optimization problem PO achieves its maximum $\mathsf{OPT}_{\mathsf{PO}}$ at $(P_1^*, P_2^*, \ldots, P_N^*)$. Let P^* be the distribution over $[N]$ such that $\Pr_{x\sim P^*}[x = i] = P_i^*$. Then, P^* is the optimal order statistic mechanism in the symmetric setting with N samples and its approximation ratio is $\mathsf{OPT}_{\mathsf{PO}}$. Notice that since the c.d.f. of the order statistic mechanism $Q(x)$ is differentiable, we use $q(t)\,dt$ instead of $dQ(t)$ for ease of computation. Here $q(x)$ is the probality density function of mechanism Q.

The Optimization Problem PO

$$\max_{P_1,\ldots,P_N} \quad \min_{x\in[0,1]} \quad \frac{\int_0^x q(t)\cdot t(1-x)\,dt + \int_x^1 q(t)\cdot(1-t)x\,dt + (1-x)}{1-x^2}$$

$$\text{s.t.} \quad q(x) = \sum_{i=1}^{N} P_i \cdot f_N^i(x)$$

$$f_N^i(x) = N\cdot\binom{N-1}{i-1}\cdot x^{i-1}(1-x)^{n-i} \quad \forall i\in[N]$$

$$\sum_{i=1}^{n} P_i = 1 \tag{42}$$

$$P_i \ge 0 \quad \forall i \in [N] \tag{43}$$

Now we aim to solve the optimization problem PO numerally with different numbers of samples N. For the inner minimization problem, it must be solved accurately so that it precisely reflect the approximation ratio of the order statistics mechanism. We use binary search to find its optimum. When we need to check whether

$$\inf_{x\in[0,1)} \frac{\int_0^x q(t)\cdot t(1-x)\,dt + \int_x^1 q(t)\cdot(1-t)x\,dt + (1-x)}{1-x^2} \ge r$$

It is equivalent to check if

$$\int_0^x q(t)\cdot t(1-x)\,dt + \int_x^1 q(t)\cdot(1-t)x\,dt + (1-x) - r(1-x^2) \ge 0 \quad \forall x\in[0,1]$$

Notice that $q(t)$ is a polynomial of degree N, thus we only need to find the minimum of a single-variable polynomial over $[0,1]$, and this could be efficiently done by finding the roots of its derivatives.

We do the binary search for 100 times, so the error caused by binary search is at most 2^{-100}, which is much smaller than the floating point errors and can be ignored. Then we use some empirical algorithms to search for parameters in the outer maximization problem. Unlike the inner minimization problem, we do not need to get an exact optimum of the outer maximization problem since it only reflects the ratio of the order statistic mechanism we have found. The code can be found at our GitHub repository(`https://github.com/BilateralTradeAnonymous/On-the-Optimal-Fixed-Price-Mechanism-in-Bilateral-Trade`).

E.2 General Instance

Now again we formally write down the optimization problem that characterizes the optimal order statistic mechanism with N samples. Similarly, as we proved in Section 5.2.2, suppose the following optimization problem QO achieves its maximum $\mathsf{OPT}_{\mathsf{QO}}$ at $(P_1^*, P_2^*, \ldots, P_n^*)$. Let P^* be the distribution over $[N]$ such that $\Pr_{x\sim P^*}[x = i] = P_i^*$. Then, P^* is the optimal order statistic mechanism in the general setting with N samples and its approximation ratio is exactly $\mathsf{OPT}_{\mathsf{QO}}$. Again notice that since the c.d.f. of the order statistic mechanism $Q(x)$ is differentiable, we use $q(t)\,dt$ instead of $dQ(t)$ where $q(x)$ is the p.d.f. of the order statistic mechanism.

The Optimization Problem QO

$$\max_{P_1,\ldots,P_N} \quad \min_{x\in[0,1]} \int_0^x q(t)\cdot t\,dt + 1 - x$$

$$\text{s.t.} \quad q(x) = \sum_{i=1}^{N} P_i \cdot f_N^i(x)$$

$$f_N^i(x) = N \cdot \binom{N-1}{i-1} \cdot x^{i-1}(1-x)^{n-i} \quad \forall i \in [N]$$

$$\sum_{i=1}^{n} P_i = 1 \tag{44}$$

$$P_i \geq 0 \quad \forall i \in [N] \tag{45}$$

Now we solve the optimization problem numerically for different fixed number N. Again the inner minimization problem could be solved efficiently by calculating the zero point of its derivatives and we search for the parameters in the outer maximization problem to get a good enough solution. The code could be found at our Github repository.