The Power of Menus in Dynamic Pricing: Optimal Regret and Equivalence Results

Shiliang Zuo

Meta
szuo.rs@gmail.com

Abstract

We study dynamic pricing where a seller repeatedly interacts with a strategic, non-myopic buyer who has a fixed private valuation and discounts future utility. Prior work focused exclusively on posted-price mechanisms, where the seller gives a take-it-or-leave-it offer. For our first result, under geometric discounting $\gamma_t = \gamma^{t-1}$, we show that menu mechanisms consisting of allocation-payment contracts achieve $O(T_\gamma)$ regret, where T_γ is the buyer's effective discounted time horizon. We also establish a $\Omega(T_\gamma)$ lower bound, demonstrating the bound is tight. Considering the geometric discounting buyer with a constant discount factor, our bound is $O(1)$, while prior bounds using posted-price mechanisms incur an unavoidable $\Omega(\log \log T)$ factor in regret. Our second contribution is more conceptual in nature. The problem of dynamic pricing sits at the intersection of two paradigms: learning with strategic agents in computer science / machine learning and revelation-principle-based mechanism design in economics, yet their relationship has remained unclear. We establish an equivalence between these two classes of mechanisms in our single-buyer setting: any adaptive indirect menu mechanism can be converted into an incentive-compatible direct mechanism with no less revenue, and conversely, any such direct mechanism can be implemented by an adaptive indirect menu mechanism with no less revenue. Consequently, the two classes achieve identical optimal regret. The adaptive, data-driven algorithms of online learning and explicit type elicitation from direct revelation type mechanisms are two languages towards solving the same problem.

1 Introduction

In this work we study the problem of *dynamic pricing*: a seller repeatedly offers a product to a single buyer over T rounds. The buyer has a private valuation v for the item that remains fixed across time, while the seller seeks to maximize cumulative revenue. This problem was introduced by Amin et al. [2013] and has since been studied extensively (Kleinberg and Leighton [2003], Mohri and Medina [2014], Drutsa [2017], Haghtalab et al. [2022], Drutsa [2020a]). Despite this rich line of work, two significant gaps remain. First, all prior work focuses exclusively on posted-price mechanisms, yet there exists a much richer class of mechanisms based on menus—and it is unclear whether this richer class can yield improved regret. Second, there is a conceptual gap. Dynamic pricing has been studied from two seemingly difference perspectives. In machine learning, the problem is treated as strategic online learning and typically designs indirect adaptive learning-based mechanisms. In economics, the problem is treated as dynamic contracting and typically works with *direct* revelation mechanisms. What is the precise relationship between these two classes of mechanisms? We elaborate on both gaps and present our contributions below.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Table 1: Regret bounds for dynamic pricing ($\gamma_t = \gamma^{t-1}$, $\gamma < 1$). For lower bounds, Amin et al. [2013] provide $\Omega(T_\gamma)$ in the non-myopic setting, while Kleinberg and Leighton [2003] provide $\Omega(\log\log T)$ in myopic setting.

Mechanism	**Any T, γ**	**Constant γ independent of T**
Posted-price [Drutsa, 2017]	$O(T_\gamma \log T_\gamma \log\log T)$	$O(\log\log T)$
Posted-price [Haghtalab et al., 2022]	$O(\log T + T_\gamma \log T_\gamma)$	$O(\log T)$
Menu, 2 items (Ours)	$\boldsymbol{O(T_\gamma)}$	$\boldsymbol{O(1)}$
Lower bound for posted-price	$\Omega(T_\gamma)$ / $\Omega(\log\log T)$	
Lower bound for menu (Ours)	$\boldsymbol{\Omega(T_\gamma)}$	

Menu-based learning mechanisms. Prior work has focused exclusively on *posted-price mechanisms*, where the seller offers a single take-it-or-leave-it price in each round. Yet in principle, the seller can employ a much richer class of mechanisms: *menus*. A menu mechanism offers the buyer a set of allocation-payment contracts, allowing the buyer to select among multiple options rather than making a binary accept/reject decision. Besides its literal reading as a probability of allocation, a fractional allocation can represent the fraction of the item delivered when the item is divisible, or a service level—for instance, the higher and lower service tiers commonly offered on online marketplaces, where a lower tier provides reduced service at a lower price.

To illustrate the power of menus, consider first the simpler case of a myopic buyer. In the seminal work of Kleinberg and Leighton [2003], it is shown that posted-price mechanisms can achieve regret $\Theta(\log\log T)$, and that such regret is necessary. However, with menus the seller can achieve regret $O(1)$. Specifically, the seller can design a menu $p : [0, 1] \mapsto \mathbb{R}^+$ specifying a payment for each allocation probability. Once $p(\cdot)$ is specified, the buyer then chooses his preferred allocation probability. For example, setting $p(a) = a^2/2$ yields buyer utility $v \cdot a - p(a)$ in a single round, which is maximized at $a = v$. Thus, the buyer's choice immediately reveals their valuation. This raises our first research question:

Can menu mechanisms achieve better regret than posted-price mechanisms against non-myopic buyers?

We answer this question affirmatively. For our regret upper bound, we focus on the geometric discounting setting where $\gamma_t = \gamma^{t-1}$ for $\gamma \in (0, 1)$. We show that menu mechanisms achieve $O(T_\gamma)$ regret, where T_γ denotes the buyer's effective time horizon. Hence, while it is believed since Kleinberg and Leighton [2003] that the dependence on T could not be removed, we show this is not the case if the principal use menu-based mechanisms. In particular, our bound improves upon the best known bounds for posted-price mechanisms: $O(\log T + T_\gamma \log T_\gamma)$ (Haghtalab et al. [2022]) and $O(T_\gamma \log T_\gamma \log\log T)$ (Drutsa [2017]). For constant $\gamma < 1$, our regret bound becomes $O(1)$, compared to the $\Omega(\log\log T)$ lower bound for posted-price mechanisms. Further, our menu only requires two entries each round. The key insight is regret measures a *fundamental tradeoff* between allocative efficiency for lower type agents and information rent paid to high type agents, and that using menus (even with just two entries) enable a fine-grained control of the two quantities.

Indirect Adaptive Mechanisms vs. Direct Revelation Mechanisms. Our second contribution is more conceptual in nature. The dynamic pricing problem has been studied from two seemingly distinct perspectives.

The aforementioned machine learning / computer science perspective, treats the seller's problem as one of *learning through interaction* Kleinberg and Leighton [2003], Amin et al. [2013]. The seller posts prices sequentially, observes the buyer's accept/reject decisions, and adapts future prices based on this feedback. This approach appears algorithmic and data-driven, requiring no upfront communication from the buyer; instead, the buyer simply best-responds to the seller's pricing strategy. In mechanism-design terms, such an algorithm is an adaptive *indirect* mechanism: no explicit reports are solicited, and future outcomes are determined by the history of the buyer's choices.

The economics perspective views this as a problem of *dynamic contracting* Bergemann and Välimäki [2019], Salanié [2005], Börgers [2015]. By the revelation principle, the seller can commit upfront to a *direct revelation mechanism*: the buyer reports a type $\hat{v}$, and the seller commits to a complete sequence of allocations and payments $(a_t(\hat{v}), p_t(\hat{v}))_{t=1}^T$ contingent on this report. The mechanism must satisfy incentive compatibility (IC)—truthful reporting maximizes the buyer's utility—and

individual rationality (IR). The seller's problem then reduces to optimizing revenue subject to these constraints.

These two formulations look quite different on the surface. A direct mechanism relies on an explicit report of the buyer's type, while an adaptive indirect mechanism relies only on preferences revealed through the buyer's choices. The technical tools also differ substantially: mechanism design employs the revelation principle and solves structured optimization problems subject to incentive constraints, whereas online learning utilizes tools from exploration-exploitation tradeoffs and sequential decision-making under partial feedback. Yet both frameworks address the same underlying economic problem. This raises a natural question:

What is the relationship between indirect adaptive mechanisms and direct revelation mechanisms for dynamic pricing against a non-myopic buyer?

We establish that the two classes are *equivalent* in our setting. Specifically, we prove that any adaptive indirect menu mechanism can be converted into an IC and PIR direct revelation mechanism with the same revenue, and conversely, any IC and PIR direct mechanism can be converted into an adaptive indirect menu mechanism with no less revenue via what we call the *price equalization procedure*. As a corollary, the optimal regret achievable by adaptive indirect mechanisms equals that achievable by direct mechanisms.

This equivalence is useful in both directions. The direct formulation characterizes the feasible outcomes through a continuum of IC and PIR constraints, which can be difficult to solve explicitly; an adaptive pricing algorithm instead provides a constructive sequential mechanism without solving this constraint system. Our first conversion shows that any such algorithm implicitly constructs a feasible direct mechanism. Our second conversion shows that, in this setting, explicit type reports provide no additional power over preferences revealed through menu choices: any IC and PIR direct mechanism can be implemented without soliciting reports.

Utilizing the equivalence result, we provide a proof of a $\Omega(T_\gamma)$ lower bound on regret. Hence, our upper bound on regret is tight. This lower bound was first established by Amin et al. [2013] for posted-price mechanisms. We show this bound also holds for menu-based mechanisms; also, our proof clearly demonstrates that at the heart of the problem is a tradeoff between two quantities: allocative inefficiency for lower types versus information rents for higher types.

1.1 Related Work and Discussion

Firstly, the specific problem we studied in this work is built on the line of work on dynamic pricing against non-myopic agents. The problem has received significant attention in the computer science community in recent years, categorized under the broader class of problems of *learning with strategic agents*. The problem was originally introduced by Amin et al. [2013], with followup works including Haghtalab et al. [2022], Drutsa [2017], Mohri and Medina [2014]. Extensions to multiple buyers include Liu et al. [2018], Drutsa [2020b]. The different lens of exploring dynamic pricing against non-myopic agents is explored in economics literature falling under the domain of *dynamic mechanism design* (Bergemann and Välimäki [2019]). For the case where the principal and agent share the same discount factor, there is a well-known result: the optimal dynamic mechanism is a static one that pose the optimal Bayesian price, which is known as *false dynamics* Börgers [2015]. [1]

There also exists work that explore the problem from a pure learning perspective where the agent is myopic and best responds each round. The early work that introduced the problem include Kleinberg and Leighton [2003] (for which the literature on pricing against non-myopic agents was built upon), and later also extended to incorporate contextual information Liu et al. [2021], Mao et al. [2018]. Robust versions of the pricing algorithms were also explored in Krishnamurthy et al. [2021], Zuo [2024].

[1] In fact, as Börgers [2015] point out, if the seller and buyer have different discounting sequence with the buyer discounting future utility more heavily, then the gains from intertemporal trade can be infinite. Indeed, the seller and buyer should sign the following debt contract: the seller pays the buyer 1 dollar in time period 1, and ask the buyer to repay $1/\gamma^{T-1}$ dollars (minus an infinitesimal small amount) in time period T. If the buyer has geometric discounting with rate γ, then the contract is strictly beneficial to both parties. We avoid this scenario by allowing the agent to reserve the right to drop out at any round (see Section 2).

Our work is also closely related the line of work on contract theory, in particular the use of menus in principal-agent models with adverse selection (see Salanié [2005], Bolton and Dewatripont [2004], Laffont and Maskin [1981]). Relatedly, some recent work in the computer science community studied computational aspects of menus in principal-agent problems (e.g. Han et al. [2024], Guruganesh et al. [2023], Castiglioni et al. [2022]).

Another related work is Albert et al. [2025]; the focus of their work was on characterizing when the *false dynamics* scenario can be escaped, i.e., what game structure allows the principal to reason about the agent's private type through repeated interactions (when they share the same discount factor). Another related work is Conitzer and Sandholm [2004]; their work was not on the dynamic setting where "learning" is involved, and focus more on the computational perspective. They discuss the direct revelation mechanism and what they call an *insincere* mechanism, and argue that in the direct revelation mechanism the computational burden is on the principal, while in the insincere mechanism the computational burden is shifted to the agent.

2 Problem Setting: Dynamic Pricing against a Non-Myopic Agent

There are two parties, a seller and a buyer. We may also refer to the seller as the principal and the buyer as the agent. They interact in discrete time for T rounds. The buyer holds a private valuation

$$v \in [0, 1]$$

for the item. At each round, the seller has one unit of the item she would like to sell to the buyer. The seller is patient and does not discount future utility. The buyer is impatient, modeled by a sequence of discount factors $\{\gamma_t\}_{t=1}^{T}$ satisfying

$$\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_T > 0.$$

We write

$$T_\gamma := \sum_{t=1}^{T} \gamma_t$$

for the buyer's effective discounted time horizon. A canonical special case is geometric discounting,

$$\gamma_t = \gamma^{t-1}, \qquad \gamma \in (0, 1),$$

for which

$$T_\gamma = \frac{1 - \gamma^T}{1 - \gamma}.$$

Our regret upper bound (Section 3) is established for geometric discounting, whereas our equivalence results and our lower bound (Section 4) hold for any discount sequence that is positive and non-increasing. We normalize $\gamma_1 = 1$; this is without loss of generality, since rescaling all discount factors by a common positive constant does not change the buyer's preferences.

At each round the outcome is a pair of payment and allocation probabilities (p_t, a_t). The seller's performance is measured by the strategic regret, as defined by the following:

$$\texttt{Reg}(v) = T \cdot v - \sum_{t=1}^{T} p_t, \quad \texttt{Reg} = \sup_{v \in [0,1]} \texttt{Reg}(v)$$

Here $\texttt{Reg}(v)$ measures the difference of the optimal payment value the principal could have collected had she known the agent's true type, versus the actual payment collected.

The buyer's utility is given according to his discounting factor and the actual allocation and payment outcomes:

$$\sum_{t=1}^{T} \gamma_t (a_t \cdot v - p_t).$$

We now describe the classes of mechanisms that determine the actual payment and allocations (p_t, a_t).

2.1 Indirect Menu-based Learning Mechanisms

We begin by describing posted-price mechanisms, the focus of prior work, and then introduce the more general menu-based mechanisms studied in this paper.

Posted-Price Mechanisms. In a posted-price mechanism, the principal posts a price each round and the buyer makes a binary decision to accept or reject. Formally, the mechanism is described by a sequence of mappings $\mathcal{M} = (\mathcal{M}_t)_{t=1}^T$, where

$$\mathcal{M}_t : (p_\tau, a_\tau)_{\tau=1}^{t-1} \mapsto p_t,$$

and $a_\tau \in \{0, 1\}$ denotes the buyer's accept/reject decision at round τ. The mechanism is announced to the buyer before the game begins; given $\mathcal{M}$, the buyer strategically responds to maximize his total utility over the T rounds. Equivalently, such a mechanism can be represented as a *binary decision tree* of depth T. Each non-leaf node is associated with a posted price. A rejection moves the buyer to the left child; an acceptance moves him to the right child. Leaf nodes encode complete interactions of length T. The buyer chooses the root-to-leaf path that maximizes his utility.

Menu-Based Mechanisms. A key departure in our work is allowing fractional allocation probabilities instead of binary allocation decisions. In a menu-based mechanism, at each round the principal posts a *menu of contracts*:

$$\{(a, p(a)) : a \in A \subset [0, 1]\}.$$

Here a is an allocation probability and $p(a)$ is the associated payment. The agent selects an allocation level $a \in A$ and pays $p(a)$, upon which the item is allocated with probability a. The payment is made ex-ante regardless of the actual allocation outcome ex-post. Formally, a menu-based mechanism is described by mappings $\mathcal{M} = (\mathcal{M}_t)_{t=1}^T$, where

$$\mathcal{M}_t : (M_\tau, a_\tau)_{\tau=1}^{t-1} \mapsto M_t,$$

with M_t denoting the menu offered at round t and $a_\tau \in [0, 1]$ the allocation probability selected by the buyer at round τ. Posted-price mechanisms are the special case where each menu contains only two options: $(1, p_t)$ (accept at price p_t) and $(0, 0)$ (reject).

2.2 Direct Revelation Mechanism

In a direct mechanism, the agent is first asked to report his private type (a "bid"). The report $\hat{v} \in [0, 1]$ may differ from the agent's true value v. For each report $\hat{v}$, the principal commits to a sequence of per-round contracts

$$\hat{v} \mapsto (p_t(\hat{v}), a_t(\hat{v}))_{t=1}^T,$$

which specify the payment $p_t(\hat{v})$ and allocation probability $a_t(\hat{v})$ in each period. We assume that payments are nonnegative, i.e., $p_t(\hat{v}) \geq 0$ for all reports $\hat{v}$ and time period t.

Each pair $(p_t(\hat{v}), a_t(\hat{v}))$ functions as a spot contract: if the buyer continues participating in round t, he pays $p_t(\hat{v})$ and receives allocation $a_t(\hat{v})$. The seller and buyer executes this contract sequence *sequentially over time*, but the agent always retains the right to *drop out at any round*. Dropping out terminates the mechanism immediately. The interaction in the direct revelation mechanism proceeds as follows.

1. The principal announces the direct mechanism to the agent.
2. The agent makes a report $\hat{v}$ on his private type.
3. At the beginning of each round t, the agent must decide whether to drop out or not. If he drops out the game immediately ends; otherwise he is charged $p_t(\hat{v})$ and receives the item with probability $a_t(\hat{v})$, then the game proceeds to the next round.

A direct mechanism must satisfy two key conditions:

- *Incentive compatibility (IC):* truthful reporting maximizes the agent's discounted utility among all possible reports.
- *Periodic individual rationality (PIR):* at every round, a truthful agent weakly prefers continuing participation over dropping out.

Together, IC and PIR ensure that an agent is never worse off by reporting his true type and remaining in the mechanism throughout. Similar PIR constraints have also appeared in prior work, e.g. Bergemann and Välimäki [2010], Balseiro et al. [2018].

To formalize these requirements, we define the discounted utility of a type-v agent who reports $\hat{v}$ and accepts all contracts from periods 1 through t as

$$U_{1:t}(\hat{v}; v) := \sum_{\tau=1}^{t} \gamma_\tau [a_\tau(\hat{v}) v - p_\tau(\hat{v})].$$

IC and PIR together then require that for all types v, all reports $\hat{v}$, and all periods $t = 0, 1, \ldots, T$, we have (defining $U_{1:0}(\cdot;\cdot) = 0$):

$$U_{1:T}(v; v) \;\geq\; U_{1:t}(\hat{v}; v). \tag{1}$$

Tie-breaking. In an indirect mechanism, the buyer may have multiple strategies that maximize his total discounted utility, and these may generate different revenue for the seller. Throughout the paper we adopt the convention that the buyer breaks ties in favor of the seller: among all strategies maximizing his total discounted utility over the complete interaction, the buyer selects one that maximizes the seller's total revenue. This is the dynamic analogue of the principal-favorable tie-breaking rule commonly used in Stackelberg equilibria. The direct formulation uses the analogous convention: IC and PIR only require truthful reporting and continued participation to be *weakly* optimal, and the buyer reports truthfully and follows the prescribed contract sequence. Under this convention, the outcome generated by each type v, and hence the seller's revenue, is well defined for both classes of mechanisms.

3 Optimal Regret via Menus

We now demonstrate that by employing menus of contracts, the seller can achieve a regret bound of $O(T_\gamma)$, hence breaking the T-dependent $\Omega(\log \log T)$ barrier present in prior work using posted-price mechanisms. We later establish an $\Omega(T_\gamma)$ lower bound, showing that this rate is tight up to constant factors.

For the upper bound in this section, we specialize to geometric discounting,

$$\gamma_t = \gamma^{t-1}, \qquad \gamma \in (0, 1).^2$$

3.1 Menu-based Learning Mechanism Design

We now present a two-item menu algorithmic learning mechanism that achieves the optimal regret bound. The mechanism is summarized in Algorithm 1. At a high level, the seller maintains an interval $I_t = [\underline{v}_t, \overline{v}_t]$ that contains the buyer's value and repeatedly uses a two-item menu to ask whether the value lies above or below the midpoint of the interval.

Let

$$\delta_t := \overline{v}_t - \underline{v}_t, \qquad m_t := \frac{\underline{v}_t + \overline{v}_t}{2}.$$

The menu offered in the phase is

$$L_t = \big(1 - \delta_t, (1 - \delta_t)\underline{v}_t\big), \qquad H_t = \big(1, \underline{v}_t + \delta_t^2/2\big).$$

[2]The phase-consistency argument can be extended beyond geometric discounting when the discount sequence has sufficiently controlled tails. In particular, for a phase beginning at round t with candidate-interval width δ, it suffices to choose a phase length D such that

$$\sum_{s=t+D}^{T} \gamma_s \leq c\,\delta^2 \sum_{s=t}^{\min\{t+D-1,T\}} \gamma_s$$

for a sufficiently small absolute constant $c > 0$. The resulting regret depends on the phase lengths required to satisfy this tail-decay condition. We do not pursue a general regret bound for arbitrary discount sequences here.

Algorithm 1: Consistent Two-Item Menu Search

Input: Initial candidate interval $[0, 1]$, discount parameter γ, constant $\rho > 0$

Initialize $\underline{v}_1 \leftarrow 0$ and $\overline{v}_1 \leftarrow 1$;
Set $t \leftarrow 1$;
while $t \leq T$ **do**
  $\delta_t \leftarrow \overline{v}_t - \underline{v}_t$;
  $m_t \leftarrow (\underline{v}_t + \overline{v}_t)/2$;
  Define the two menu entries

$$L_t \leftarrow \big(1 - \delta_t, (1 - \delta_t)\underline{v}_t\big), \qquad H_t \leftarrow \big(1, \underline{v}_t + \delta_t^2/2\big).$$

  Set

$$D_t \leftarrow \left\lceil \frac{\log\left(\frac{1+\rho\delta_t^2}{\rho\delta_t^2}\right)}{\log(1/\gamma)} \right\rceil, \qquad \overline{D}_t \leftarrow \min\{D_t, T - t + 1\}.$$

  Post the two-item menu $M_t = \{L_t, H_t\}$ in the first round of the phase;
  Let $C_t \in \{L_t, H_t\}$ be the agent's choice in the first round of this phase;
  for $s = t + 1, \ldots, t + \overline{D}_t - 1$ **do**
    Post the same menu $M_t = \{L_t, H_t\}$;
    if *the agent's choice is not* C_t **then**
      Offer only the outside option $(0, 0)$ in every future round;
  if $C_t = L_t$ **then**
    Update

$$[\underline{v}_{t+D_t}, \overline{v}_{t+D_t}] \leftarrow [\underline{v}_t, m_t + \delta_t/8] \cap [\underline{v}_t, \overline{v}_t].$$

  else
    Update

$$[\underline{v}_{t+D_t}, \overline{v}_{t+D_t}] \leftarrow [m_t - \delta_t/8, \overline{v}_t] \cap [\underline{v}_t, \overline{v}_t].$$

  $t \leftarrow t + D_t$;

For a buyer of value v, the utility difference between the two options is

$$\begin{aligned} U_{H_t}(v) - U_{L_t}(v) &= \left(v - \underline{v}_t - \frac{\delta_t^2}{2}\right) - (1 - \delta_t)(v - \underline{v}_t) \\ &= \delta_t(v - \underline{v}_t) - \frac{\delta_t^2}{2} \\ &= \delta_t(v - m_t). \end{aligned}$$

Thus a myopic buyer with $v > m_t$ prefers H_t, while a myopic buyer with $v < m_t$ prefers L_t.

The mechanism proceeds in phases, and repeats this same two-item menu for an entire phase. The buyer's first choice determines the branch of the binary search. If the buyer ever switches between L_t and H_t within the phase, the seller offers only the outside option $(0, 0)$ in every future round. This consistency requirement amplifies the cost of inducing the wrong branch: a buyer who wants to be misclassified must choose the wrong menu item throughout the phase, rather than only once.

Fix a sufficiently small absolute constant $\rho > 0$. For a phase beginning at round t with interval width δ_t, define the *planned* phase length

$$D_t := \left\lceil \frac{\log\left(\frac{1+\rho\delta_t^2}{\rho\delta_t^2}\right)}{\log(1/\gamma)} \right\rceil$$

and the *actual* phase length

$$\overline{D}_t := \min\{D_t, T - t + 1\}.$$

The seller holds the menu $\{L_t, H_t\}$ fixed for $\overline{D}_t$ rounds. If $\overline{D}_t < D_t$, then the phase reaches the end of the horizon and the mechanism terminates without performing another interval update. Otherwise

the full planned phase is completed, and the interval is updated according to the buyer's consistent choice:

$$I_{t+D_t} = [\underline{v}_t, m_t + \delta_t/8]$$

after choosing L_t, and

$$I_{t+D_t} = [m_t - \delta_t/8, \overline{v}_t]$$

after choosing H_t. Both successor intervals have width at most

$$\frac{5}{8}\delta_t.$$

Although the planned length D_t need not be bounded by the finite horizon, the actual phase length satisfies

$$\overline{D}_t = O\left(T_\gamma\left(1 + \log\frac{1}{\delta_t}\right)\right).$$

3.2 Analysis

We state the key results for the main regret bound, the full proof is deferred to Appendix A. We first show that the candidate interval $[\underline{v}_t, \overline{v}_t]$ always contains the true hidden value v.

Lemma 1 (Consistency of the phase)**.** *There exists an absolute constant $\rho > 0$ such that the following holds in every phase. If $v \geq m_t + \delta_t/8$, then the buyer weakly prefers choosing H_t consistently throughout the phase to choosing L_t consistently and being placed in the left branch. Similarly, if $v \leq m_t - \delta_t/8$, then the buyer weakly prefers choosing L_t consistently throughout the phase to choosing H_t consistently and being placed in the right branch.*

The overlap between the two successor intervals handles types close to the midpoint. If

$$v \in [m_t - \delta_t/8, m_t + \delta_t/8],$$

then either successor interval still contains v, so either branch is acceptable for the purpose of future learning.

Next, the regret of each phase can bounded by the phase's duration and the length of the candidate interval during the phase.

Lemma 2 (Regret of one phase)**.** *The regret incurred during a phase of width δ_t and actual duration $\overline{D}_t$ is $O(\delta_t \overline{D}_t)$.*

Theorem 1 (Optimal regret of two-item menus)**.** *The two-item menu mechanism described above achieves regret $O(T_\gamma)$.*

4 Equivalence Results

Here we show that *indirect menu-based learning* mechanisms are equivalent to IC and PIR direct revelation mechanisms in the single-buyer dynamic pricing model. Given a mechanism $\mathcal{M}$ (either a direct revelation type or an indirect learning type), we define the revenue to the seller as the total collected monetary payment when an agent with private type (value) v best responds to the mechanism: $\mathsf{Rev}(v; \mathcal{M}) = \sum_{t=1}^{T} p_t$. We show that for any indirect mechanism, there exists a direct mechanism whose revenue is at least no less than the indirect one for any type; and vice versa. Results in this section does not require geometric discounting and hold for any nonincreasing and positive discount sequence.

4.1 Converting Indirect Menu Mechanism to Direct Mechanism

The menu-based indirect learning mechanism can be converted to a direct revelation type mechanism by the revelation principle. For each type $v \in [0, 1]$, let

$$H(v) := (a_t(v), p_t(v))_{t=1:T}$$

be the trajectory generated by the principal's menu mechanism and the agent's best response strategy in the indirect mechanism.

Theorem 2. *For any indirect menu-based learning mechanism $\mathcal{I}$, there exists a direct mechanism $\mathcal{D}$ with the same outcome and revenue, i.e,* $\mathsf{Rev}(v, \mathcal{D}) = \mathsf{Rev}(v, \mathcal{I}), \forall v$. *In particular the trajectory $H(v)$ from the indirect mechanism is itself an IC and PIR direct mechanism.*

The proof is by an application of the revelation principle, which we defer to Appendix B. An interesting implication is that any indirect adaptive learning-based mechanism implicitly constructs a feasible solution to the IC and PIR constraint system of the direct formulation.

4.2 Converting Direct Mechanism to Indirect Menu Mechanism

Algorithm 2: Converting direct mechanism to menu-based learning mechanism

Input: direct mechanisms $v \mapsto (p_t(v), a_t(v))_{t=1}^T$;
`// Construct a tree. Each node has two attributes: values and menu`
Construct root node root (depth is 1), set root.values $= [0, 1]$;
`// Price Equalization for Menu Construction`
for $t \leftarrow 1$ **to** T **do**
  for node *in depth* t **do**
    `/* Iterate over all nodes at level` t`, compute the correct menu for the node */`
    $S =$ node.values;
    Let A be the set of allocation probabilities that occur in the set S, i.e., $A = \{a_t(v) : v \in S\}$;
    for $a \in A$ **do**
      Let $S(a) = \{v \in S : a_t(v) = a\}$ be the set of reports whose allocation probability is a;
      Let $p_{\inf}(a) = \inf_{v \in S(a)} p_t(v)$ be the infimum of the prices that occur in the direct mechanism;
      Set $p_{\inf}(0) = 0$;
      `/* Adjustment Steps: equalize prices and defer differences */`
      for $v \in S(a)$ **do**
        if $t < T$ **then**
          Set $p_{t+1}(v) := p_{t+1}(v) + \frac{\gamma_t}{\gamma_{t+1}} \cdot (p_t(v) - p_{\inf}(a))$;
        Set $p_t(v) := p_{\inf}(a)$;
    Set the menu for the current node as node.menu $= \{(a, p_{\inf}(a)) : a \in A\}$;
    for $a \in A$ **do**
      Create child node child_a with $\mathsf{child_a}$.values $= S(a)$;

We now establish the converse direction: any direct mechanism can be implemented as an indirect learning mechanism.

Theorem 3. *Given any direct mechanism $\mathcal{D} : v \mapsto (a_t(v), p_t(v))_{t=1:T}$ that is IC and PIR, Algorithm 2 produces an indirect menu-based mechanism $\mathcal{I}$ satisfying:* $\forall v$, $\mathrm{Rev}(v, \mathcal{I}) \geq \mathrm{Rev}(v, \mathcal{D})$.

The conversion from direct to indirect is more involved than the reverse direction (which follows directly from the revelation principle). The core difficulty is the following: in a direct mechanism, the seller knows the buyer's reported type, so she can charge *different payments to different types even when they receive the same allocation*. In an indirect mechanism, the seller never learns the buyer's type explicitly—she can only post a menu of (allocation, payment) pairs, and every buyer facing the same menu must pay the same price for the same allocation. We introduce a procedure (Algorithm 2) that converts any direct mechanism into an indirect one. Note that in the procedure, for direct mechanisms, payments $p_t(\hat{v})$ are indexed by the reported type, whereas in the indirect mechanism, payments $p(a)$ depend only on the chosen allocation—this asymmetry is precisely what makes the conversion non-trivial.

The procedure essentially shows how to "unfold" any direct mechanism into a *tree-structured* indirect mechanism. Each node in the tree corresponds to a particular interaction history, and branches correspond to the buyer's allocation choices. At each node we track:

1. **Candidate set** node.values: this is the set of types which would have reached this node in the interaction history. At each depth, the sets are disjoint and the union is the full set $[0, 1]$.
2. **Menu** node.menu: the (allocation, payment) pairs offered at this history.

The central challenge is *price equalization*. Consider a node whose candidate set contains types v and v' that receive the same allocation $a_t(v) = a_t(v') = a$ but pay different prices $p_t(v) \neq p_t(v')$. Since the indirect mechanism must post a single price for allocation a, we cannot directly replicate both contracts. Our solution is to set the menu price to the *lowest* price in each allocation group,

$$p_{\inf}(a) = \inf_{v \in S(a)} p_t(v), \quad \text{where } S(a) = \{v \in \mathsf{node.values} : a_t(v) = a\},$$

and *defer* the remaining payment to the next round. Specifically, for any type v that originally paid more than $p_{\inf}(a_t(v))$, we add the difference—adjusted for discounting by the factor γ_t/γ_{t+1}—to that type's payment in round $t + 1$. This deferral is feasible because IC guarantees that each type has sufficient continuation utility to absorb the extra charge. After each adjustment step, the buyer's total discounted utility is unchanged, while the seller's revenue weakly increases (since payments are moved earlier in effective terms). The complete proof appears in Appendix B.

The conversions (Theorem 2, Theorem 3) between indirect learning mechanisms and direct mechanisms indeed implies that the two classes of mechanisms are equivalent in power.

Corollary 1. *Given any direct mechanisms $\mathcal{D}$ with regret $\mathsf{Reg}(\mathcal{D})$, there exists an indirect menu-based learning mechanism with regret $\mathsf{Reg}(\mathcal{I})$ no larger than that of the direct one; and vice versa.*

Note that while the conversion in Theorem 3 can strictly *increase* revenue, it does not mean that indirect mechanisms are more powerful; it is simply because the input direct mechanism was suboptimal. The revenue gain comes from deferring early payments to later rounds, which increases revenue while keeping the buyer indifferent (his discounted utility is unchanged). An optimally designed direct mechanism would already exploit this, leaving no slack for the conversion to improve upon. Thus neither class dominates the other, and the optimal regret is identical.

If we focus on the class of (indirect learning) posted-price mechanisms, one can show (via the same procedure) that they are equivalent to direct revelation mechanisms *with binary allocation outcomes*.

Corollary 2. *The conversions (Theorem 2, Theorem 3) holds with "indirect menu-based learning mechanisms" replaced by "posted-price mechanisms", and "direct mechanisms" by "direct mechanisms with binary allocation outcomes".*

4.3 Lower Bound on Regret

In Amin et al. [2013], they prove a lower bound of $\Omega(T_\gamma)$ for non-myopic agents in the posted-price mechanism setting. We show this lower bound still holds for menu-based mechanisms. Hence, our $O(T_\gamma)$ upper bound is tight up to constant factors. Our proof clearly pinpoints how the regret is measuring a *fundamental trade-off* between allocation inefficiencies for lower-type agents and information rents paid to higher-type agents. We defer the full analysis to Appendix C.

Theorem 4. *For any discount sequence $1 = \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_T > 0$, the regret for indirect menu-based learning mechanisms (and equivalently, direct revelation mechanisms) can be lower bounded by $\Omega(T_\gamma)$.*

5 Conclusion

In this work, we have shown two main results. For our first result, we show that menu-based mechanisms break the $\Omega(\log \log T)$ barrier for the dynamic pricing problem, achieving regret $O(T_\gamma)$ under geometric discounting, matching the lower bound up to constant factors. Our second result is an equivalence between indirect menu mechanisms and direct revelation mechanisms in the single-buyer dynamic pricing setting. We mention two directions for future research. The first is to characterizing when direct mechanisms can be implemented indirectly. While we showed that any IC and PIR direct mechanism can be converted to an indirect mechanism in our setting, the general question remains: what structural properties of a mechanism design problem enable this conversion? The second would be to extend our results to multiple buyers and richer type spaces. In particular, does our equivalence results still hold for multiple buyers or multi-dimension type spaces?

References

Michael Albert, Quinlan Dawkins, Minbiao Han, and Haifeng Xu. Is learning effective in dynamic strategic interactions? evidence from stackelberg games. *arXiv preprint arXiv:2504.15568*, 2025.

Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Learning prices for repeated auctions with strategic buyers. *Advances in neural information processing systems*, 26, 2013.

Santiago R Balseiro, Vahab S Mirrokni, and Renato Paes Leme. Dynamic mechanisms with martingale utilities. *Management Science*, 64(11):5062–5082, 2018.

Dirk Bergemann and Juuso Välimäki. The dynamic pivot mechanism. *Econometrica*, 78(2):771–789, 2010.

Dirk Bergemann and Juuso Välimäki. Dynamic mechanism design: An introduction. *Journal of Economic Literature*, 57(2):235–274, 2019.

Patrick Bolton and Mathias Dewatripont. *Contract theory*. MIT press, 2004.

Tilman Börgers. *An introduction to the theory of mechanism design*. Oxford university press, 2015.

Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Designing menus of contracts efficiently: The power of randomization. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 705–735, 2022.

Vincent Conitzer and Tuomas Sandholm. Computational criticisms of the revelation principle. In *Proceedings of the 5th ACM conference on Electronic commerce*, pages 262–263, 2004.

Alexey Drutsa. Horizon-independent optimal pricing in repeated auctions with truthful and strategic buyers. In *Proceedings of the 26th International Conference on World Wide Web*, pages 33–42, 2017.

Alexey Drutsa. Optimal non-parametric learning in repeated contextual auctions with strategic buyer. In *International Conference on Machine Learning*, pages 2668–2677. PMLR, 2020a.

Alexey Drutsa. Reserve pricing in repeated second-price auctions with strategic bidders. In *International Conference on Machine Learning*, pages 2678–2689. PMLR, 2020b.

Guru Guruganesh, Jon Schneider, Joshua Wang, and Junyao Zhao. The power of menus in contract design. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 818–848, 2023.

Nika Haghtalab, Thodoris Lykouris, Sloan Nietert, and Alexander Wei. Learning in stackelberg games with non-myopic agents. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 917–918, 2022.

Minbiao Han, Michael Albert, and Haifeng Xu. Learning in online principal-agent interactions: The power of menus. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17426–17434, 2024.

Robert Kleinberg and Tom Leighton. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 594–605. IEEE, 2003.

Akshay Krishnamurthy, Thodoris Lykouris, Chara Podimata, and Robert Schapire. Contextual search in the presence of irrational agents. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 910–918, 2021.

Jean-Jacques Laffont and Eric Maskin. *The theory of incentives: An overview*. Université des sciences sociales, Faculté des sciences économiques, 1981.

Allen Liu, Renato Paes Leme, and Jon Schneider. Optimal contextual pricing and extensions. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1059–1078. SIAM, 2021.

Jinyan Liu, Zhiyi Huang, and Xiangning Wang. Learning optimal reserve price against non-myopic bidders. *Advances in Neural Information Processing Systems*, 31, 2018.

Jieming Mao, Renato Leme, and Jon Schneider. Contextual pricing for lipschitz buyers. *Advances in Neural Information Processing Systems*, 31, 2018.

Mehryar Mohri and Andres M Medina. Optimal regret minimization in posted-price auctions with strategic buyers. *Advances in Neural Information Processing Systems*, 27, 2014.

Bernard Salanié. *The economics of contracts: a primer*. MIT press, 2005.

Shiliang Zuo. Corruption-robust lipschitz contextual search. In *International Conference on Algorithmic Learning Theory*, pages 1234–1254. PMLR, 2024.

A Proofs for Section 3

Proof of Lemma 1. We first show that an optimal buyer never switches between L_t and H_t within a phase. Suppose the true value v lies in the current candidate interval I_t, and let $C^* \in \{L_t, H_t\}$ be the myopically preferred contract (breaking ties in favor of H_t).

Since the menu remains fixed throughout the phase, consistently choosing C^* weakly dominates any switching strategy in terms of utility earned during the phase. Moreover, the successor interval corresponding to C^* contains v. In every subsequent phase, the buyer can secure nonnegative utility by consistently choosing a myopically preferred contract, since

$$U_{L_t}(v) = (1 - \delta_t)(v - \underline{v}_t) \geq 0$$

whenever $v \in I_t$. In contrast, switching triggers termination and yields zero continuation utility.

We now prove the lemma for $v \geq m_t + \delta_t/8$; the other case is symmetric. Suppose $v \geq m_t + \delta_t/8$. Then choosing L_t instead of H_t in any round of the phase causes current utility loss

$$U_{H_t}(v) - U_{L_t}(v) = \delta_t(v - m_t) \geq \frac{\delta_t^2}{8}.$$

Let

$$\Gamma_{\text{phase}} := \sum_{i=0}^{D_t - 1} \gamma^i = \frac{1 - \gamma^{D_t}}{1 - \gamma}$$

and

$$\Gamma_{\text{tail}} := \sum_{i=D_t}^{\infty} \gamma^i = \frac{\gamma^{D_t}}{1 - \gamma}.$$

If the buyer falsely chooses L_t consistently throughout the phase, his discounted utility loss during the phase is at least

$$\frac{\delta_t^2}{8}\Gamma_{\text{phase}}.$$

We next upper bound the possible continuation gain from being placed in the wrong branch. Since values, allocations, and payments are bounded, the buyer's per-round continuation utility can change by at most an absolute constant C. Hence the total discounted continuation gain from forcing the wrong branch is at most

$$C\Gamma_{\text{tail}}.$$

By the definition of D_t,

$$\gamma^{D_t} \leq \frac{\rho\delta_t^2}{1 + \rho\delta_t^2}.$$

Therefore

$$\begin{aligned}\Gamma_{\text{tail}} = \frac{\gamma^{D_t}}{1 - \gamma} &\leq \rho\delta_t^2 \frac{1 - \gamma^{D_t}}{1 - \gamma} \\ &= \rho\delta_t^2 \Gamma_{\text{phase}}.\end{aligned}$$

Thus the continuation gain from false branching is at most

$$C\rho\delta_t^2\Gamma_{\text{phase}}.$$

Choosing $\rho \le 1/(8C)$ makes this no larger than the current utility loss from choosing L_t rather than H_t throughout the phase. Hence false branching is not profitable. □

Proof of Lemma 2. Fix a phase with interval $I_t = [\underline{v}_t, \overline{v}_t]$ and width δ_t. Since $v \in [\underline{v}_t, \overline{v}_t]$, the full-information benchmark revenue in any round is at most

$$v \le \underline{v}_t + \delta_t.$$

The low menu option generates revenue

$$(1-\delta_t)\underline{v}_t \ge \underline{v}_t - \delta_t,$$

and the high menu option generates revenue

$$\underline{v}_t + \delta_t^2/2 \ge \underline{v}_t.$$

Thus every nonzero menu choice generates revenue at least $\underline{v}_t - \delta_t$, and the seller's per-round regret is $O(\delta_t)$. Since the actual phase duration is $\overline{D}_t$, the total regret incurred in the phase is

$$O(\delta_t \overline{D}_t).$$

□

Proof of Theorem 1. Index the phases by $k = 0, 1, 2, \ldots$, and let δ_k denote the candidate-interval width at the beginning of phase k. Every completed phase shrinks the interval width by a factor at most

$$\beta := \frac{5}{8}.$$

There is at most one truncated phase, namely the final phase. Consequently, for every phase that is actually reached,

$$\delta_k \le \beta^k.$$

We first bound the actual duration of a phase. Write

$$A_k := \log\left(\frac{1+\rho\delta_k^2}{\rho\delta_k^2}\right) = \log\left(1 + \frac{1}{\rho\delta_k^2}\right).$$

Since $\rho > 0$ is an absolute constant,

$$A_k = O\left(1 + \log\frac{1}{\delta_k}\right).$$

By definition,

$$D_k \le 1 + \frac{A_k}{\log(1/\gamma)}, \qquad \overline{D}_k = \min\{D_k, T - t_k + 1\},$$

where t_k is the first round of phase k. Hence

$$\overline{D}_k = O\left(\left(1 + \log\frac{1}{\delta_k}\right)\min\left\{T, \frac{1}{\log(1/\gamma)}\right\}\right).$$

Using

$$\log(1/\gamma) \ge 1 - \gamma,$$

we obtain

$$\min\left\{T, \frac{1}{\log(1/\gamma)}\right\} \le \min\left\{T, \frac{1}{1-\gamma}\right\}.$$

We next relate the latter quantity to T_γ. Let $x := 1 - \gamma$. Since

$$\gamma^T = (1-x)^T \le e^{-xT},$$

we have

$$T_\gamma = \frac{1-\gamma^T}{1-\gamma} \geq \frac{1-e^{-xT}}{x}.$$

For every $y \geq 0$,

$$1 - e^{-y} \geq (1 - e^{-1}) \min\{y, 1\}.$$

Applying this with $y = xT$ gives

$$T_\gamma \geq (1 - e^{-1}) \min\left\{T, \frac{1}{1-\gamma}\right\}.$$

Therefore

$$\overline{D}_k = O\left(T_\gamma \left(1 + \log \frac{1}{\delta_k}\right)\right).$$

By Lemma 2, the regret in phase k is thus

$$O\left(T_\gamma\, \delta_k \left(1 + \log \frac{1}{\delta_k}\right)\right).$$

The function

$$x \mapsto x \left(1 + \log \frac{1}{x}\right)$$

is increasing on $(0, 1]$. Since $\delta_k \leq \beta^k$,

$$\delta_k \left(1 + \log \frac{1}{\delta_k}\right) \leq \beta^k \left(1 + k \log \frac{1}{\beta}\right).$$

Summing over all phases, including the possibly truncated final phase,

$$\begin{aligned} \mathsf{Reg} &\leq O(T_\gamma) \sum_{k \geq 0} \beta^k \left(1 + k \log \frac{1}{\beta}\right) \\ &= O(T_\gamma), \end{aligned}$$

because both $\sum_{k\geq 0} \beta^k$ and $\sum_{k \geq 0} k\beta^k$ are finite. Thus the mechanism achieves regret $O(T_\gamma)$ for every finite horizon T. □

B Proof of Section 4

Proof of Theorem 2. Fix any true type v, any potential misreport $\hat{v}$, and any round $t \in \{1, \ldots, T\}$. We must show:

$$U_{1:T}(v; v) \geq U_{1:t}(\hat{v}; v).$$

Consider the indirect menu mechanism. At each round τ and history h_τ, the principal posts a menu $M_\tau(h_\tau) = \{(a, p_\tau(a; h_\tau)) : a \in [0, 1]\}$.

By construction, $(a_\tau(v), p_\tau(v))_{\tau=1:T}$ is the trajectory generated when a type-v agent plays their best response in the indirect mechanism. This means at each round τ, given the history $h_\tau(v)$ generated by type v's play through round $\tau - 1$, the agent selects allocation $a_\tau(v)$ from menu $M_\tau(h_\tau(v))$ and pays $p_\tau(v) = p_\tau(a_\tau(v); h_\tau(v))$.

Now consider an alternative strategy in the indirect mechanism where the type-v agent:

1. Mimics type $\hat{v}$'s behavior for rounds 1 through t: at each round $\tau \leq t$, follows history $h_\tau(\hat{v})$ and selects allocation $a_\tau(\hat{v})$ from menu $M_\tau(h_\tau(\hat{v}))$, paying $p_\tau(\hat{v})$.
2. Drops out at round $t + 1$ by selecting the outside option $(0, 0)$ in all subsequent rounds.

This deviation strategy yields utility exactly $U_{1:t}(\hat{v}; v)$ to the type-v agent.

Since $(a_\tau(v), p_\tau(v))_{\tau=1:T}$ is generated by the agent's best response in the indirect mechanism, it must weakly dominate all alternative strategies, including the deviation described above:

$$U_{1:T}(v; v) \geq U_{1:t}(\hat{v}; v).$$

This holds for all v, all $\hat{v}$, and all $t \in \{1, \ldots, T\}$, establishing that the direct mechanism satisfies IC and PIR. The argument applies to any selection of best responses $H(v)$, so this direction does not depend on how the buyer breaks ties. □

Proof of Theorem 3. The bulk of the proof is showing the mechanism still satisfies IC and PIR after each adjustment step (Parts (i) and (ii)) and that no adjustment is needed in the final period (Part (iii)). Then, in the final indirect mechanism, the prescribed path is a best response for each type, and by the tie-breaking convention the buyer's selected best response yields at least the revenue of this path (Part (iv)).

Consider a single adjustment at round $t < T$ where we set $p_t(v) \leftarrow p_{\min}(a_t(v))$ and $p_{t+1}(v) \leftarrow p_{t+1}(v) + \frac{\gamma_t}{\gamma_{t+1}}(p_t^{\text{old}}(v) - p_{\min}(a_t(v)))$, where $p_t^{\text{old}}(v)$ denotes the value before adjustment.

Recall that the utility for a type-v' agent reporting $\hat{v}$ and participating through round τ is:

$$U_{1:\tau}(\hat{v}; v') = \sum_{s=1}^{\tau} \left[\gamma_s a_s(\hat{v}) \cdot v' - \gamma_s p_s(\hat{v})\right].$$

We must verify that the adjusted mechanism still satisfies IC and PIR: for all types v', all reports $\hat{v}$, and all rounds τ,

$$U_{1:T}(v'; v') \geq U_{1:\tau}(\hat{v}; v').$$

Part (i). Utility for truthful reporting for type v is unchanged. For a type-v agent reporting v and participating through all T rounds, the utility change from the adjustment is:

$$\begin{aligned}
\Delta U_{1:T}(v; v) &= \left[-\gamma_t p_{\min}(a_t(v)) - \gamma_{t+1}\left(p_{t+1}(v) + \frac{\gamma_t}{\gamma_{t+1}}(p_t^{\text{old}}(v) - p_{\min}(a_t(v)))\right)\right] \\
&\quad - \left[-\gamma_t p_t^{\text{old}}(v) - \gamma_{t+1} p_{t+1}(v)\right] \\
&= -\gamma_t p_{\min}(a_t(v)) - \gamma_t(p_t^{\text{old}}(v) - p_{\min}(a_t(v))) + \gamma_t p_t^{\text{old}}(v) \\
&= 0.
\end{aligned}$$

Therefore, $U_{1:T}(v; v)$ remains unchanged. The allocation benefits $\sum_{s=1}^{T} \gamma_s a_s(v) \cdot v$ are unaffected, and the payment stream is adjusted to preserve its present value.

Part (ii). Deviations remain suboptimal.

Next, We need to verify that for any type v' and any dropout round τ, the constraint $U_{1:T}(v'; v') \geq U_{1:\tau}(v; v')$ continues to hold after adjusting v.

We need to check how $U_{1:\tau}(v; v')$ changes. This depends on whether $\tau < t$, $\tau = t$, or $\tau \geq t + 1$.

- If $\tau < t$. Since we are adjusting price in round t and the next round, dropping out before round t have the exact same utility as before the adjustment. If the original mechanism were IC and PIR, dropping out before round t would not benefit the agent.
- If $\tau = t$. We must show that a type-v' agent cannot benefit from reporting v and dropping out at round t after the price adjustment. Since $S(a)$ may be infinite, the infimum defining $p_{\min}(a)$ in Algorithm 2 may not be attained; to emphasize this, in this step we write

  $$p_{\inf}(a) := \inf_{w \in S(a)} p_t(w).$$

  Fix any $\epsilon > 0$. By definition of infimum, there exists $v^\sharp \in S(a)$ such that

  $$p_t(v^\sharp) \leq p_{\inf}(a) + \epsilon.$$

Consider a type-v' agent who reports v and drops out at round t after the adjustment. His utility from this deviation is:

$$U^{\text{new}}_{1:t}(v; v') = U_{1:t-1}(v; v') + \gamma_t[a_t(v) \cdot v' - p_{\inf}(a_t(v))].$$

Now, in the *original* mechanism (before adjustment), the agent could have reported $v^\sharp$ and dropped out at round t, obtaining:

$$U^{\text{orig}}_{1:t}(v^\sharp; v') = U_{1:t-1}(v^\sharp; v') + \gamma_t[a_t(v^\sharp) \cdot v' - p_t(v^\sharp)].$$

Since $v^\sharp \in S(a)$, we have $a_t(v^\sharp) = a_t(v) = a$, and since $v, v^\sharp$ both reach the same node (i.e., have the same interaction history up to round t), we have $U_{1:t-1}(v; v') = U_{1:t-1}(v^\sharp; v')$. Therefore:

$$U^{\text{new}}_{1:t}(v; v') = U^{\text{orig}}_{1:t}(v^\sharp; v') + \gamma_t[p_t(v^\sharp) - p_{\inf}(a)] \le U^{\text{orig}}_{1:t}(v^\sharp; v') + \gamma_t\epsilon.$$

By IC and PIR of the original mechanism:

$$U_{1:T}(v'; v') \ge U^{\text{orig}}_{1:t}(v^\sharp; v').$$

Combining:

$$U_{1:T}(v'; v') \ge U^{\text{new}}_{1:t}(v; v') - \gamma_t\epsilon.$$

Since this holds for all $\epsilon > 0$, taking $\epsilon \to 0$ yields:

$$U_{1:T}(v'; v') \ge U^{\text{new}}_{1:t}(v; v'),$$

as desired.

- If $\tau > t$. By our adjustment step, the utility for an agent remains unchanged after proceeding beyond round $t+1$, since the adjustment had accounted for discounting. Hence any deviation still would not benefit the agent.

Part (iii). No adjustment is needed in the final period. Fix a node at depth T and two reports v, v' in its candidate set with $a_T(v) = a_T(v') = a$. Since v and v' reach the same node, after the equalization steps in rounds $1, \dots, T-1$ we have $a_s(v) = a_s(v')$ and $p_s(v) = p_s(v')$ for all $s < T$. By Parts (i) and (ii), the mechanism obtained after these steps is still IC, so $U_{1:T}(v; v) \ge U_{1:T}(v'; v)$, which reduces to $p_T(v) \le p_T(v')$. Exchanging the roles of v and v' gives $p_T(v) = p_T(v')$. Hence all reports in $S(a)$ already pay a common price in round T, this price equals $p_{\min}(a)$ for $a \neq 0$, and no deferral is needed.

Part (iv). Best response and revenue. After all adjustments, all reports in the candidate set of a node at depth t share the same allocations and payments in rounds $1, \dots, t-1$, and the menu at that node offers exactly the contract $(a_t(\hat{v}), p_t(\hat{v}))$ of each report $\hat{v}$ in the candidate set. Hence any strategy of a type-v buyer in $\mathcal{I}$ traces a path in the tree and yields utility $U_{1:T}(\hat{v}; v)$ for some report $\hat{v}$ (or $U_{1:t}(\hat{v}; v)$ if he drops out after round t). By Parts (i)–(iii), the final adjusted direct mechanism is IC and PIR, so following the path $(a_t(v))_{t=1}^T$, which yields $U_{1:T}(v; v)$, is a best response for type v. Along this path the buyer pays the adjusted payments $p_t(v)$. Each adjustment step moves an amount $\Delta = p^{\text{old}}_t(v) - p_{\min}(a_t(v)) \ge 0$ from round t to round $t+1$ and scales it by $\gamma_t/\gamma_{t+1} \ge 1$, so it changes the undiscounted revenue from report v by $(\gamma_t/\gamma_{t+1} - 1)\Delta \ge 0$. Hence the revenue along this path is at least $\mathsf{Rev}(v; \mathcal{D})$. Finally, under the tie-breaking convention of Section 2, the buyer selects, among all his best responses, one maximizing the seller's revenue, and therefore $\mathsf{Rev}(v; \mathcal{I}) \ge \mathsf{Rev}(v; \mathcal{D})$.

□

C Proof of Lower Bound

Consider the case with only two types $0 < \underline{v} < \overline{v}$. In general, for a truthful direct mechanism, the following principles hold:

- **Lower type:** Receives an inefficient allocation (the item is withheld in certain rounds), but the seller need not pay information rent to the buyer.

- **Higher type:** Receives efficient allocation (the item is allocated in every round), but the buyer extracts information rent from the seller.

The above principal generally holds for usual adverse selection problems where the agent's type is one-dimensional, and a more detailed discussion can be found in Salanié [2005]. To understand why, consider the incentive compatibility constraint: the higher-type agent $\overline{v}$ must weakly prefer reporting truthfully over misreporting as type $\underline{v}$. In a first-best scenario, the item would always be allocated and the agent always charged their true valuations. Such is not possible in the second-best scenario when the agent's type is unknown. The seller must adopt the following to ensure incentive compatibility:

1. *Withhold allocation to the lower type report*: By withholding the item from type $\underline{v}$ in certain rounds (setting $a_t(\underline{v}) < 1$ for some t), the seller makes this allocation less attractive to the higher type. If the item is always allocated at the lower type, he cannot be made to pay more than his true valuation, and hence the higher type should deviate to the lower type and be sure to enjoy the surplus.
2. *Leave information rent to the higher type*: The seller can give the higher type a utility premium beyond his participation constraint. If the higher type is left with no information rent, he would of course again deviate to reporting the lower type. The lower type would not have this issue, since deviating upward would generally net him no gain.

We demonstrate that the regret lower bound $\Omega(T_\gamma)$ measures a fundamental tradeoff between the two quantities.

We state the lower bound when there are two possible types. The result of course extends to richer type spaces (e.g., $v \in [0, 1]$), since the mechanism can only have more restrictions.

Theorem 5. *Suppose there are only two possible types $\underline{v}$ and $\overline{v}$. The seller's strategic regret satisfies*

$$\mathrm{Reg} \geq T_\gamma \cdot \frac{\underline{v}(\overline{v} - \underline{v})}{\overline{v}} = \Omega(T_\gamma).$$

Proof. We break the proof into several steps.

Step 1: Bounding payments from the lower type. Since the lower-type agent participates voluntarily (periodic individual rationality), we have

$$\sum_{t=1}^{T} \gamma_t [a_t(\underline{v})\underline{v} - p_t(\underline{v})] \geq 0.$$

Rearranging yields

$$\sum_{t=1}^{T} \gamma_t p_t(\underline{v}) \leq \sum_{t=1}^{T} \gamma_t a_t(\underline{v})\underline{v}. \tag{2}$$

Step 2: Information rent for the higher type. The utility that a higher-type agent obtains by misreporting to $\underline{v}$ (and never dropping out) is

$$\sum_{t=1}^{T} \gamma_t [a_t(\underline{v})\overline{v} - p_t(\underline{v})] \geq \sum_{t=1}^{T} \gamma_t [a_t(\underline{v})\overline{v} - a_t(\underline{v})\underline{v}], \tag{3}$$

where the inequality follows from (2).

By incentive compatibility, the principal must provide the higher type with information rent at least equal to this deviation utility. Define

$$\mathrm{Reg}(\overline{v}) := \sum_{t=1}^{T} \gamma_t \cdot [a_t(\underline{v}) \cdot (\overline{v} - \underline{v})],$$

which lower bounds the regret due to information rent paid to the higher type.

Step 3: Allocation inefficiency for the lower type. The regret from allocating inefficiently to the lower type is at least

$$\begin{aligned} T\underline{v} - \sum_{t=1}^{T} p_t(\underline{v}) &\geq T\underline{v} - \sum_{t=1}^{T} a_t(\underline{v}) \cdot \underline{v} \\ &= \sum_{t=1}^{T} (1 - a_t(\underline{v})) \cdot \underline{v} \\ &\geq \sum_{t=1}^{T} \gamma_t \cdot (1 - a_t(\underline{v})) \cdot \underline{v}. \end{aligned}$$

The first line above is by Lemma 3 below. Define

$$\mathrm{Reg}(\underline{v}) := \sum_{t=1}^{T} \gamma_t \cdot (1 - a_t(\underline{v})) \cdot \underline{v}.$$

which lower bounds the allocation inefficiency for the lower type.

Step 4: Combining the bounds. The total regret satisfies

$$\mathrm{Reg} \geq \max\left(\mathrm{Reg}(\underline{v}), \mathrm{Reg}(\overline{v})\right).$$

Observe that

$$\frac{1}{\underline{v}}\mathrm{Reg}(\underline{v}) + \frac{1}{\overline{v} - \underline{v}}\mathrm{Reg}(\overline{v}) = T_\gamma.$$

By the AM-GM inequality applied to the weighted harmonic mean,

$$\mathrm{Reg} \geq T_\gamma \cdot \frac{\underline{v}(\overline{v} - \underline{v})}{\overline{v}} = \Omega(T_\gamma).$$

This completes the proof. □

In step 3 above we used the below lemma. which shows that the non-discounted payment from the buyer must be upper bounded by the non-discounted utility that the buyer receives.

Lemma 3. *Fix any type v. We have*

$$\sum_{t=1}^{T} p_t(v) \leq \sum_{t=1}^{T} a_t(v) \cdot v$$

Proof. By PIR, we have a sequence of participation constraints. The inequality with the summand $t = s$ to T means that the agent prefers continuing participating in the contracts rather than dropping out.

$$\sum_{t=s}^{T} \gamma_t \cdot [a_t(v) \cdot v - p_t(v)] \geq 0$$

Since the sequence γ_t is decreasing, we can find a sequence of appropriate positive constants $\{c_t\} > 0$ such that for any $s \in [1, T]$,

$$\gamma_s \sum_{t=1}^{s} c_t = 1.$$

Then multiplying the participation constraint in round s by c_s and summing the sequence of inequality together gives the desired bound. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The main claims made in the abstract and introduction accurately reflection the contribution and scope.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Limitations and futures directions are discussed in the Conclusion section.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: We provide full set of assumptions and complete and correct proof.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: This is a theoretical paper and we do not have experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: This does not apply to this paper.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: This does not apply to this paper.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: This does not apply to this paper.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: This does not apply to this paper.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We conform to the code of ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: We do not foresee societal impacts for this theory paper.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: This does not apply to the current paper.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: This does not apply to this paper.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: This does not apply to this paper.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This item does not apply to this paper.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This does not apply to this paper.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: We did not use LLM beyond the scope of writing, editing, or formatting purposes, and use of LLM does not impact the core methodology.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.